PROCEEDINGS OF THE VIDYA COMPUTER APPLICATIONS DEPARTMENTAL

SEMINAR

(VCADS - 2017)

25-26 April 2017

DEPARTMENT OF COMPUTER APPLICATIONS
VIDYA ACADEMY OF SCIENCE AND TECHNOLOGY
THALAKKOTUKARA P.O | THRISSUR | KERALA 680501

Proceedings of the

Vidya CA Departmental Seminar (VCADS - 2017)

25 - 26 April 2017

Editors

Dr V N Krishnachandran
Professor of Computer Applications

Reji C Joy
Associate Professor of Computer Applications

Sajay K R
Associate Professor of Computer Applications

Siji K B
Assistant Professor of Computer Applications

Department of Computer Applications
Vidya Academy of Science & Technology (VAST)
Thalakkottukara, Thrissur – 680501

Mission

Progress through Education

Vision

To seek, strive for and scale greater heights of quality education

Vidya Academy of Science & Technology

Contents

Preface

This volume contains the proceedings of the two-day departmental seminar organised by the Computer Applications Department of Vidya Academy of Science & Technology during 25 – 26 April 2017. The seminar was the culmination of a coursework (with course code MCA 2010 506(P) Seminar) to be completed by the MCA students of Calicut University during the Fifth Semester of the MCA programme. As part of the course, each student has to prepare and present a paper on any topic in the field of computer science.

In the previous years the students completed this coursework in a certain format. The students chose the topics for study and presentation. A Department Committee scrutinised the topics and approved the topics with modifications if necessary. In the evaluation of students' work, the focus was only on the presentation and communication skills.

As part of the decennial celebrations of the Department (the Computer Applications Department of the College was established in the academic year 2006 – 07), it was decided to introduce some steps to enhance the learning experience of the students. And as part of this, it was decided to offer the seminar course in a new format. In the new format, it was the teachers who identified the areas in which the students are to work and the teachers provided the students with some initial learning materials in the form of papers. After the initial reading of these materials, the students had to search for additional reading materials themselves. The students were required to study the papers and present a "study paper" in a Departmental Seminar. The papers collected in this volume are the study papers prepared by the students and presented in the two-day seminar. They are indicative of the level of achievements of the students.

As part of the learning process, the students were also required to present the paper in the IEEE conference paper format. To facilitate this, the students were given a basic introduction to the LaTeX software and the `IEEEtran` document style.

The emphasis in the whole exercise was to give the students a hands on experience in preparing a conference/seminar paper and not on making the students learn a new topic or subject in depth. The expected learning outcomes include:

- understanding of the structure of a research paper,
- awareness about the process of literature survey,
- basic knowledge about the accurate preparation of bibliography and their citations in the paper,
- exposure to the IEEE format for the preparation of conference/journal papers,
- introduction to the concepts of "Abstracts", "Keywords", and the like,
- experience in applying these concepts by actually preparing a paper, and

- methodology of presenting a multi-author paper in a seminar/conference.

The articles compiled in this Proceedings are not even moderately edited. The editors have only ensured that the basic learning outcomes outlined above have been met. However, the editors have tried to ensure that the titles of chapters, sections, etc., the abstract, figure and table captions, and the like are as per IEEE guidelines. The references have not been checked for accuracy and completion. The papers have not been edited for grammar, punctuation, spelling or style.[1]

The present work is only a record of the activities of the course referred to above and it is prepared only for private circulation. To the best of our understanding the authors of the papers have given proper attribution to ideas and material presented in the papers. If there are no attributions or improper attributions, it was unintentional. Hence the contents have not been subjected to plagiarism tests.

It is believed that the teachers as well the students enjoyed very much the new format of the seminar course. There are still much scope for improvement. It is our hope that the future batches of students will have a stronger and wider learning experience from a similar seminar courses.

April 2017 Editors

[1] For different models of editing, see, for example "IEEE Editorial Style manual", [Online] Available: `https://www.ieee.org/documents/style_manual.pdf` (April 2017).

Introduction to Natural Language Processing

Alpha John T J, Muhammed Ishaq, Senil Seby K
Department of Computer Applications
Vidya Academy of Science & Technology
Thrissur 680501

Aparna S Balan
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur 680501

***Abstract*—Natural language processing is a field that covers computer understanding and manipulation of human languages. The field of study that focus on simplify the interaction between human languages and computer is called natural language processing or NLP. This seminar paper is an attempt to analyze the research studies in NLP techniques for south Indian languages and also it tells about the tools and concept of studies related to it.**

***Index Terms*—Natural language processing, Machine transaction, Information Extraction, Summarization, Sentiment analysis, Question answering system, Text categorization, Prediction system/Decission making system, Text mining, POS tagging (Parts of Speech)**

I. Introduction

NLP is a way for computers to analyze, understand, and derive meaning from human language in a smart and useful way. By utilizing NLP, developers can organize and structure knowledge to perform tasks such as automatic summarization, translation, named entity recognition, relationship extraction, sentiment analysis, speech recognition, and topic segmentation. This paper aims to address about the concept of natural language processing and the analysis of how effective it in case of south Indian languages.Studies on south Indian languages are done which include Tamil, Telugu and Malayalam. As we are using regional languages to communicate with the system ,it increases the bond between the user and the computer. They are mainly used for popular languages as it is very difficult task to accomplish. machine translation offers an even more scalable alternative to harmonizing the world's information. Summarizing branch of NLP will become increasingly useful as a valuable marketing asset. Typical text mining tasks include text categorization, text clustering, concept/entity extraction, production of granular taxonomies, sentiment analysis, document summarization, and entity relation modeling. To understand human language is to understand not only the words, but the concepts and how theyre linked together to create meaning. As this is a complex task to accomplish these are mainly used for popular languages.

II. Machine Transaction

As the world's information is online, the task of making that data accessible becomes increasingly important. The challenge of making the world's information accessible to everyone, across language barriers, has simply outgrown the capacity for human translation. Innovative companies like Duolingo are looking to recruit large amounts of people to contribute, by coinciding translation efforts with learning a new language. But machine translation offers an even more scalable alternative to harmonizing the world's information. Google is a company at the forefront of machine translation, using a proprietary statistical engine for its Google translate service. The challenge with machine translation technologies is not in translating words, but in preserving the meaning of sentences, a complex technological issue that is at the heart of NLP.

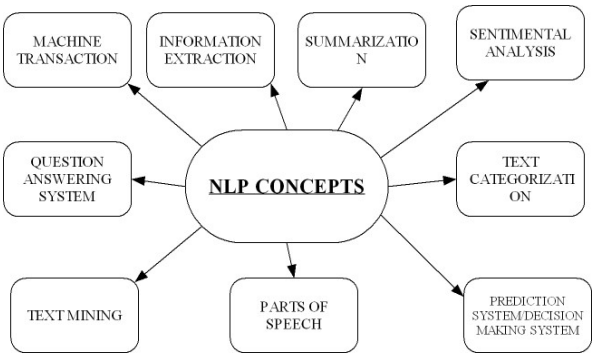


Fig. 1. NLP Concepts

A. Malayalam machine translation using hybrid approach

Machine translation is one of the challenging field in NLP. The goal of machine translation is to automatically translate a sentence from one language to another. In this paper we can introduce an hybrid approach for translating Malayalam sentences to corresponding English sentence. The hybrid approach combines both example based machine translation technique and transfer approaches to deliver greater quality and functionality. Machine translation requires a deep and rich understanding of the source language and input text.[1]

B. Text to speech synthesis system for English to Malayalam translation

Speech recognition and Speech synthesis are the two emerging technologies in the communication field. Speech recognition is a system which generates text for the given speech, while a speech synthesizer is a system that should be able to read any text aloud. Research is conducting all over the world, to develop an efficient text to speech synthesis (TTS) system for minority languages. India has the largest democracy system in the world, also known as land of unity in diversity and has more than 22 official languages. It becomes difficult to understand the English and the other European languages to common people. This works aims to help people to translate English text to their own language and implement a TTS for the minority language, Malayalam. It is achieved by combining both Machine Translation and TTS. When an English text is given, it is translated to Malayalam with the help of a parser, using grammatical rules, applying morphology and a bilingual dictionary. From each of the translated Malayalam text, syllables are separated. A good number of syllables are recorded and stored in the syllable corpus. Syllables are concatenated to generate a synthesized Malayalam speech. Machine translation of English to Malayalam text is tested and achieved 73 percentage accuracy. For the TTS system, accuracy is verified by checking the naturalness and intelligibility. 87 percentages of the sentences are uttered correctly.[2]

C. Conversion of Malayalam text to Indian sign language using synthetic animation

This paper presents a machine translation system for translating Malayalam text to Indian Sign Language (ISL) using synthetic animation approach. The proposed system makes use of HamNoSys structure as an intermediate representation for signs. The system accepts Malayalam input either as single words or as multiple words and the output is delivered as 3D character animation. This system also provides the facility to input new words to the database through an interactive sign editor that transcribes signs into HamNoSys structure. It has been implemented as a tutoring system for sign language and hence promote sign language education among common people of Kerala.[3]

D. SVM based feature set analysis in dynamic malayalam handwritten character recognition

Dynamic or Online handwritten character recognition is a challenging field in Human Computer Interfaces. The classification success rate of current techniques decreases when the dataset involves the similarity and complexity in stroke styles, number of strokes and stroke characteristics variations. Malayalam is a complex south Indian language spoken about 35 million people especially in Kerala and Lakshadweep islands. In this paper, a classification scheme based on support vector machines (SVM) is proposed to improve the accuracy in classification and recognition of online malayalam handwritten characters. SVM Classifier is a popular one in academy as well as in industry. This Classifiers are more suitable in a real world applicative problem, if we have major concern on the speed of recognition per character. The contribution of various features towards the accuracy in recognition is analyzed. Performance for different kernels of SVM are also studied. A graphical user interface has developed for reading and displaying the character. Different writing styles are taken for each of the 44 alphabets. Various features are extracted and used for classification after the preprocessing of input data samples. Feature Selection is carried out by choosing of different combinations of extracted features versus accuracy. Highest recognition accuracy of 97% is obtained for the best selected features in SVM with polynomial kernel. Recognition speed of a single stroke is obtained 0.52 secs.[4]

E. Online handwritten malayalam character recognition using LIBSVM in matlab

This paper proposes an experimental technique to find the Malayalam Handwritten Character Recognition using support vector machines (SVM). Malayalam is a south indian language originated from brahmi script spoken about 35 million people especially in Kerala and Lakshadweep islands. It has the largest number of alphabets of 128 characters among indian languages. Online Malayalam Handwritten Character Recognition has been studied from the past several years, demands to replace a large keyboard because of huge alphabet size for texting applications and easy convenience of writing in our own style, signature verifications etc. Real time (x,y) coordinates per stroke are acquired and preprocessed. Directional and Curvature features are extracted and trained in LIBSVM, a tool for SVM Classifiers. Testing alphabet is given online to the trained SVM network and the recognized label is displayed in Notepad. Experiments are done for handwritten basic vowel alphabets (8) in Malayalam. Recognition speed of a single stroke is attained 0.52 secs.[5]

F. Zone based handwritten Kannada character recognition using crack code and SVM

Efficient methods have been proposed in the literature for recognition of printed characters of Indian scripts. However, much less attention has been received for handwritten characters of Indian scripts with only a few work available in the literature. In this paper, a novel zone based method has been presented for recognition of handwritten characters written in Kannada language, a major South Indian language. The normalized character image is divided into 64 zones each of size 8x8 pixels. For each zone, from left to right and from top to bottom, the crack code, representing the line between the object pixel and the background (the crack), is generated by traversing it in anticlockwise direction. A feature vector of size 512 is obtained for each character. A multi-class SVM is used for the classification purpose. Experiments are performed on handwritten Kannada character images consisting of 24500 images with 500 samples of each character. Five-fold cross validation is used for result computation that yielded 87.24% recognition accuracy.[6]

G. Domain Specific Sentence Level Mood Extraction from Malayalam Text

Natural Language Processing (NLP) is a field which studies the interactions between computers and natural languages. NLP is used to enable computers to attain the capability of manipulating natural languages with a level of expertise equivalent to humans. There exists a wide range of applications for NLP, of which sentiment analysis(SA) plays a major role. In sentimental analysis, the emotional polarity of a given text is analysed, and classified as positive, negative or neutral. A more difficult task is to refine the classification into different moods such as happy, sad, angry etc. Analysing a natural language for mood extraction is not at all an easy task for a computer. Even after achieving capabilities of massive amount of computation within a matter of seconds, understanding the sentiments embodied in phrases and sentences of textual information remains one of the toughest tasks for a computer till date. This paper focuses on tagging the appropriate mood in Malayalam text. Tagging is used to specify whether a sentence indicates a sad, happy or angry mood of the person involved or if the sentence contains just facts, devoid of emotions. This research work is heavily dependent on the language since the structures vary from language to language. Mood extraction and tagging has been successfully implemented for English and other European languages. For the south Indian language Malayalam, no significant work has yet been done on mood extraction. We will be focusing on domain-specific sentence-level mood extraction from Malayalam text. The extraction process involves parts-of-speech tagging of the input sentence, extracting the patterns from the input sentence which will contribute to the mood of the sentence, such as the adjective, adverb etc., and finally scoring the sentence on an emotive scale by calculating the semantic orientation of the sentence using the extracted patterns. Sentiment classification is becoming a promising topic with the rise of social media such as blogs, social networking sites, where people express their views on various topics. Mood extraction enables computers to automate the activities performed by human for making decisions based on the moods of opinions expressed in Malayalam text.[7]

H. Shallow parser for Malayalam language using finite state cascades

Various methods have been proposed for chunking sentences in agglutinative languages. For Malayalam a South Indian language, chunking methods proposed are mainly statistical. This paper describes a chunking method for Malayalam sentences based on morpheme based augmented transition network. For the trial set of sentences the system works with good accuracy with the set of chunk rules proposed. The chunking system has good potential for use as a full fledged parser for Malayalam language.[8]

III. Information Extraction

Many important decisions in financial markets are increasingly moving away from human oversight and control. Algorithmic trading is becoming more popular, a form of financial investing that is entirely controlled by technology. But many of these financial decisions are impacted by news, by journalism which is still presented predominantly in English. A major task, then, of NLP has become taking these plain text announcements, and extracting the pertinent info in a format that can be factored into algorithmic trading decisions. For example, news of a merger between companies can have a big impact on trading decisions, and the speed at which the particulars of the merger, players, prices, who acquires who, can be incorporated into a trading algorithm can have profit implications in the millions of dollars.

A. Suffix stripping algorithm for Kannada information retrieval

Due to the explosion of usage of the internet and websites a huge amount of data on the web is available in languages other than English. Hence, it is important to develop Information Retrieval (IR) tools for other languages too for web searches. This is as challenging as developing an IR tool for English since each language has unique characteristics of its own. In the development of an IR tool for a particular language we need to consider the specifics of that language. Due to unique characteristics of each language, many efficient algorithms developed for the IR in English language cannot be used directly. Here we consider a south Indian language Kannada and propose a suffix stripping algorithm. This algorithm is for the Kannada text available on-line in unicode. It is a rule based approach that strips fourteen different major classifications of suffixes (pratyaya in Kannada) and some sub classes. It also covers suffixes associated with nouns, verbs, articles, adjectives and stop words. This algorithm will be very useful to rank Kannada documents(that are represented by a bag-of-words) based on relevance upon a query in web searches. It can also be used in Kannada (i) text extraction, (ii) natural language processing tools and (iii)speech recognition engines. We have implemented this suffix stripping stemming algorithm and have evaluated it using Kannada documents from Kendasampige a web based magazine with and without our stemming algorithm. We used several metrics for the evaluation. Our results indicate that the recall factor is much better after stemming. This promising preliminary results imply the applicability of this algorithm to the above mentioned applications.[9]

B. Haar features based handwritten character recognition system for Tulu script

Automatic recognition of handwritten characters from scanned images helps to convert characters in an image into convenient editable and readable form. Tulu is a south Indian Dravidian language with rich set of handwritten patterns. This paper presents an approach to recognize the Tulu script using automatic character recognition mechanism. The recognition of handwritten Tulu characters is based on the AdaBoost algorithm using Haar features. Finally, recognized characters are mapped into an equivalent editable document of Kannada

characters. Hence, make it to readable for the next generation by digital technology.[10]

C. Information extraction by an abstractive text summarization for an Indian regional language

The Internet provides many sources of different opinions, expressed through user reviews of products, blogs, and forum discussions. Systems which could automatically summarize these opinions would be immensely useful for those who wish to use this information to make decisions. The previous work in automatic summarization has completely focused on extractive summarization, in which key sentences are identified from the source text and extracted to form the output. An alternative solution is abstractive summarization in which the information from the source text is first extracted into the form of abstract data which is then post processed to infer the most important message from the original text. This work is built upon past work of extractive summarization methods to create abstractive summaries by creating new sentences in it. This paper conveys the methodology for the abstractive summarization process and its evaluation considering Telugu, a south Indian regional language, as the language of study.

IV. SUMMARIZATION

Information overload is a real phenomenon in our digital age, and already our access to knowledge and information far exceeds our capacity to understand it. This is a trend that shows no sign of slowing down, and so an ability to summarize the meaning of documents and information is becoming increasingly important. This is important not just in allowing us the ability to recognize and absorb the pertinent information from vast amounts of data. Another desired outcome is to understand deeper emotional meanings, for example, based on aggregated data from social media, can a company determine the general sentiment for its latest product offering? This branch of NLP will become increasingly useful as a valuable marketing asset.

A. Summarization of customer reviews for a product on a website using natural language processing

In the recent past, e-commerce sites have made rapid growth. There are thousands of products and various websites sell these products. Massive growth in the number of reviews and their availability along with the advent of opinion-rich review forums for the products sold online, choosing the right one from a large number of products has become difficult for the users. HELPME-BUY APP is an android application that assists buyers in online shopping. It is imminent for buyers to verify for genuineness and quality of products. What better way is there than to ask people who have already bought the product? This is when customer reviews come into picture. The major hitch here is popular products have thousands of reviews-we do not have the time or patience to read all thousands of them. Hence, our application eases this task by analyzing and summarizing all reviews which will help the user decide what other buyers have experienced on buying this product. We carry out this process by a number of modules that include feature extraction and opinion extraction which improves the process of analysis and helps in the formation of an efficient summary.[11]

B. Document Summarization in Malayalam with sentence framing

Document Summarization is a technique of conveying important information in a given document. It is one of the most important chores of Natural Language Processing as the summary produced is helpful for information retrieval systems, question answering systems, medical domain and news domain etc. Most of the summarization works in Indian languages are of extractive nature and not much work is oriented towards the abstractive summarization approaches in Indian languages as they need more linguistic processing. As Indian languages belong to several language families and since they are morphologically rich and agglutinative in nature, a lot of challenges are faced while doing abstractive summarization because it requires natural language generation techniques. The proposed work is an approach for abstractive text summarization that will accept single document as input in Malayalam, processes the input by building a suitable semantic representation and then use sentence framing techniques to generate the final summary. The entire framework is composed of eight modules that mainly deals with constructing a suitable semantic representation called Karaka tree and a sentence framing module to generate the natural summary.[12]

C. Malayalam text summarization: An extractive approach

Automatic summarization of text is on of the area of interest in the field of natural language processing. The proposed method utilizes the sentence extraction in a single document and produces a generic summary for a given Malayalam document (Extractive summarization). Sentences in the document are ranked based on the word score of each word present in it. Top N ranked sentences are extracted and arrange them in their chronological order for summary generation, where N represents the size of summary with respect to the percentage of original document size (condensation rate). The standard metric ROUGE is used for performance evaluation. ROUGE calculates the n-gram overlap between a generated summary and reference summaries. Reference summaries were constructed manually. Experiments show that the results are promising.[13]

D. Named entity recognition in Malayalam using fuzzy support vector machine

Named entities in a text are the atomic elements that represent the name of something, and the name can be a person name, name of an organization, name of a place or location etc. In the field of information extraction the identification and classification of named entities are quite an important task. The identification and classification of the named entities in a text into some predefined classes is known as named entity recognition. The commonly used pre-defined

classes are the person name, place, name of organization, date, numericals, measurements, and so on. It is a subtask of information extraction and many other natural language processing like text summarization, text categorization, question answering etc. In the case of language processing of Malayalam documents no effective tools are readily available. Through this paper a named entity recognition for Malayalam language is presented. The system proposed follows machine learning approach using support vector machine integrated with fuzzy module for improving the performance. The design is a kind of One-Against-All-Multi classification technique to solve the ambiguity caused by traditional SVM classifier. The system is based on contextual semantic rules and linguistic grammar rules. Malayalam NER is a challenging work as there is no specific feature for identifying named entities like capitalization feature in English. Also no named entity tagged corpus for Malayalam language is available for training the system. The system defines four primary named entity classes, i.e, Name, Organization, Place and Date.[14]

E. HMDSAD: Hindi multi-domain sentiment aware dictionary

Sentiment Analysis is a fast growing sub area of Natural Language Processing which extracts user's opinion and classify it according to its polarity into positive, negative or neutral classes. This task of classification is required for many purposes like opinion mining, opinion summarization, contextual advertising and market analysis but it is domain dependent. The words used to convey sentiments in one domain is different from the words used to express sentiments in other domain and it is a costly task to annotate the corpora in every possible domain of interest before training the classifier for the classification. We are making an attempt to solve this problem by creating a sentiment aware dictionary using multiple domain data. The source domain data is labeled into positive and negative classes at the document level and the target domain data is unlabeled. The dictionary is created using both source and target domain data. The words used to express positive or negative sentiments in labeled data has relatedness weights assigned to it which signifies its cooccurrence frequency with the words expressing the similar sentiments in target domain. This work is carried out in Hindi, the official language of India. The web pages in Hindi language is booming very quickly after the introduction of UTF-8 encoding style. The dictionary can be used to classify the unlabeled data in the target domain by training a classifier[15].

V. SENTIMENT ANALYSIS

Sentiment Analysis (SA), also known as opinion mining, is a powerful tool you can use to build smarter products. Its a natural language processing algorithm that gives you a general idea about the positive, neutral, and negative sentiment of texts. Sentiment analysis is often used to understand the opinion or attitude in tweets, status updates, movie/music/television reviews, chats, emails, comments, and more. Social media monitoring apps and companies all rely on sentiment analysis and machine learning to assist them in gaining insights about mentions, brands, and products. There are so many applications for natural language processing (NLP). Sentiment analysis is one of the major application of NLP. In sentimental analysis the emotional polarity of a given text is analyzed and classified as positive, negative, or neutral .The difficult task is to refine the classification in to different modes such as happy, sad, angry, etc Analyzing a natural language for mood extraction is not at all an easy task for a computer.[32]

A. Sentiment Analysis for Kannada using mobile product reviews: A case study 2015 (2015)

In this paper, a case study of Kannada SA for mobile product reviews is proposed as there are many user generated Kannada product reviews available online. In this approach a lexicon based method for aspect extraction has been developed. Furthermore, the Naive Bayes classification model is applied to analyze the polarity of the sentiment due to its computational simplicity and stochastic robustness. This was the first attempt in Kannada to the best of author's knowledge. Therefore, a customized corpus has been developed. The weekly reviews from the column 'Gadget Loka' by U.B Pavanaja are considered to develop this corpus. The source for this is the famous Kannada news paper 'Prajavani'. The preliminary results indicate this approach is an efficient technique for Kannada SA.[16]

B. Domain Specific Sentence Level Mood Extraction from Malayalam Text (2012)

This paper proposes a method of extracting the moodfrom a Malayalam sentence and this paper focuses on tagging the appropriate mood in Malayalam text. Tagging is used to specify whether a sentence indicates a sad, happy or angry mood of the person involved or if the sentence contains just facts, devoid of emotions. This research work is heavily dependent on the language since the structures vary from language to language. Mood extraction and tagging has been successfully implemented for English and other European languages. For the south Indian language Malayalam, no significant work has yet been done on mood extraction. They will be focusing on domain-specific sentence-level mood extraction from Malayalam text. The extraction process involves parts-of-speech tagging of the input sentence, extracting the patterns from the input sentence which will contribute to the mood of the sentence, such as the adjective, adverb etc., and finally scoring the sentence on an emotive scale by calculating the semantic orientation of the sentence using the extracted patterns. Mood extraction enables computers to automate the activities performed by human for making decisions based on the moods of opinions expressed in Malayalam text. In this paper mainly tells about a domain specific sentence level mood extraction from Malayalam text. For an input Malayalam sentence belonging to a particular domain and extract the corresponding mood from the sentence and this paper also focus on the sentence-level mood extraction.It is significant

because in most websites, user comments are just a single sentence because they focusing on a specific domain because different domains may use different words to express the mood.[17]

VI. Question Answering System

Question answering (QA) is a computer science discipline within the fields of information retrieval and natural language processing (NLP), which is concerned with building systems that automatically answer questions posed by humans in a natural language. A QA implementation, usually a computer program, may construct its answers by querying a structured database of knowledge or information, usually a knowledge base. More commonly, QA systems can pull answers from an unstructured collection of natural language documents.[31]

A. An empirical study of the impact of E-Learning Tool developed for dyslexic children with special reference to selective schools in Tamil Nadu, South India (2016)

This paper contain analyses impact of E-Learning Tool named R-U-LEXIC for dyslexic children . The objective of the work is to create a combined and interactive environment, where children may screened on a mass scale for dyslexia by means of an online tool named R-U-LEXIC. The research focuses on developing software platforms, integrating man-machine interfaces in the screening and remediation process and elaborating their technical specifications in view of their later integration within a local or national network. The research analysis was performed using SPSS Statistic 17.0. The statistical techniques applied for drawing statistical inferences and conclusions about the study included descriptive statistics, mean and standard deviation, one sample t test, one-way ANOVA and reliability test. The results of this study clearly revealed that there is a positive relationship between the data collected from Parents and Teachers and the students were excited and happy to take the test and could understand and use the E-learning tool easily.[18]

B. Named entity recognition in Malayalam using fuzzy support vector machine (2016)

Named entities in a text are the atomic elements that represent the name of something, and the name can be a person name, name of an organization, name of a place or location etc. In the field of information extraction the identification and classification of named entities are quite an important task. The identification and classification of the named entities in a text into some pre-defined classes is known as named entity recognition. The commonly used pre-defined classes are the person name, place, name of organization, date, numericals, measurements, and so on. In the case of language processing of Malayalam documents no effective tools are readily available. Through this paper a named entity recognition for Malayalam language is presented. The system proposed follows machine learning approach using support vector machine integrated with fuzzy module for improving the performance. The design is a kind of One-Against-All-Multi classification technique to solve the ambiguity caused by traditional SVM classifier. The system is based on contextual semantic rules and linguistic grammar rules. Malayalam NER is a challenging work as there is no specific feature for identifying named entities like capitalization feature in English. Also no named entity tagged corpus for Malayalam language is available for training the system. The system defines four primary named entity classes, i.e, Name, Organization, Place and Date. Named Entity Recognition is sometimes also known as named entity identification or named entity chunking or named entity extraction. NER in Malayalam is difficult as it lacks certain features like capitalization. Also the Indian languages are mostly morphologically rich, agglutinative and resource scarce language which makes the study on it difficult. The major difficulties are that the atomic elements which are the named entities may firstly be difficult to locate, and once identified, difficult to classify . And also making distinctions based on world knowledge is difficult (eg. organization name or person name). The system proposed in the paper works on Malayalam language, which is an Indian language. Indian languages belong to several major language families like the Indo-European languages, Indo-Aryan and the Dravidian languages. The difficulties in recognizing NEs occur due to some language properties .[19]

VII. Text Categorization

Text categorization (a.k.a. text classification) is the task of assigning predefined categories to free-text documents. It can provide conceptual views of document collections and has important applications in the real world. For example, news stories are typically organized by subject categories (topics) or geographical codes; academic papers are often classified by technical domains and sub-domains; patient reports in health-care organizations are often indexed from multiple aspects, using taxonomies of disease categories, types of surgical procedures, insurance reimbursement codes and so on. Another widespread application of text categorization is spam filtering, where email messages are classified into the two categories of spam and non-spam, respectively.[30]

A. A computational framework for Tamil document classification using Random Kitchen Sink (2015)

The main objective of this paper is to develop an computational framework for supervised Tamil document classification task. This paper highlights the performance of Random Kitchen Sink, a randomization algorithm, in Grand Unified Regularized Least Squares (GURLS), a Machine Learning Library, is proven to be comparably better than the conventional kernel based classifier in terms of accuracy. Henceforth, we claim that Random Kitchen Sink can be an effective alternative to the kernels for a classifier The manual classification of the Tamil documents is expensive, requires man-work, chiefly domain specialist and highly time-consuming, for the reason, it undergoes redundant process for a new set of documents. Because of the limitations of manual document classification, the automatic classification came to limelight. An automated document classification lies at the crossroads of Machine

Learning (ML), Natural Language Processing (NLP) and Information Retrieval (IR) . Information Retrieval is the task of extracting the documents that is appropriate to the search query from within the large collection . Automatic summarization, web searching, information filtering are some applications of IR . Natural Language Processing is utilized to represent the document semantically for improving the classification task. Certain NLP techniques are used in pre-processing phase, which include stemming, lemmatization, stop words removal, Parts Of Speech tagging etc The classification algorithms largely fall under three categories: supervised learning, unsupervised learning and semi supervised learning . The classifier is trained with the labeled data, in supervised learning algorithm.[20]

B. An Automated System for Tamil Named Entity Recognition Using Hybrid Approach (2014)

Named Entity Recognition is the process of identifying and recognizing named entities such as person, organization, location, date, time and money in the text documents. Named Entity Recognition is a subtask of Information Extraction. Information Extraction is the process of extracting the relevant data from documents. In this project implement a named entity recognizer using the hybrid approach that uses both Rule based and Hidden Markov Model in succession, which identifies only person, location and organization names respectively. Input data for proposed Named Entity Recognition system is any text document related to the any domain but limited size corpora respectively in Tamil language. In this system are tagging each word by using POS tagger and then imposing certain rules such as Lexical features and use some Gazetteers. HMM model using E-M algorithm is taken output data from trained as input to recognition system. The main purpose of this system identifies unknown entities and solves the problem of same name entity in different positions in the same document. The system is measuring the recall and precision parameters calculate the F-measure score. Goal of this project is to improve the performance of NER system to achieving high F-measure score.[21]

C. Kernel based part of speech tagger for Kannada (2010)

The proposed paper presents the development of a part-of-speech tagger for Kannada language that can be used for analyzing and annotating Kannada texts. POS tagging is considered as one of the basic tool and component necessary for many Natural Language Processing (NLP) applications like speech recognition, natural language parsing, information retrieval and information extraction of a given language. In order to alleviate problems for Kannada language, we proposed a new machine learning POS tagger approach. Identifying the ambiguities in Kannada lexical items is the challenging objective in the process of developing an efficient and accurate POS Tagger. We have developed our own tagset which consist of 30 tags and built a part-of-speech Tagger for Kannada Language using Support Vector Machine (SVM). A corpus of texts, extracted from Kannada news papers and books, is manually morphologically analyzed and tagged using our developed tagset. The performance of the system is evaluated and we found that the result obtained was more efficient and accurate compared with earlier methods for Kannada POS tagging. The development of POS tagger for a language with limited electronic resources can be very demanding. In the proposed study we have presented a Part-Of-Speech tagger for Kannada language which is modeled using SVM kernel. First we have conducted a linguistic study to determine the internal linguistic structure of the Kannada sentence and based on this developed a suitable tagset. A corpus size of fifty four thousand words was used for training and testing the accuracy of the tagger generators. From the experiment we found that accuracy increased with increasing the number of words in the corpus and the result obtained was more efficient and accurate compared with earlier methods for Kannada POS tagging. We conclude that the proposed part of speech tagger will helpful to develop natural language application like parser, bilingual machine translation, and in many areas in Kannada language.[22]

VIII. Prediction system / Decision making system

Prediction system allow to process and identify the accurate minimum words given in a context for machines it is difficult to predict Some extent of intelligence may add to the machine for an accurate prediction.

A. Malayalam word sense disambiguation using Nave Bayes classifier

Word Sense Disambiguation is the process of identifying accurate meaning of a polysemous words given in a context. The paper proposes a Supervised Malayalam word sense disambiguation system using Naive Bayes classifier. Word Sense Disambiguation is a complex problem in NLP because a particular word may have different meanings in different situations. For all human beings it is very easy to find out the accurate sense in a particular context but for machines it is very difficult to predict. Some extent of intelligence may add to the machine for an accurate prediction. Here this proposed system provide us 95% reliability using a corpora of 1 lakh words.[23]

B. A system for identification of idioms in Hindi

Idioms are extensively used in everyday language. They carry a metaphorical sense that makes their comprehension difficult as their meaning cannot be deduced from the meaning of their constituent parts. They pose a challenge for Natural language processing (NLP) applications like machine translation, information retrieval and question answering as their translation and meaning needs to be derived logically rather than literally. A lot of research work has been carried out into automatic extraction of multi-word expressions, but no comprehensive work has been reported on idioms in Hindi. In this paper, an attempt has been made to study the linguistic and morphological variations that are usually encountered in idioms in Hindi. Based on this study, a methodology for

deriving rules for representation of idioms and their search has been developed. The rules representing the idioms are hand crafted. For the idiom identification, rule-base has been used to mark the input text for probable presence of idiom. Our system is limited to use only intra-sentential context. The experimental results demonstrate feasibility and scalability of our methodology.[24]

IX. Text mining

Text mining, also referred to as text data mining, roughly equivalent to text analytics, is the process of deriving high-quality information from text. High-quality information is typically derived through the devising of patterns and trends through means such as statistical pattern learning. Text mining usually involves the process of structuring the input text (usually parsing, along with the addition of some derived linguistic features and the removal of others, and subsequent insertion into a database), deriving patterns within the structured data, and finally evaluation and interpretation of the output. 'High quality' in text mining usually refers to some combination of relevance, novelty, and interestingness. Typical text mining tasks include text categorization, text clustering, concept/entity extraction, production of granular taxonomies, sentiment analysis, document summarization, and entity relation modeling (i.e., learning relations between named entities).

A. Information extraction and text mining of Ancient Vattezhuthu characters in historical documents using image zoning

The aim of this paper is to develop a system that involves character recognition of Brahmi, Grantha and Vattezuthu characters from palm manuscripts of historical Tamil ancient documents, analyzed the text and machine translated the present Tamil digital text format. Though many researchers have implemented various algorithms and techniques for character recognition in different languages, ancient characters conversion still poses a big challenge. Because image recognition technology has reached near-perfection when it comes to scanning English and other language text. But optical character recognition (OCR) software capable of digitizing printed Tamil text with high levels of accuracy is still elusive. Only a few people are familiar with the ancient characters and make attempts to convert them into written documents manually. The proposed system overcomes such a situation by converting all the ancient historical documents from inscriptions and palm manuscripts into Tamil digital text format. It converts the digital text format using Tamil unicode. Our algorithm comprises different stages: i) image preprocessing, ii) feature extraction, iii) character recognition and iv) digital text conversion. The first phase conversion accuracy of the Brahmi script rate of our algorithm is 91.57% using the neural network and image zoning method. The second phase of the Vattezhuthu character set is to be implemented. Conversion accuracy of Vattezhuthu is 89.75%.[25]

B. SentiMa - Sentiment extraction for Malayalam

This paper proposes a rule based approach for sentiment analysis from Malayalam movie reviews. The research in Sentiment Analysis nowadays become one among active research areas in natural language processing. Sentiment Analysis is the cognitive process in which the user's feeling and emotions are extracted. The growing importance of sentiment analysis coincides with the growth of social media such as reviews, forum discussions, blogs, and social networks. Sentiment analysis enables computers to automate the activities performed by human for making decisions based on the sentiment of opinions, which has wide applications in data mining, web mining, and text mining. Negation Rule has been applied for extracting the Sentiments from a given text. This system gives the polarity at the sentence level for the movie reviews with an accuracy of 85%, when analysed.[26]

C. An Extractive Malayalam Document Summarization Based on Graph Theoretic Approach

Text summarization is a way to condense the large amount of information into a concise form by the process of selection of important information and discarding unimportant and redundant information. The need for Text summarization has increased much due to the abundance of documents in the internet.Even though a lot of text summarization systems have been developed for summarizing documents in various languages, there is no such well performing system for Malayalam.In this paper, we propose the use of Graph theoretic approach for summarizing Malayalam documents that is motivated by the method of identification of themes. After the common preprocessing steps, namely, stop word removal and stemming, sentences in the documents are represented as nodes in an undirected graph. There is a node for every sentence. Two sentences are connected with an edge if the two sentences share some common words, or in other words, their (cosine, or such) similarity is above some threshold. This representation yields two results: The partitions contained in the graph (that is those sub-graphs that are unconnected to the other sub graphs), form distinct topics covered in the documents. The second result yielded by the graph-theoretic method is the identification of the important sentences in the document. We apply graph theoretic approach on Malayalam text summarization task and achieve comparable results to the state of the art.[27]

X. Parts of speech

A Part-Of-Speech Tagger (POS Tagger) is a piece of software that reads text in some language and assigns parts of speech to each word (and other token), such as noun, verb, adjective, etc., although generally computational applications use more fine-grained POS tags like 'noun-plural'. This software is a Java implementation of the log-linear part-of-speech taggers described in these papers (if citing just one paper, cite the 2003 one).

A. Evaluation for POS tagger, chunk and resolving issues in word sense disambiguate in machine translation for Hindi to English languages

Our paper develops innovative algorithms for machine translation system based on the innovative algorithms for parts of speech tagger, chunking, word sense disambiguate and word translation in English. Parts of speech tagging and chunking for 1657 tokens with 990 phrases for Hindi languages and to calculate the accuracy we created confusion matrix an evaluate Precision, Recall, F-score, Accuracy for chunk accuracy: 81.23%; precision: 66.57%; recall: 73.03%; F-score: 69.65, 90.31, POS accuracy: 94.75%; precision: 90.67%; recall: 93.23%; F-score 91.93. Rule based and learning algorithm (Conditional Random Fields) is used to develop the system.[28]

B. A hybrid Parts Of Speech tagger for Malayalam language

Parts of speech tagging is an important research topic in Natural Language Processing research are. Since it is one among the first steps of any natural language processing (NLP) techniques such as machine translation, if any error happens for tagging the same will repeat in the whole NLP process. So far works had been done on POS tagging based on SVM, MBLP, HMM, Ngram. All of these methods were not fixing the problem of ambiguity. So for fixing ambiguity, we put forward a new Hybrid tagger for Malayalam. The combination of traditional rules and n-gram may produce better result compared to other methodologies. And also the ambiguity will be reduced by enriching the bigram dictionary. A bigram dictionary of co-occurring words are built with their tags. About 100000 more words are there in bigram dictionary. A corpus for Malayalam must be built which may be supposed to access by the model. It contains about 100000 words which are Malayalam words as well as the words originated from English. Since it's a hybrid tagger, we can take advantage of both traditional rules as well as bigrams. Also the heart of the research is the rule set, which contains 267 manually created rules. Rules can be applied with help of a morph analyzer. Rules are also used for tagging if bigram and corpus can't be referred for tagging. The proposed method when tested on 150 words, only 11 words were not identified, and obtained 90.5% accuracy. For the unidentified words, it can be caused by either the root word may not be in corpus or bigram, or the absence of rule. So adding the word, bigram or rule, we can improve the result and enhance the work. Addition is simple task. The size of bigram dictionary, corpus, and rule set and accuracy of morph analyzer influences the performance of the system.[29]

XI. Conclusion

This paper aims to address about the concept of natural language processing and the analysis of how effective it in case of south Indian languages.Studies on south Indian languages are done which include Tamil, Telugu and Malayalam. Most of the reserach studies are undergone in Malayalam language. As we are using regional languages to communicate with the system ,it increases the bond between the user and the computer.They are mainly used for popular languages as it is very difficult task to accomplish.

References

[1] Rosna P Haroon, T A Shaharban, *Malayalam machine translation using hybrid approach*, Electrical, Electronics, and Optimization Techniques (ICEEOT), International Conference on 2016.

[2] Ancy Anto,K. K. Nisha, *Text to speech synthesis system for English to Malayalam translation* , Emerging Technological Trends (ICETT), International Conference on 2016.

[3] T. A. Shaharban, *Conversion of Malayalam text to Indian sign language using synthetic animation*, Next Generation Intelligent Systems (ICNGIS), International Conference on 2016.

[4] Steffy Maria Joseph, *SVM based feature set analysis in dynamic malayalam handwritten character recognition* , Signal and Image Processing Applications (ICSIPA), 2015 IEEE International Conference on 2015.

[5] Abdul Hameed, *Online handwritten malayalam character recognition using LIBSVM in matlab*, Communication, Signal Processing and Networking (NCCSN), 2014 National Conference on 2014.

[6] Ganpat Singh G Rajput , *Zone based handwritten Kannada character recognition using crack code and SVM* , Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on 2013.

[7] Janardhanan P.S. Nair, *Domain Specific Sentence Level Mood Extraction from Malayalam Text* , Advances in Computing and Communications (ICACC), 2012 International Conference on 2012.

[8] Latha R. Nair , *Shallow parser for Malayalam language using finite state cascades* , Advances in Computing, Communications and Informatics (ICACCI), 2013 International Conference on 2013.

[9] Yashaswini Hegde, *Suffix stripping algorithm for Kannada information retrieval* , Emerging Technological Trends (ICETT), International Conference on 2016.

[10] P. J. Antony , *Haar features based handwritten character recognition system for Tulu script* , Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE International Conference on 2016.

[11] Jagadish S Kallimani , *Summarization of customer reviews for a product on a website using natural language processing* , Natural Language Processing andKnowledge Engineering (NLP-KE), 2011 7th International Conference on 2011. 2015.

[12] Greeshma N Gopal, *Document Summarization in Malayalam with sentence framing*, Information Science (ICIS), International Conference on 2016.

[13] A Sooryanarayanan , *Malayalam text summarization: An extractive approach* , Next Generation Intelligent Systems (ICNGIS), International Conference on 2016.

[14] Janu R. Panicker, *Named entity recognition in Malayalam using fuzzy support vector machine* , Information Science (ICIS), International Conference on 2016

[15] Vandana Jha, Savitha R., *HMDSAD: Hindi multi-domain sentiment aware dictionary* , Computing and Network Communications (CoCoNet), 2015 International Conference on 2015.

[16] S.K Padma, Yashaswini Hegde *Sentiment Analysis for Kannada using mobile product reviews: A case study 2015 (2015)*, Advance Computing Conference (IACC), 2015 IEEE International Conference on 2015.

[17] Priyanka, *Domain Specific Sentence Level Mood Extraction from Malayalam Text (2012)* , Advances in Computing and Communications (ICACC), 2012 International Conference on 2012.

[18] Dr.V.Thulasibai, Dr.P.M.Beulah Devamalar, Dr.B.SenthilKumar, Mr.Arun Marx , *An empirical study of the impact of E-Learning Tool developed for dyslexic children with special reference to selective schools in Tamil Nadu, South India (2016)*, Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), 2016 2nd International Conference on 2016.

[19] Lakshmi G,Janu R.panicker,Meera M , *Named entity recognition in Malayalam using fuzzy support vector machine (2016)* , Information Science (ICIS), International Conference on 2016.

[20] Sanjanasri J.P,Anand Kumar M , *A computational framework for Tamil document classification using Random Kitchen Sink (2015)*, Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on 2015.

[21] Mrs.N.Jeyashenbagavalli,Dr.K.G.Srinivasagan,Mrs.S.Suganthi , *An Automated System for Tamil Named Entity Recognition Using Hybrid Approach (2014)* , Intelligent Computing Applications (ICICA), 2014 International Conference on 2014.

[22] E.K. Vellingiriraj, *Kernel based part of speech tagger for Kannada (2010)*, Machine Learning and Cybernetics (ICMLC), 2010 International Conference on 2010.

[23] Antony P.J,Soman K.P , *Malayalam word sense disambiguation using Nave Bayes classifier*, Advances in Human Machine Interaction (HMI), 2016 International Conference on 2016

[24] R. M. K. Sinha, *A system for identification of idioms in Hindi* , Contemporary Computing (IC3), 2014 Seventh International Conference on 2014.

[25] E.K. Vellingiriraj, *Information extraction and text mining of Ancient Vattezhuthu characters in historical documents using image zoning*, Asian Language Processing (IALP), 2016 International Conference on 2016

[26] Elizabeth Sherly, *SentiMa - Sentiment extraction for Malayalam* , Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on 2014.

[27] Ajmal E. B, *An Extractive Malayalam Document Summarization Based on Graph Theoretic Approach* , e-Learning (econf), 2015 Fifth International Conference on 2015.

[28] Umesh Chandra Jaiswal, *Evaluation for POS tagger, chunk and resolving issues in word sense disambiguate in machine translation for Hindi to English languages* , Computing for Sustainable Global Development (INDIACom), 2016 3rd International Conference on 2016.

[29] Anisha Aziz T , *A hybrid Parts Of Speech tagger for Malayalam language* , Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on 2015.

[30] "sitename"[online].available:*http://www.scholarpedia.org/article/Text categorization [Accessed : 08-04-2017]*

[31] "sitename"[online].available:*https://en.wikipedia.org/wiki/Question answering [Accessed : 12-04-2017]*

[32] "sitename"[online].available:*http://blog.algorithmia.com/benchmarking-sentiment-analysisalgorithms [Accessed : 12-04-2017]*

[33] "sitename"[online].available:*https://en.wikipedia.org/wiki/Institute of Electrical and Electronics Engineers [Accessed : 15-04-2017]*

[34] "sitename"[online].available:*https://www.ieeefoundation.org/ [Accessed : 17-04-2017]*

[35] "sitename"[online].available:*https://www.ieee.org/ [Accessed : 17-04-2017]*

[36] "sitename"[online].available:*https://www.ieeefoundation.org/ [Accessed : 17-04-2017]*

[37] "sitename"[online].available:*https:// www.theiet.org Resources Library and Archives Virtual Library [Accessed : 17-04-2017]*

Mobile Cloud Computing: A Survey

Anjana P, Arunima K M, Thuhin Mol P S
Department of Computer Applications
Vidya Academy of
Science and Technology
Thrissur-680501

Siji K B
Assistant Professor of Computer Applications
Vidya Academy of
Science And Technology
Thrissur-680501

***Abstract*—Mobile Application and cloud computing concept lead to the exponential growth of Mobile Cloud Computing. Mobile Cloud Computing integrates cloud computing into mobile environment and deals with the obstacles related to performance, security etc discussed in Mobile Cloud Computing. This paper gives a survey of Mobile Cloud Computing definition, architecture,advantages, applications, issues and approaches are presented.**

***Index Terms*—Cloud Computing Mobile Cloud Computing Mobile Computing**

I. Introduction

Now days we are no longer restricted to saving all of our mission information on one physical device, such as a floppy disk, CD or USB flash drive. Cloud Computing has forever changed the way in which companies store and access their data. The cloud is safe and secure, it helps to boost productivity and it is cost effective. For this reasons, countless companies from numerous industries have made the switch to the cloud. But what about file access? Whether in the office or on the road, your employees need the ability to securely access vital technology resources and company data. The solution: Mobile Cloud Computing, which is simply the use of Cloud Computing technology on a mobile device. The rest of the paper is organized as follows. Section 2: Cloud Computing, Mobile Computing, Mobile Cloud Computing, architecture. Section 3 discuss about advantages. Section 4: applications. The section 5 presents several issues and approaches in Mobile Cloud Computing. Finally we summarize and conclude in section 6.

II. Overview of mobile Cloud Computing

Before going through Mobile Cloud Computing at headmost we have to get a better understanding of cloud computing and Mobile Computing.

A. Cloud Computing

Cloud Computing is a technology that uses the internet and central remote servers to maintain data and applications. A simple example of Cloud Computing is Yahoo email or Gmail etc. A style of computing where massively scalable (and elastic). It related capabilities are provided as a service to external customers using the internet technologies. [1]Clod Computing is referred to the use of network structure software and capability to provide resources to users in an on demand environment as shown in figure 1.

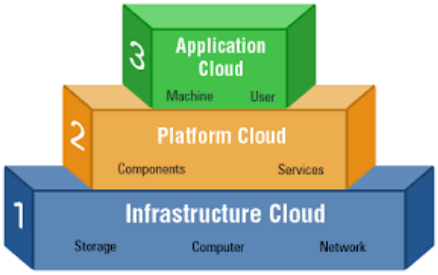


Fig. 1. Cloud Computing

B. Mobile Computing

In Wikipedia, it is described as a form of human-computer interaction by which a computer is expected to be transported during normal usage. The following are features of Mobile Computing.

1) Mobility: In a mobile network, it is possible for the mobile nodes to establish connection with others, even with fixed nodes in wired network through Mobile Support Station(MSS).

2) Diversity of Network conditions: Normally, network used by mobile nodes are not unique. Such as network can be a wired network with high-bandwidth or a Wireless Wide Area Network (WWAN).

3) Frequent Consistency and Disconnection: Due to the limitation of battery power, charge of wireless communication, network conditions and so on, mobile nodes will not always keep the connection, but disconnect and consistent with the wireless network passively or actively.

4) Challenges: Compared to traditional wired network, mobile computing network may face various problems and challenges in different aspects, such as security, limited power, signal disturbance etc.

C. Mobile Cloud Computing

The mobile cloud computing forum defines Mobile Cloud Computing as follows [2]:Mobile Cloud Computing at its

simplest refers to an infrastructure where both the data storage and the data processing happen outside of the mobile device. Mobile Cloud Applications move the computing power and data storage away from mobile phones and into the cloud, bringing applications and Mobile Computing to not just Smartphone users but a much broader range of mobile subscribers. Aepona[3] describes Mobile Cloud Computing as a new paradigm for mobile applications where by the data processing and storage are moved from the mobile device to powerful and centralized computing platform located in clouds. These centralized applications are then accessed over the wireless connection based on a thin native client or web browser on the mobile devices. Alternatively, Mobile Cloud Computing can be defined as a combination of mobile web and Cloud Computing which is the most popular tool for mobile users to access applications and services on the internet. Briefly, Mobile Cloud Computing provides mobile users with the data processing and storage services in clouds. The mobile devices do not need a powerful configuration (eg, CPU speed and memory)

From the concept of Mobile Cloud Computing, the general architecture of Mobile Cloud Computing can be shown in figure 2. In figure, the mobile devices are connected to the mobile networks via base stations (eg, base transceiver station (BTS), access point or satellite) that establish and control connections (air links) and functional interfaces between the networks and mobile devices. Mobile users requests and information (eg id and location) are transmitted to the central processors that are connected to servers providing mobile network services. Here, mobile network operator can provide services to mobile users as AAA (for authentication, authorization and accounting) based on the home agent (HA) and subscribers data stored in databases. After that, the subscribers requests are delivered to a cloud through the internet. In the cloud, cloud controllers process the request to provide mobile users with the corresponding cloud services. These services are developed with the concept of utility computing, virtualization, and service-oriented architecture (eg web, application and database servers). The details of cloud architecture could be different in different contexts. For example, four layer architecture is explained in [4] to compare cloud computing with grid computing.

III. Advantages of Mobile Cloud Computing

1) Improving Processing Power and Data Storage Capacity: One of the main constraints for mobile devices is also storage capacity. Mobile Cloud Computing is developed to enable mobile clients to store or access the large amount of data on the cloud through wireless networks. First example is the Amazon S3 (Simple Storage Device) [5] which supports file storage service. Another example is Image Exchange which utilizes the large storage space in clouds for mobile clients [6] Flicker [7] and Shozu[8] is also the successful mobile photos sharing applications based on Mobile Cloud Computing. Facebook[9] is the most successful social network application today, and it is also a typical example of using cloud in sharing images. Mobile Cloud Computing also helps in reducing the running cost for compute-intensive applications that take long time and large amount of energy when performed in the limited-resource devices [10] Cloud Computing can efficiently support various tasks for data warehousing, managing and synchronizing multiple documents online for example, clouds can be used for transcoding, playing chess, or broadcasting multimedia services to mobile devices. In these things, all the examples calculations for transcoding or offering an optimal chess move that take a long time when perform on mobile devices will be processed efficiently on the cloud.

2) Improved Reliability: Storing data/information or running applications on clouds is an effective way to improve the reliability because the data and application are stored and backup on a number of computer systems. This one reduces the chances of data and application last on the mobility devices. In addition, Mobile Cloud Computing can be designed as a comprehensive data security model for both service providers and users. For example, the cloud can be used to protect copyrighted digital contents (eg video, clip and music) from being abused and unauthorized distribution [11].

3) Extending battery lifetime: Battery is one of the main concerns for mobile devices. Numerous solutions have been proposed. However, these solutions require changes in the structure of mobile devices. But these changes may not be feasible for all mobile devices. In order to execute the large computations and complex processing from resource limited devices like mobile devices to resourceful machines such as servers in clouds several computations offloading technique is proposed. Mobile Cloud Computing avoids taking a long application execution time on mobile devices which may results in large amount of power consumption.

4) Dynamic Provisioning: Dynamic provisioning of resources is a flexible way for service providers and mobile users to run their various applications without advanced reservations of resources without storing data in mobile devices it be stored in cloud and can be accessed dynamically.

IV. Applications of Mobile Cloud Computing

Mobile Cloud Computing has a large number of applications [12] in various fields and a wide range of potential mobile cloud applications have been recognized in the present literature.

A. Image Processing

In [12], the authors try to experiment with running GOCR and an optical character recognition (OCR) program on a collection of different mobile devices. A similar scenario is given in [13]. If user/subscribers visit foreign museum, he cant perceive the language written in each object of the museum. He can take picture of the object and using Mobile Cloud Computing can understand the language written over the object.

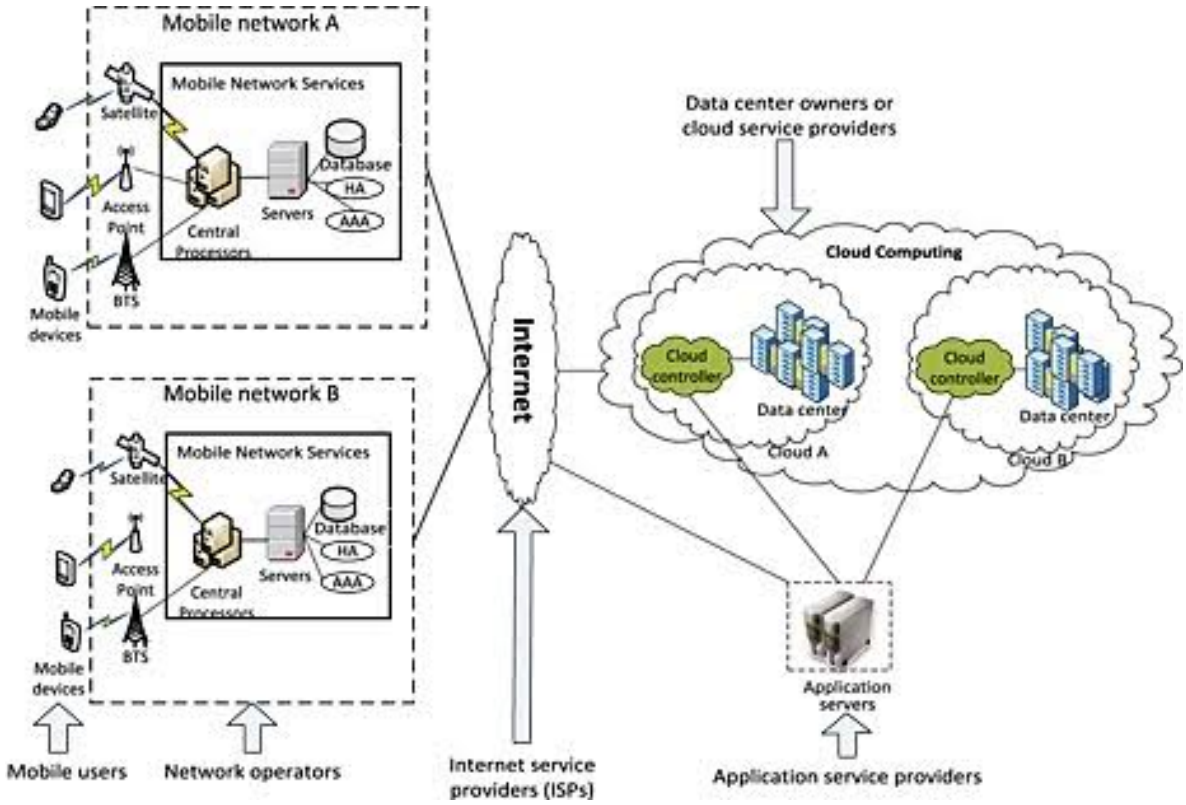


Fig. 2. Mobile Cloud Computing (MCC) Architecture

B. Natural Language Processing

Language translation is one possible application for Mobile Cloud Computing. Translations is viable candidate for language processing since different sentences and paragraphs can be translated independently and this is experimentally explored in [12] using Tangloss-Lite[13].

C. Sharing in GPS/Internet Data

Through local area or peer-to-peer network data can be store among a group of mobile devices that are near each other. It is faster as well as cheaper [14].

D. Sensor Data Applications

Now a days almost every mobile devices are built with sensors which are used to read data. Some sensors such as GPS, accelerometer, thermo sensors, light sensors, clock and compass may be time stamped and associated with each other phone readings.

E. Mobile Healthcare

The purpose of applying Mobile Cloud Computing in medical applications is to minimize the limitations of traditional medical treatment (eg small physical storage, security and privacy). Existing processors for patients vital data collection require a great deal of labor work to collect, input and analyze the information. These processes are usually slow and errors prone. This scenario restrains the clinical diagnostics and monitoring capabilities [16] propose a solution to automate this process by using sensors attached to existing medical equipment that are inter-connected to exchanged service. The proposal is based on the concepts of utility computing and wireless sensor networks. Mobile healthcare systems focus towards achieving two specific goals: the availability of e-health applications and medical information anywhere and anytime and the invisibility of computing [17]. Mobile pervasive healthcare technologies can support a wide range of applications and services including mobile telemedicine, patient monitoring, location based medical services, emergency responses and management, personalized monitoring and pervasive access to healthcare information, providing great benefits to both patients and medical personnel[17][18]. The realization however of health information management through mobile devices introduces several challenges, like data storage and management. One potential solution for addressing all issues is the introduction of Cloud Computing concept in electronic Healthcare system.

1) Healthcare Big Data: A big data analytics is motivated in healthcare through the following aspects [19]:

- Healthcare data is now growing very rapidly in terms of size, complexity, and speed of generation and traditional databases and data mining techniques are no longer efficient in storing, processing and analyzing these data.
- The patients behavioral data is captured through several

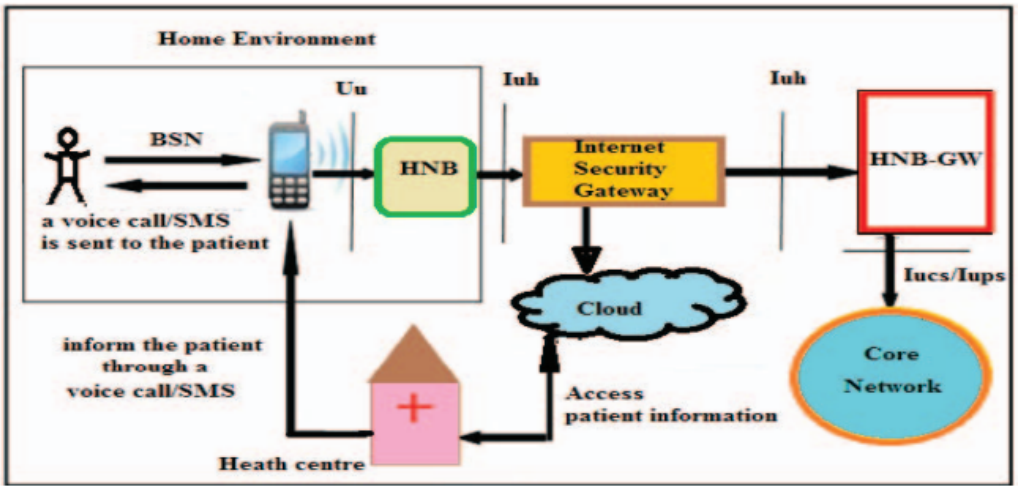


Fig. 3. Working model of proposed m-health monitoring scheme

sensors, patients various social interactions and communications.

- Understanding unstructured clinical notes in the right context.
- Efficiently handling large volumes of medical imaging data and extracting potentially useful information and bio markers.
- Inferring knowledge from complex heterogeneous patient sources and leveraging the patient/data correlations in longitudinal records.

2) Architecture of Mobile Cloud Computing:

3) Using Femtocell: In Femtocell and MCC-based M-Health Monitoring scheme, health iformation of every users are caputured using sensors and send it to the corresponding mobile devices. From the mobile device the data is transfered to the Femtocell under which the mobile devices are registered. In the Femtocell it is verified whether user's health is normal using a database stored inside the Femtocell. If any abnormality is detected the data will be send to the cloud. The health data are securely stored on the cloud and accessed by the health centre[20]. This scheme requires the following componenets:

- Body Sensor Network
- Mobile station(MS)
- Femtocell ie.,HNB
- Internet Connectivity

The working model of proposed m-health is pictorially depicted in figure 3. Mobile users and HNB are connected via Vu interface. HNB is connected to internet and Home Node Base station-Gateway(HNB-GW) via Iub interface. By HNB-GW, the HNB is connected with the core network. HNB-GW is connected with the core network via Iucs/Iups interface. To provide proper security between HNB and HNB-GW over the internet a security gateway(SeGW) is maintained[21].

In the proposed scheme if only the health data seems to be abnormal then the data are sent to the cloud through HNB. The health data are accessed on the cloud through th HNB in a network. If the health is normal, then no health data is sent to the cloud.

V. Issues and Approaches in Mobile Cloud Computing

A. Low Bandwidth

Bandwidth is one kind of the big issues in Mobile Cloud Computing. Since the radio resources for wireless network is limited as compared to traditional wired network.

B. Availability

Service availability is one of the major issue that need to be considered in the Mobile Cloud Computing than that in the Cloud Computing with wired networks. Mobile users may not be able to connect to the cloud to obtain service due to out-of-signal, network failures.

C. Security

The security related issues in Mobile Cloud Computing are introduced in two categories: the security for mobile users and the security for data.

1) Security for mobile users: Installing and running security software such as kaspersky, Mc Afee on mobile devices are simplest ways to detect security threats. However, mobile devices are concentrated in their processing and power. Cloud consists of cloud AV platform that provides an incloud service for malware detection. To apply cloud AV platform for the mobile environment, a mobile agent should be improved and customized to fit in the mobile devices [20].

2) Securing Data on clouds: Wherever you store nobody can give assurances that your data will be secured all the time. But if you are storing your data on clouds your data is 100% secured. Both the mobile users and application developers benefit from storing a large amount of data/application on a cloud. But you need to be careful dealing with the data/applications in terms of their integrity, authentication and digital rights [20]

D. Quality of Services

Mobile users satisfaction must be the aim of service providers. To achieve this aim the aim the need to monitor their preferences and should provide appropriate services to each of the users.

VI. Conclusions

The cloud is safe and secure, it helps to boost productivity and it is cost effective. For this reason the mobile applications are shifted to the cloud. Cloud Computing provides a brand new opportunities for the enlargement of mobile applications. A study by Junipr Research states that the consumer and enterprise market for cloud-based mobile applications is expected to rise to $ as billion by 2014. With this importance, this article has provided a survey of Mobile Cloud Computing in which its architecture, definitions and advantages have been presented.

References

[1] M.R. sudha,Ibrahim V.Llema ,"A study on Emerging Trends and Challenges in Mobile Cloud Computing"*International Journal on Recent and Innovation Trends in Computing and Communication*,January 2016
[2] http://www.mobilecloudcomputingforum.com/(2013)
[3] White Paper, "Mobile Cloud Computing Solution Brief",*AEPONA*,November 2010.
[4] I.Foster, Y.Zhao, I.Raicu, and S.Lu,"Cloud Computing and Grid computing 360-Degree Compared", *In Proceedings of Workshop on Grid Computing Environments(GCE)*,pp.1,Janury 2009
[5] http://aws.amazon.com/s3/(2013)
[6] Vartiainen E, Mattila KV-V, "User experience of mobile photo sharing in the cloud",*In Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia(MUM)*,2010
[7] http://www.flicker.com/(2013)
[8] P.Mell and T Grance,"The NIST definition of cloud computing(draft)",*NIST special publication,*vol 800,p.145,2011
[9] http://www.mobilecloudcomputingforum.com(2015)
[10] Surabhi S.Golechha, Prof.R.R Keole,"Mobile Cloud Computing :Worlds Leading Technology for Mobile Devices.",*International Journal of Computer Science and Information Technologies,Vol.5(2),*2014
[11] P.ZouC Wang, Z.Liu and D.Bao,"Phosphor: A cloud Based DRM Scheme with sim card",*in Proceedings of the 12th Intrnational Asia.Pacific web conference(APWC)*,pp.459,June 2010
[12] Chang,Balan,R K & Satyanarayanan,M.2005,"Exploiting rich mobile environment",2005
[13] Huang,Y,SuH,Sun,W.Zhong,J.M,Guo.C.J,Xu,J M & Zhu,J.2010" Framework for building a low-cost,scalable and secured platform for web-delivered business services", *IBM journal of Research and Development*,S.4(6),4.1
[14] Vallina,Redriguiz,N, & Croucraft, J.2011,"Erdos:achieving energy savings in mobile O S",*In preceedings of the sixth international workshop on MobileArch(pp.32.42)ACM*
[15] Marinelli,E.E2009.Hydra cloud computing on mobile devices using MapReduce(No CMU-CS-09-164)
[16] A Cloud Computing Solution for patients Data collection in Healthcare Institutions,Carlos Oberdous Rolion Fernando Luiz k
[17] Upkar Varshney,"Pervasive Healthcare",*Computer Magazine vol.36,no-12,2003,pp 138-140*
[18] Maglogiannis I,Doukas C,Korentzas G, Pliakas T, "warelet-Based Compression with ROI coding support for Mobile Access to DICOM Images Over Heterogeneous Radio Networks",*IEEE Transactions on Information Technology in Biomedicine,vol.13,no 4,pp.458-466 July 2009*
[19] J.Sun and C K Reddy "Big Data Analitics for Healthcare",*Tutorial Presentation at the SIAM International Conference on Data Mining,Auston,Tr.2013*
[20] Debashis De, Anwesha Mukherjee,"Femtocell Based Economic Health Monitoring Scheme Using Mobile Cloud Computing",*2014 IEEE International Advance Computing Conference(IACC)*
[21] J.Zhang ,G.Roche, and L.De,*Femtocells technologies and deployment*, A John Wiley and sons,Ltd.,Wiley,2010
[22] Jamshed Siddiqui, Shahab Saquib Sohail and Zaki Ahmad Khan, "A Extensive Survey on Mobile Cloud Computing"
[23] S.Perez "Mobile Cloud Computing & 9.6 billion by 2014",http://explnet.eu/catalog.php,2010

An Overview of Privacy Preserving Data Mining

Anju I S, Aswathy C B, Swathy M P
Dept of Computer application,
Vidya Academy of Science & Technology,
Thrissur-680501

Jisha Jose Panackal
Associate Professor of Computer Applications,
Vidya Academy of Science & Technology,
Thrissur-680501

***Abstract*—Data mining can be used to extract important knowledge from large databases. Privacy preserving data mining(PPDM) deals with protecting the privacy of individual data or sensitive knowledge without sacrificing the utility of the data. The success of privacy preserving data mining algorithms is measured in terms of its performance, data utility, level of uncertainty or resistance to data mining algorithms etc. Privacy in data mining can be obtained by various techniques like Perturbation, Anonymization and Cryptographic. Privacy preserving data mining works on two scenarios distributed scenario and publishing scenario So, the aim of this paper is to present overview of Privacy Preserving Data mining ,PPDM techniques and the privacy issues in distributed scenario using cryptography and in publishing scenario using anonymization approach the process of data mining.**

***Index Terms*—PPDM,Distortion Association,Clustering, Outsourcing K-anonymity, Anonymization, Perturbation, Pseudonymization, Cluster.**

I. INTRODUCTION

Data collection and storage technology have enabled organization to accumulate vast amounts of data. Traditional analysis tools and techniques cannot be used because of massive size of data. Data mining is a set of automated techniques used to extract hidden or buried information from large databases. The term data mining refers to the nontrivial extraction of valid, implicit, potentially useful and ultimately understandable information in large databases with the help of the modern computing devices. The objective of data mining is to generalize across populations, rather than reveal information about individuals. But, the problem is that data mining works by evaluating individual data. The goal of preserving privacy in data mining is to lower the risk of misuse of data and at the same time produce results same as that produced in the absence of such privacy preserving techniques.

In contrast to the centralized model, the Distributed Data Mining (DDM) model assumes that the data sources are distributed across multiple sites. Algorithms developed within this field address the problem of efficiently getting the mining results from all the data across these distributed sources. Most of the privacy preserving distributed data mining algorithms reveal nothing other than the final results. However, this will often fail to give globally valid results. Issues that cause a disparity between local and global results include:

- Values for a single entity may be split across sources. Data mining at individual sites will be unable to detect cross-site correlations.
- The same item may be duplicated at different sites, and will be over-weighted in the results.
- At a single site, it is likely to be from a homogeneous population. Important geographic or demographic distinctions between that population and others cannot be seen on a single site.

Data mining techniques have been widely used in many research disciplines such as medicine, life sciences, and social sciences to extract useful knowledge (such as mining models) from research data. Research data often needs to be published along with the data mining model for verication or reanalysis. However, the privacy of the published data needs to be protected because otherwise the published data is subject to misuse such as linking attacks.

The main aim of this paper is to give an overview about privacy methods in data mining. In this paper we just use various methods for privacy of individual information for the purpose of data mining. Here we suggest a new method for anonymizing data in the area of social networks and health care. Various PPDM methods helps in the data protection in a simpler way. The rest of this paper composed as:

In the section II we review related work. section III various PPDM concepts, section IV anonymization in Social networking and Health care services. Result and discussion work after by conclusion VI.

II. RELATED WORKS

To protect privacy of individuals, several methods can be applied on data before the process of data mining. The branch of study which include the privacy concern are referred as Privacy Preserving Data Mining(PPDM).

In the distributed scenario the data is encrypted by ensuring privacy of data by the studies like [1,2] it shows the privacy identified by various techniques including cryptography, perturbation, condensation etc [3,4].The techniques such as data hiding, compression, generalization and permutation are the common mechanisms to protect the privacy in this scenario[4,5]. In the centralized scenario data is published, so that it can be used for a variety of studies like Social networking

and Health care data[6 to 9]. By using anonymization the data in health care is protection by Re-Identification strategies[9,10].

III. PPDM Concepts

Privacy preserving data mining has gained increasing popularity in data mining research community. PPDM has become an important issue in data mining research . As a result, a whole new set of approaches were introduced to allow mining of data, while at the same time prohibiting the leakage of any private and sensitive information. The majority of the existing approaches can be classified into two broad categories :

- methodologies that protect the sensitive data itself in the mining process, and
- methodologies that protect the sensitive data mining results (i.e. extracted knowledge) that were produced by the application of the data mining.

The first category refers to the methodologies that apply perturbation, sampling, generalization or suppression, transformation, The second category deals with techniques that prohibits the disclosure sensitive knowledge patterns derived through the application of data mining algorithms as well as techniques for downgrading the effectiveness of classifiers in classification tasks, such that they do not reveal sensitive knowledge. PPDM tends to transform the original data so that the result of data mining task should not defy privacy constraints. Following is the list of five dimensions on the basis of which different PPDM Techniques can be classified

1) Data distribution
2) Data modification
3) Data mining algorithms
4) Data or rule hiding
5) Privacy preservation

The first dimension is related to distribution of data: Centralized or Distributed. The second dimension refers to the modification of original values of data that are to be released for data mining task .The third dimension is that of data mining algorithms. The data mining algorithm are applied on the transformed data to get useful nuggets of information that were hidden previously. The fourth dimension refers to whether the raw data or aggregated data should be hidden. The fifth and the final dimension refers to the techniques that are used for protecting privacy based on these dimensions, different PPDM techniques may be classified into following five categories

1) Cryptography based PPDM
2) Perturbation based PPDM
3) Randomized Response based PPDM
4) Anonymization based PPDM
5) Condensation approach based PPDM

The PPDM techniques can be broadly classified on two scenarios: the centralized server and distributed scenario. The distributed scenario is the cryptography based PPDM and central server or published scenario is the annonymization based PPDM. This paper explores various PPDM techniques based on cryptography and we just open an overview of annonymization techniques in different areas like Social Networking and Health Care.

IV. Cryptography based PPDM

Data mining can extract important knowledge from large database - sometimes this database is split among various parties. Here, the main aim of privacy preserving data mining is to find the global mining results by preserving the individual sites private data/information .Consider a scenario in which two or more parties owning confidential databases wish to run a data mining algorithm on the union of their databases without revealing any unnecessary information. In particular, although the parties realize that combining their data has some mutual benefit, none of them is willing to reveal its database to any other party.

Defining privacy- The common definition of privacy in the cryptographic community limits the information that is leaked by the distributed computation to be the information that can be learned from the designated output of the computation. A protocol that is run in order to compute the function does not leak any unnecessary information.

A. Privacy Preserving

Privacy preserving protocols are designed in order to preserve privacy even in the presence of adversarial participants that attempt to gather information about the inputs of their peers. There are, however, different levels of adversarial behavior. Cryptographic research typically considers two types of adversaries:

(a) Semi-honest adversary
(b) Malicious adversary

It is of course easier to design a solution that is secure against semi-honest adversaries, than it is to design a solution for malicious adversaries.

B. Pivacy Preserving Computation

In this section we will describe the various computation techniques which we are using for data.

1) Classification
2) Perturbation based PPDM
3) Randomized Response based PPDM
4) Mining Association Rules
5) Data Generalization,Summarization and Characterization
6) Profile Matching
7) Fraud Detection

C. Secure Computation and Privacy Preserving Data Mining

There are two distinct problems that arise in the setting of privacy-preserving data mining. The first is to decide which functions can be safely computed, where safety means that the privacy of individuals is preserved. This question of privacy-preserving data mining is actually a special case of a long-studied problem in cryptography called secure multi party computation. This problem deals with a setting where a set of parties with private inputs wish to jointly compute some function of their inputs.

D. Cryptography : Oblivious Transfer

Oblivious transfer is a basic protocol that is the main building block of secure computation. Oblivious transfer is often the most computationally intensive operation of secure protocols, and is repeated many times. Each invocation of oblivious transfer typically requires a constant number of invocations of trapdoor permutations.

Yao's presented a constant-round protocol for privately computing any probabilistic polynomial-time function (where the adversary may be either semi-honest or malicious). Yaos protocol is obviously too costly. On the other hand, a specialized protocol can be designed for computing this algorithm, which uses Yaos protocol as a primitive. Yaos protocol works by having one of the parties (say Alice) first generate an encrypted or garbled circuit computing f and send its representation to Bob. The overhead of the protocol involves:

1) Alice and Bob engaging in an oblivious transfer protocol for every input wire of the circuit

2) Alice sending to Bob tables of size linear in the size of the circuit

3) Bob computing a pseudo-random function a constant number of times for every gate

The computation overhead is dominated by the oblivious transfer stage, since the evaluation of the gates uses pseudo-random functions which are very efficient compared to the oblivious transfer protocol. A common belief with regard to Yaos protocol is that it is inherently inefficient, since it uses a circuit representation of the function.

E. The Multi Party Case

In the multi-party scenario, there are protocols that enable the parties to compute any joint function of their inputs without revealing any other information about the inputs.In this, some additional drawbacks, compared to the two-party cases.

In multi party case, the number of communication rounds depends on the depth of the circuit, but in two-party case , the number of rounds is constant.

It is impossible to require all pairs of parties to communicate. it is impossible to ensure that the number of corrupt parties is smaller than such a threshold.

In such cases the security of the protocol is not guaranteed. Privacy preserving multi-party computation can be reduced to the two-party case.The advantages of this approach are :

a) Trust
b) Independence of Inputs
c) Communication
d) Privacy
e) Correctness
f) Efficiency
g) Guaranteed Output Delivery
h) Fairness

V. ANONYMIZATION BASED PPDM

In privacy preserving data publishing, in order to prevent privacy attacks, data should be anonymized properly before it is released. Anonymization methods should take into account the privacy models of the data and the utility of the data. Generalization and perturbation are the two popular anonymization approaches for relational data .Anonymization approach uses several methods and approaches to provide privacy in the publishing scenario. Here we consider various anonymization approaches on two different publishing scenarios, health care and social networks.

A. ANONIMIZATION TECHNIQUES for PRIVACY PRESERVING in SOCIAL NETWORKS

Privacy preservation on relational data has been studied extensively. A major category of privacy attacks on relational data is to re-identify individuals by joining a published table containing sensitive information with some external tables modeling background knowledge of attackers. To battle the re-identification attacks, the mechanism of k-anonymity was proposed. Although k-anonymity has been well adopted, but later it showed that a k-anonymous table may still have some subtle but severe privacy problems due to the lack of diversity in the sensitive attributes.

we categorize the state-of-the-art anonymyzation methods on social network data into two categorizes as follows:

1) Clustering-Based Approaches:: A clustering-based method clusters vertices and edges into groups and anonymyzes a subgraph into a super-vertex. In such a way, the details about individuals can be hidden properly. The methods in this category can be further divided into vertex clustering methods, edge clustering methods and vertex and edge clustering methods.

Generalization is a popular way to anonymize relation data. Essentially, generalization can be regarded as clustering vertices and edges into groups and generalize all members in a group to the same.

2) Vertex Clustering Methods: The anonymization technique proposed is a vertex clustering approach. It generalizes an input network by grouping vertices into partitions and publishing the number of vertices in each partition along with the densities of edges within and across partitions. Data analysts can still use the anonymized graphes to study macro-properties of the original graph. The partitioning of vertices is chosen such that the generalized graph satisfies the privacy preservation goals and maximizes the data utility. To ensure anonymity, we need to make sure that any adversary has at least a minimum level of uncertainty about the re-identification of any target vertex.

3) Edge Clustering Methods: In general, a social network can have different types of vertices and different types of edges. Among all types of edges, one type is assumed sensitive and should be protected against link re-identification attacks. The privacy breach is measured by counting the number of sensitive edges that can be inferred from the anonymized data. Another anonymization strategy is to remove some observed edges. Generally, a particular type of observations which significantly contributes to the overall likelihood of a sensitive relationship, or a certain percentage of observations

that meet some pre-specified criteria (for example, at random,connecting high-degree vertices, etc.) can be removed. The most conservative anonymization strategy is to remove all edges in the network. Assumes that the vertices are divided into equivalence classes and each class is anonymized properly using some existing relational data anonymization method. Then, a more effective approach to anonymize the social network is to collapse all vertices in an equivalence class into a single vertex, and consider which edges to be included in the collapsed graph. One feasible way is to publish for each edge type the number of edges of the type between two equivalence class vertices.This approach is called cluster-edge anonymization.

4) Vertex and Edge Clustering Methods: To protect privacy in social network data, using the k-anonymity model. Every vertex should be indistinguishable with at least another (k 1) vertices in terms of both the attributes and the associated structural information such as neighborhood of vertices. The anonymization method disturbs as little as possible the social network data, both the attribute data associated to the vertices, and the structural information. The method for anonymizing vertex attribute data uses generalization, which has been well studied in relational data. For structure anonymization, the proposed method is called edge generalization, which is similar to the one described in to some extent. The critical difference is that the method in takes into account both the generalization information loss and the structural information loss during the clustering procedure. This process can be tuned by users to achieve a desirable trade off between preserving more structural information of the network and preserving more vertex attribute information. vertices's are partitioned into clusters in anonymization. To anonymize edges,vertices in the same cluster are collapsed into one single vertex, labeled with the number of vertices and edges in the cluster. The edges between two clusters are collapsed into a single edge, labelled with the number of edges between them.

5) Graph Modification Approach: The clustering-based approaches reduce a cluster of vertices and edges into a super-vertex. Thus ,the graph may be shrunk considerably after anonymization, which may not be desirable for analyzing local structures. To preserve the scale and the local structures of the original graph, graph modification approaches try to locally modify the graph structure to achieve the privacy preservation requirement.

6) Randomized Graph Modification Approach: To anonymize,modifies a graph by randomly adding and deleting edges.consider a graph having specific degree sequence having of vertices in a descending order. A degree sequence is K-degree anonymous if, for each vertex,there are at least other vertices carrying the same degree.By providing a privacy parameter, the anonymization method proceeds as follows:

In first step,starting from the original degree sequenced d, develops a dynamic programming method to construct a new degree sequence that is K-degree anonymous which minimizes the degree anonymization cost.In the second step, constructs a graph with new degree sequence.

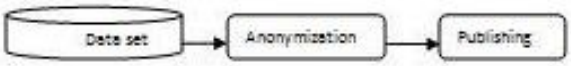


Fig. 1. Privacy Model

B. ANONYMIZATION TECHNIQUES for PRIVACY PRESERVING in HEALTH CARE

Privacy preserving in medical field means protection of individuals from being associated with undesirable conditions, diagnoses and all sensitive information. Although medical researchers often describe their research plans when they request anonimyzed.

Even though Health Data mining has this much of applications, it affects individuals privacy. Many organizations publish these medical data. The outside vendors and the insurance person can sell this data for various companies as commodity. Also the health data can be sold by the person who can access the cloud where this data is stored [1]. This may affect individuals privacy. Therefore preserving privacy is important. To publish the medical data without affecting the privacy we need to anonimyze the medical data. We have to anonimyze the data before publishing.

Data mining is exploring of large quantities of data and analyses it into understandable patterns, many factors have motivated the used of data mining applications. Privacy has always been a great concern of patients and medical service providers. As a result of recent advances in information technology and Government push for Electronic Health Record (EHR) systems, a large amount of data is collected and stored electronically. Health data mining is a process of extracting previously known information from a large volume of health data. Main aim of health mining is to improve patients care. Health data mining helps in fraud detection and abuse , decision management in customer relation best treatments and practices.

Privacy preserving in medical field means protection of individuals from being associated with undesirable conditions, diagnoses and all sensitive information. Although medical researchers often describe their research plans when they request anonymized. Even though Health Data mining has this much of applications, it affects individuals privacy. Many organizations publish these medical data. The outside vendors and the insurance person can sell this data for various companies as commodity. Also the health data can be sold by the person who can access the cloud where this data is stored [1]. This may affect individuals privacy. Therefore preserving privacy is important. To publish the medical data without affecting the privacy we need to anonymize the medical data. We have to anonymize the data before publishing. The fig.1 shows the basic model for privacy.

1) **Data anonymization :** Data anonymization is a type of information sensitization whose intent is privacy protection.

It is a process of either encrypting or removing personally identifiable information from datasets, so that the people whom the data describe remain anonymous. Anonymization is a cognitive process, practitioners must understand what could lead to the identification of an individual besides the obvious; direct data access either physically or legally carelessness etc.

2) **Re-identification Method:** Each person has some natural identifies, ie, data which characterizes an individual; name, social security number, passport number, phone number etc. some of them may not identify a person uniquely. There are a set of natural identifiers called indirect identifiers which together provide a unique identification. eg. Birth date. Mother's name.

In practice, direct and indirect identifiers are replaced with one way hash functions, i.e, functions that cannot allow original data to be restored since they have no inverse. Such a non reversible value replacement of direct identifiers called de-identification method.

Possible Re-Identification strategies:

1) Direct re-identification method:themselves without any further action reveal data subject identity.
2) Re-identification through linking:data set is believed to be de-identified while using publicly available.
3) Publishing anonymization algorithms or settings for predictive algorithms
4) Re-identification through extremities
5) Background knowledge based re-identification
6) Re-identification through event sequencing information misuse

The following strategies have been applied for anonymization:

1) Access limitations
   a) Limitation of data access : the most common procedure to limit the number of queries and to be run over a controlled environment through proper authentication and information hiding.
   b) Ciphering algorithms : change of data values in a way that makes impossible to retrieve original values
2) Obfuscation : these algorithms cut or aggregate parts of the databases I order to avoid re-identification
   a) Dynamic sampling : limited number of data elements are published, which meet the functional anonymization criteria
   b)Aggregation oriented anonymization : enforcing data aggregations and micro aggregation to achieve functional anonymization.
3) Functional anonymization : it aims at reducing the confidence about a piece of information related to specific individual.

3) **Problem statement:** Functional data anonymization techniques:

Let us consider Table 1 as a database to illustrate anonymization and related problems. The relation itself currently contains one sensitive information, the list of the salaries. Additionally, it can be directly connected with individuals as the ID attribute is present. Publishing such database could be strongly resisted with respect to data protection. Although this database is not The easiest de-identification is omitting ID column; the result can be seen in Table 2. Note that, selecting any two of the Date of birth, Sex and Postal code attributes can identify the set of individuals in that relation. In general, these attribute pairs are not enough for unique identification,

however in different countries they can identify a very large percent of the whole population , i.e. they quasi identify people.

TABLE I
EXTRACT FROM A DATABASE

ID	DOB	sex	PinCode	SALARY
Alice	21-01-82	Female	2201	40000
Bob	24-03-82	Male	2201	45000
Catherin	27-02-82	Female	2227	50000
Diya	21-01-82	Female	2227	55000
Ethu	24-03-82	Male	2237	60000
Freddi	27-02-82	Male	2237	65000

TABLE II
DE-IDENTIFIED DATA

DOB	sex	PinCode	SALARY
21-01-82	Female	2201	40000
24-03-82	Male	2201	45000
27-02-82	Female	2227	50000
21-01-82	Female	2227	55000
24-03-82	Male	2237	60000
27-02-82	Male	2237	65000

TABLE III
DE-IDENTIFIED DATA

DOB	sex	PinCode	SALARY
#	#	2201	40000
24-03-82	Male	2201	45000
27-02-82	Female	2227	50000
#	#	#	#
24-03-82	Male	2237	60000
27-02-82	Male	2237	65000

By using background knowledge on quasi-identifier and by having information about individuals from public sources, researchers can join records on quasi-identifier to the sensitive data items. A solution for this problem is to make data values ambiguous. Either one can delete some of the quasi-identifier and/or sensitive data values as shown in Table 3

VI. CONCLUSION

Cryptographic protocols for secure computation achieved remarkable results: it was shown that generic constructions

can be used to compute any function securely and it was also demonstrated that some functions can be computed even more efficiently using specialized constructions. Still, a secure protocol for computing a certain function will always be more costly than a naive protocol that does not provide any security. By making use of cryptographic techniques to store sensitive data and providing access to the stored data based on an individual's role, we ensure that the data is safe from privacy breaches.

we surveyed a few recent studies on anonymization techniques for privacy preserving publishing of social network data. Although privacy preserving data publishing and analysis techniques in relational data have been well explored, the research and development of anonymization techniques on social network data is still in its infancy. We reviewed the anonymization methods for privacy preservation in two categories: clustering-based approaches and graph modification approaches.As social network data is much more complicated than relational data, privacy preserving in social networks is much more challenging and needs many serious efforts in the near future. Particularly, modeling adversarial attacks and developing corresponding privacy preservation strategies are critical.

This paper also introduces the new method for preserving privacy in health data mining. The proposed system has mainly focuses anonimzation, Clustering by this preserving the privacy of health care sensitive data. Our system extracts the information from medical data. The data is completely extracted and the required data is obtained. The extracted data is clustered. Then data anonymized in order to ensure privacy.

REFERENCES

[1] Majid Bashir Malik ,M. Asger Ghazi,Rashid Privacy Preserving Data Mining Techniques: Current Scenario and Future Prospects 2012 Third International Conference on Computer and Communication Technology

[2] Xinjun Qi , Mingkui Zong ,2011 International Conference on Environmental Science and Engineering (ICESE 2011) An Overview of Privacy Preserving Data Mining

[3] Alex Gurevich , Ehud Gudes Privacy preserving Data Mining Algorithms without the use of Secure Computation or Perturbation

[4] J. Han, M. Kamber. Data Mining: Concepts and Techniques", Morgan Kaufmann Publishers

[5] Elisa, B., N.F. Igor and P.P. Loredana. A Framework for Evaluating Privacy Preserving Data Mining Algorithms, Published by Data Mining Knowledge Discovery, 2005, pp.121

[6] Brinal Colaco,Shamsudin S Khan-"Privacy Priverving Data Mining for Social Networks"-2014 International Conference on Advances in communication and computing technologies.

[7] Bin Zhou, Jain Pei"Priverving Privacy in Social Networks against Neighbourhood Attacks."

[8] Mingxuan Yuan,Lei Chen,Philip s Yu "Presonalised Privacy Protection in Social Networks"

[9] Jisha Jose Panackal, Anitha S Pillai, V N Krishnachandran Disclosure Risk of Individuals: a k-anonymity Study on Health Care Data Related to Indian Population, IEEE International Conference on Data Science and Engineering (ICDSE), 2014.

[10] Jisha Jose Panackal, Anitha S Pillai Adaptive Utility-based Anonymization Model for Privacy Preserving Data Mining, International Conference on Data Mining and Warehousing (ICDMW), Data Mining Algorithms, Elsevier, 2014.

[11] Soy M.S,Gyatri K.S,Ashwini.B "Privacy Preserving Health Data Mining",IJCST Vol.6,Issue 4,Oct-DEC 2015

[12] Li Liu, Murat Kantarcioglu and Bhavani Thuraisingham, The applicability of the perturbation based privacy preserving data mining for real-world data, Data and Knowledge Engineering, 2008.

[13] Samarati P, Protecting respondents privacy in Microdata release, IEEE Transactions on Knowledge and Data Engineering, 13:10101027

[14] Tiancheng Li, Ninghui Li, Towards Optimal k-anonymization, Data and Knowledge Engineering, Elsevier, 2008.

[15] Marina Blanton, Achieving Full Security in Privacy-Preserving Data Mining, IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing, 2011.

[16] R. Mukkamala and V.G. Ashok, Fuzzy-based Methods for Privacy-Preserving Data Mining, Eighth International Conference on Information Technology: New Generations, 2011.

[17] Jisha Jose Panackal and Anitha S Pillai, An Intelligent Framework for Protecting Privacy of Individuals Empirical Evaluations on Data Mining Classification, Proc. of International Conference on Hybrid Intelligence Systems (HIS), IEEE, 2014.

Different E-payment Systems : An Introduction

Anju Pius, Hafsa M H, Taniya Ignatious
Department of Computer Applications
Vidya Academy of Science and Technology
Thrissur - 680501

Dijesh P
Associate Professor of Computer Applications
Vidya Academy of Science and Technology
Thrissur - 680501

***Abstract*—The human never stop to build up new things. Day by day humans are adding new technology and improvement in the world. While electronic commerce (e-commerce) continues to have a deeply impact on the global business environment, technologies and applications have begun to focus more on more. Now a days customers and businesses are using the Internet to conduct business and to run it which gives an incredible growth for e-commerce. E-commerce engages different kinds of phases like WWW, digital rights management, security, privacy issues and electronic payment systems. Now days in market areas there are different types of electronic payment systems available.**

In this paper, we will discuss the different types of payment system which are currently being used in the market-places. It presents the working of each payment systems in a brief manner.

***Index Terms*—MMID, Wallets,IFSC,Octopus card,Mondex card**

I. Introduction

Commerce is a part of business which is concerned with the exchange of goods and services and include all those activities which directly or indirectly facilitate that exchange. In todays world there has been major changes to the commerce industry.The most important of it is the introduction of computers into commerce industry.

Payment started with barter system centuries ago. Goods were exchanged directly between the people in this method. But the major drawback of this system was that both the buyer and seller must be satisfied with the goods they had in surplus. This led to the next generation of payment method called Commodity Money System. Here, the buyer would bye goods from the seller in exchange of some commodity in the form of gold,silver,corn etc. Commodity money slowly evolved into standard of having paper notes. This method is simple and there is no bank involvement. It is mostly used for low-value payments. But this cash payment method is very insecure and there is no record of transaction maintained. Later the check payment method is employed for medium to high value payments and a record of transaction is maintained at the bank.

The introduction of computers into this field changed the scenario. Electronic commerce is defined as a monetary transaction that occurs electronically as opposed to the physical exchange of money or checks. In this method, tangible money is eliminated.

Electronic payment systems can be categorized into the following:

- Card-based payment system
- Electronic wallets
- Others

A. Card-based Payment Systems

1) Credit Card: Payment using credit card is one of most common mode of electronic payment. Credit card is small plastic card with a unique number attached with an account. It has also a magnetic strip embedded in it which is used to read credit card via card readers. When a customer purchases a product via credit card, credit card issuer bank pays on behalf of the customer and customer has a certain time period after which he/she can pay the credit card bill. It is usually credit card monthly payment cycle. Following are the actors in the credit card system. The card holder - Customer The merchant - seller of product who can accept credit card payments. The card issuer bank - card holder's bank The acquirer bank - the merchant's bank The card brand - for example , visa or mastercard.

Credit card payment process:-
Step 1- Bank issues and activates a credit card to customer on his/her request.
Step 2-Customer presents credit card information to merchant site or to merchant from whom he/she want to purchase a product/service.
Step 3-Merchant validates customer's identity by asking for approval from card brand company.
Step 4-Card brand company authenticates the credit card and paid the transaction by credit. Merchant keeps the sales slip.
Step 5-Merchant submits the sales slip to acquirer banks and gets the service chargers paid to him/her.
Step 6-Acquirer bank requests the card brand company to clear the credit amount and gets the payment.
Step 7-Now card brand company asks to clear amount from the issuer bank and amount gets transferred to card brand company.

Advantages of credit card are they allow us to make purchases on credit without carrying around a lot of cash,

they allow accurate record-keeping by consolidating purchases into a single statement, they allow convenient ordering by mail or phone, they allow you to pay for large purchases in small, monthly installments, under certain circumstances, they allow you to withhold payment for merchandise which proves defective. It also have some disadvantages. They are the ease of using credit cards, combined with impulsive buying, may result in over-spending. High interest rates, as well as other costs make credit cards a relatively expensive method of obtaining credit. Lost or stolen cards may result in some expense ($50 and inconvenience. The use of multi-credit cards can get you even further into debt. Fraudulent or unauthorized charges may take months to dispute , investigate, and resolve.

2) Debit Card: Debit card is a prepaid card and also known as ATM card. An individual has to open an account with the issuing bank which gives debit card with a personal id number, when he makes a purchase he enter his pin number on shop pin pad. When the card is slurped through the electronic terminal it dial the acquire a banking system either master card or visa card that validate the pin and finds out from the issuing bank whether to accept or decline the transaction the customer can never overspend because the system reject any transaction which exceeds the balance in his account the bank never face a default because the amount spent is debited immediately from the customer account.

3) Smart Card: Smart card was first introduce in Europe most of these method are known as stored value card.A smart card is about the size of a credit card, made of a plastic with an embedded microprocessor chip that holds important financial and personal information. The microprocessor chip is loaded with the relevant information and periodically recharged. In addition to these pieces of information, systems have been developed to store cash onto the chip. The money on the card is saved in an encrypted form and is protected by a password to ensure the security of the smart card solution. In order to pay via smart credit is necessary to introduce the card into a hardware terminal. The device requires a special key from the issuing bank to start a money transfer in either direction. Smart cards can be disposable or rechargeable.
Some of the advantages of smart cards include the following:

- Stored many types of information
- Not easily duplicated
- Not occupy much space
- Portable
- Low cost to issuers and users
- Included high security

The disadvantages of smart cards are the lack of universal standards for their design and utilization. On the other hand, smart card applications are expected to increase as a result of the resolution of these disadvantages in the near future.

4) Gift Card: A gift card (also known as gift certificate in North America, or gift voucher or gift token in the UK)is a prepaid stored-value money card usually issued by a retailer or bank to be used as an alternative to cash for purchases within a particular store or related businesses. Gift cards are also given out by retailers and marketers as part of a promotion strategy, to entice the recipient to come in or return to the store, and at times such cards are called cash cards. Gift cards are generally redeemable only for purchases at the relevant retail premises and cannot be cashed out, and in some situations may be subject to an expiry date or fees. Visa and MasterCard credit cards produce generic gift cards which need not be redeemed at particular stores, and which are widely used for cashback marketing strategies. A feature of these cards is that they are generally anonymous and are disposed of when the stored value on a card is exhausted.

5) Octopus Card: The octopus card is a reusable contactless stored value card smart card for making electronic payments in online or offline systems in Hong Kong. Launched in September 1997 to collect fares for the territorys mass transit system, the octopus card system is the second contactless smart card system in the world, after the Korean Upass, and has since grown into a widely used payment system for all public transport in Hong Kong, leading to the development of Oyster card in London. Octopus card has also grown to be used for payment in many retail shops in Hong Kong,from convenience stores, supermarkets, fast-food restaurants, street parking meters, car parks, to other point-of sale applications such as service stations and vending machines. Octopus card Limiteds and t he cards slogan is Making Everday Life Easier, which is also part of the corporations mission statement.

B. Electronic Wallets

E-wallet is a type of electronic card which is used for transactions made online through a computer or a smartphone. Its utility is same as a credit or debit card. An E-wallet needs to be linked with the individuals bank account to make payments. An E-wallet is a type of pre-paid account in which a user can store his/her money for any future online transaction. An E-wallet is protected with a password. With the help of an E-wallet, one can make payments for groceries, online purchases, and flight tickets, among others. E-wallet has mainly two components, software and information. The software component stores personal information and provides security and encryption of the data. The information component is a database of details provided by the user which includes their name, shipping address, payment method, amount to be paid, credit or debit card details, etc. For setting up an E-wallet account, the user needs to install the software on his/her device, and enter the relevant information required. After shopping online, the E-wallet automatically fills in the users information on the payment form. To activate the E-wallet, the user needs to enter his password. Once the online payment is made, the consumer is not required to fill the order form on any other website as the information gets stored in the database and is updated automatically.

1) AirtelMoney: With the AirtelMoney app, users can easily recharge prepaid and postpaid bills. You can also shop online if your digital Wallet has cash loaded in it.

2) Paytm: Launched in 2010, is currently the largest mobile Wallet app in India. Paytm Wallet is the digital

payment instrument where you can transfer money from your bank account or credit card to use for transaction.

Steps to use Paytm:-

1) Setup a Paytm account using your mobile number and email ID.
2) Add some money to your Paytm Wallet using netbanking, debit card, or credit card.
3) To transfer money to someone else, select the pay or send option.
4) You can make payments to others or to bank account on Paytm by scanning a QR code.
5) Alternatively, you can send money to another Paytm user via their phone number.

3) JioMoney: Launched recently in 2016 by Jio, is a digital payment app. With this, one can receive great discounts and offers. User can also bookmark their frequently visited retailers. So shopping can be made quicker than usual.

4) Yahoo!Wallet: This is an electronic wallet offered by web portal site Yahoo. Yahoo!wallet functions in the same way as most other electronic wallet- by completing order forms automatically with identifying information and credit card payment information. This wallet is used to pay for airplane tickets and hotel reservations booked through the yahoo.

C. Other Payment Systems

1) SWIFT: SWIFT stands for the Society for Worldwide Interbank Financial Telecommunications. SWIFT payments are a type of international transfer sent via the SWIFT international payment network. The SWIFT international payment network is one of the largest financial messaging systems in the world. TransferWise can send or receive certain currencies via SWIFT payment. SWIFT assigns each financial organization a unique code that has either eight characters or 11 characters. The code is called interchangeably the bank identifier code (BIC), SWIFT code, SWIFT ID, or ISO9362code.

- First four characters: the institute code
- Next two characters: the country code
- Next two characters: the location/city code
- Last three characters: optional, but organizations use it to assign codes to individual branches.

SWIFT payments usually take 1-3 working days to reach their destination, however it is possible that they can take longer due to circumstances such as time differences between the sending and receiving country.

2) NEFT: National Electronic Funds Transfer (NEFT) is a nation-wide payment system facilitating one-to-one funds transfer. Under this Scheme, individuals, firms and corporates can electronically transfer funds from any bank branch to any individual, firm or corporate having an account with any other bank branch in the country participating in the Scheme. For being part of the NEFT funds transfer network, a bank branch has to be NEFT- enabled. Even such individuals who do not have a bank account (walk-in customers) can also deposit cash at the NEFT-enabled branches with instructions to transfer funds using NEFT. However, such cash remittances will be restricted to a maximum of Rs.50,000/- per transaction. Such customers have to furnish full details including complete address, telephone number, etc. NEFT, thus, facilitates originators or remitters to initiate funds transfer transactions even without having a bank account. There is no limit either minimum or maximum on the amount of funds that could be transferred using NEFT. However, maximum amount per transaction is limited to Rs.50,000/- for cash based remittances and remittances to Nepal.
The working of NEFT system is given below:

Step 1-An individual / firm / corporate intending to originate transfer of funds through NEFT has to fill an application form providing details of the beneficiary (like name of the beneficiary, name of the bank branch where the beneficiary has an account, IFSC of the beneficiary bank branch, account type and account number) and the amount to be remitted. The application form will be available at the originating bank branch. The remitter authorizes his/her bank branch to debit his account and remit the specified amount to the beneficiary. Customers enjoying net banking facility offered by their bankers can also initiate the funds transfer request online. Some banks offer the NEFT facility even through the ATMs. Walk-in customers will, however, have to give their contact details (complete address and telephone number, etc.) to the branch. This will help the branch to refund the money to the customer in case credit could not be afforded to the beneficiarys bank account or the transaction is rejected / returned for any reason.
Step 2-The originating bank branch prepares a message and sends the message to its pooling centre (also called the NEFT Service Centre).
Step 3-The pooling centre forwards the message to the NEFT Clearing Centre (operated by National Clearing Cell, Reserve Bank of India, Mumbai) to be included for the next available batch.
Step 4-The Clearing Centre sorts the funds transfer transactions destination bank-wise and prepares accounting entries to receive funds from the originating banks (debit) and give the funds to the destination banks(credit). Thereafter, bank-wise remittance messages are forwarded to the destination banks through their pooling centre (NEFT Service Centre).
Step 5-The destination banks receive the inward remittance messages from the Clearing Centre and pass on the credit to the beneficiary customers accounts.

3) Unified Payments Interface (UPI):

- The United Payments Interface is a system for instant, electronic payments through your smart phone.
- It authenticates the identity of the user like a debit card does using the phone as a tool instead of a separate card.
- It is an advanced version of Immediate Payment Service (IMPS) which was used to transfer money between bank accounts.
- Like IMPS, UPI will facilitate round-the-clock funds

transfer service.
- It works 24x7, 365 days, unlike RTGS or NEFT services which have specific working hours.

Working of UPI is given below

Let us assume a person named Ram has to make payments using UPI (lets explain this with an example). To make UPI money transfer, Ram needs 2 basic things:

- A smartphone with UPI application (app)
- A bank account

Ram has to download the UPI app and get a UPI ID by registering on the app with his bank details. UPI ID is a virtual identity (a payment identifier) like an email address. It can be a name or a mobile number along with the name of your bank.

For example-

- Ram@sbi or Ram@icici
- 9900000099@hdfc or 9900000099@axis

The payment is verified instantly through the smart phone, without needing to rely on debit card payment or net banking. Ram has to buy a book online. He can initiate the e-commerce purchase by selecting UPI as the payment mode and providing his UPI ID Ram@sbi. He then receives a pop-up notification on his smartphone through the UPI App requesting confirmation of the payment. He has to enter his secure pin on the app to authenticate the purchase transaction. He will then receive a confirmation of a successful online purchase from the merchant within seconds.

4) Immediate Payment Service (IMPS): Immediate Payment Service (IMPS) is an instant interbank electronic fund transfer service through mobile phones. IMPS facilitate customers to use mobile instruments as a channel for accessing their banks accounts and remitting funds there from. The customer need to have a bank account for availing IMPS.

The benefits of IMPS are,

- Instant
- Available 24 x7 (functional even on holidays)
- Safe and secure, easily accessible and cost effective
- Channel Independent can be initiated from Mobile/ Internet / ATM channels ? Debit & Credit Confirmation by SMS

Account Number & IFSC Or Mobile Number & MMID required for the money transfer through IMPS.Mobile Money Identifier (MMID) is a seven digit unique number issued by the bank upon registration. Remitter (customer who wants to send money) and Beneficiary (customer who wants to receive the money) should have this MMID for doing this interbank funds transfer through Mobile Number and MMID.

5) PayPal: PayPal is a service that allows national and international money transactions through internet. The whole system of PayPal is based on email ID. First of all, you would register with PayPal using your email address as your user ID. You will be identified with your email address. All the time of registration, you will need IFSC and branch code of your banks branch. After registration is done, you will need to enter your bank or credit card details in your profile. To verify that you have entered the correct bank account details, PayPal will transfer one or two very amounts of money into your account. Once you receive these amounts, you would need to go back to your PayPal profile and enter the amounts you have received. If you enter correct values, your bank/credit card account would be considered verified. When you need to receive payment from someone-just tell that person your email ID which is registered with PayPal. When your payer has transferred money to your PayPal account, you will get an email informing that you have received this much amount from so and so person. You can also use your PayPal account to send money as well. You can go to your PayPal account and send the required amount to the email ID of the person whom you want to pay. PayPal will withdraw that amount from your bank account. Then PayPal will cut its fee and transfer rest of the money to the payee.

6) RTGS: Stands for Real Time Gross Settlement. It allow electronic transfer of funds from the remitter, who has an account in one bank, to the beneficiary, who has an account in any other bank/branch. The transfer can be carried out using the internet banking facility. Payer select beneficiary name, amount and the reason of transfer. On submission of details and security transaction password, the transfer instruction is processed. RTGS transfer is carried out on a real-time basis. The minimum amount that can be transferred by RTGS is Rs.2lakh and there is no upper limit.

II. Conclusion

Most of payment systems described above offer a secure means directly related to transfer credit/debit details for settlement in the existing financial systems. This also suffers from transaction processing costs, ensuring that low value transactions cannot be cost-effective. Well known institutions are able to aid in EPS (electronic payment system) adoption through the provision of a large installed base of customers. This study has also found that these institutions play other crucial roles in EPS adoption. Large partners are able to provide EPS with association with trusted brand names and marketing boom. These result in the system gaining credibility and public awareness. Once this has been achieved the system is assessed by users on factors such as simplicity, security and mutuality of stakeholder benefits.

An electronic cash scheme, such as visa, Mondex and PayPal offers the user the ability to pay retailers and other consumers on the Internet as well as in the high street, over the phone and in the home. The payment requires no other participants than the payer and payee, so by having no transaction processing fees and allowing low value transactions to be cost-effective. This uses inherent security mechanisms to ensure the safety of transactions independent of the transmission protocol being used.

E-commerce on the Internet needs payment mechanisms that can serve for as much diversity as commerce in the real world. Large value transactions will require secure ways to use existing bank card mechanisms.

Characteristics of payment system includes:

Applicability:The usefulness of a payment mechanism is dependent upon what one can buy with it. Applicability (or acceptability, as it often mentioned in literature) of a payment system is defined as the extent to which it is accepted for payments. For instance, cash is accepted widely and thus has high level of applicability. Applicability of a payment system may vary from country to country. Quite high applicability have debit cards (bankcards) and credit cards, while cheques are not longer common in several European countries.

Ease of use:Paying with an electronicpayment system should not be a complex task; we will call this characteristic ease of use or usability. Usability is an important characteristic and defined as the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use,. Payments should be automated and done in an easy, seamless way. In such a responsible task as a payment process users should have minimum factors that make it difficult to pay or distract them.

Security:Since Internet services are provided today on the open networks, the infrastructure supporting electronic commerce, and payment systems in particular, must be resistant to attacks in the Internet environment. For electronic cash systems the issue of security has a special angle of counterfeiting, which means that no one should be able to produce electronic tokens on their own. Another angle is double spending; design should ensure that electronic tokens couldnt be spent twice.

Reliability:Naturally, users would like to see that system is reliable, because the smooth running of an enterprise will depend on the availability of the payment infrastructure. Scalability:As the commercial use of the Internet grows, the demands placed on payment infrastructure will also increase. The payment infrastructure as a whole should be scalable, to be able to handle the addition of users and merchants, so that systems will perform normally without performance degradation. Among the least scalable systems are those that require from users and vendors purchase and installation of additional hardware; this often hampers development of electronic cash systems.

Interoperability:A payment system is interoperable if it is not dependent on one organization, but is open and allows as many as necessary interested parties to join. This can be achieved by means of open standards for the technology that is used. It is natural, though, that companies that implement new technologies treat them as knowhow, because of the added value they create by investing in the technologies; therefore it is not always sensible to demand interoperability. Examples of interoperable initiatives are the CAFE project,and SEMPER project.

REFERENCES

[1] Paul and Sarah Edwards,Linda Rohrbough,*Making Money in Cyberspace*
[2] Henry Dreifus,J.Thomas Monk,*Smart Cards*
[3] Donal O'Mahony,Michael Peirc,Hitesh Tiwari,*Electronic Payment Systems*
[4] Jack. M.Kaplan,*Smart Cards,The Global Information Passport*
[5] Asokan.N, Janson .P.A,Steiner.M and Waidner.M,*The state of the Art in Electronic Payment Systems*,1997
[6] Low.S.H,Maxemchuk.N.F and Paul.S,*Anonymous Credit Cards*,1994
[7] Wayner.P,*Digital cash.:commerce on the net*,1997
[8] Srivalli Arkalgud,"Electronic Payment Systems"

Big Data Privacy: Challenges to Privacy Principles and Models

Aswathy A S, Aswathy Rajan, Aswathi C A
Department of Computer Application
Vidya Academy of Science and Technology
Thrissur-680501

Jisha Jose Panackal
Associate Professor of Computer Application
Vidya Academy of Science and Technology
Thrissur - 680501

***Abstract*—This paper explores the challenges raised by big data in privacy-preserving data management.First,we see the definition of big data then examine the conflicts raised by big data such as consent, purpose limitation, individual rights of access, rectification and erasure. Anonymization appears as the best tool to remove such conflicts. For this reason we evaluate how well the two privacy models used in anonymization (k-anonymity and -differential privacy) meet the requirements of big data, namely composability, low computational cost and linkability. Big Data is gaining more and more attention since the number of devices connected to the so called Internet of Things (IoT) is still increasing to unforeseen levels, producing large amount of data which needs to be transformed into valuable information.**

***Index Terms*—Big data, Consent, Privacy models, k-anonymity, -differential privacy,IOT,Health Care,User Profiling,Social Media**

I. Introduction

Big data have become a reality in recent years: The term big data refers to the collection of large and complex data sets, which exceeds existing computational, storage and communication capabilities of conventional methods or systems. Data are being collected from multiple independent sources, and then they are fused and analyzed to generate knowledge.Big Data may well be the next big thing in the IT world.Big Data concept come in the first decade of 21st century. The first organisations to embrace it were online and start up firms. Firms like Google, eBay, Linkedln and facebook were built around big data forms the beginning. Big Data may well be the next big thing in the IT world. Big Data concept come in the first decade of 21st century. The first organisations to embrace it were online and start up firms. Firms like Google, eBay, Linkedln and facebook were built around big data forms the beginning.

A. Aspects of Big Data

Big data includes three aspects, that is 3Vs:

1) **Volume**
   Organizations collect data from variety of sources, including business transactions, social media etc.
2) **Variety**
   Data comes in all forms both structured and unstructured.
3) **Velocity**
   This refers to the speed at which data arrives from different sources, the speed at which data flows inside the systems and how fast the data is processed.

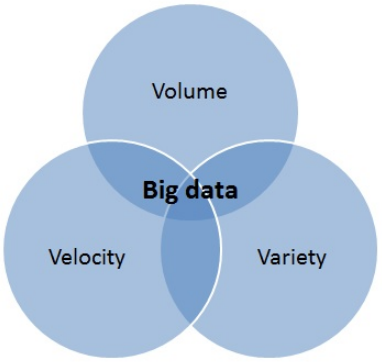


Fig. 1. Three Aspects of Big Data

Main benefits of big data is that it helps for cost reductions, time reductions, smart decision making.Many new information technology, big data can bring about dramatic cost reduction substantial improvements in the time required to perform a computing task, or new product and service offerings. Most traditional data sources are structured, many sources of big data are semi-structured and video data, audio data are unstructured. Growth of big data is needed increase of storage capacities, increase of processing power and availability of data(different data types).

B. Complexity of big data

In todays digital world, where lots of information is stored in big datas, the analysis of the databases can provide the opportunities to solve big problems of society like healthcare and others. As the database contains personal information, it is vulnerable to provide the direct access to researchers and

analysts. Since in this case, privacy of individuals is leaked, it can cause threat and it is also illegal. This is an important side effect of big data.

C. Two privacy models of bigdata

- **k-Anonymity**
  k-Anonymity seeks to limit the disclosure risk of a data set by limiting the capability of intruders to re-identify a record, that is, k-anonymity seeks to prevent identity disclosure . To that end, k-anonymity assumes that record re-identification is performed based on a fixed set of combinations of attributes. Each such combination is known as a quasi-identifier. This data seems harmless as it cannot reveal the identity of individual capacity. But these attributes can be combined with some external informations to uniquely identify an individual.
- **Evaluating Differential Privacy**
  Differential privacy is a privacy model offering strong privacy guarantees. It seeks to limit the impact of any single individual subjects contribution on the outcome of any analysis. Its primary setting is based on a trusted party that holds the original data set, receives queries and returns randomized answers for those queries so that the following differential privacy guarantee is satisfied. Several approaches have been proposed to generate differentially private data sets, although the main purpose of differential privacy remains to provide privacy-preserving answers for queries performed on the original data..

In this paper, we have highlighted the security and privacy issues in big data technology. Section 2 includes the related work based on this paper. Third section includes the big data privacy issues in various domain such as, Health care, Social media, User profiling and also add the new technology Internet of Things .

II. Related Works

With the ever-increasing cost for healthcare and increased health insurance premiums, there is a need for proactive healthcare and wellness. In addition, the new wave of digitizing medical records has seen a paradigm shift in the healthcare industry. As a result, the healthcare industry is witnessing an increase in sheer volume of data in terms of complexity, diversity and timeliness. As healthcare experts look for every possible way to lower costs while improving care process, delivery and management, big data emerges as a plausible solution with the promise to transform the healthcare industry. This paradigm shift from reactive to proactive healthcare can result in an overall decrease in healthcare costs and eventually lead to economic growth. While the healthcare industry harnesses the power of big data, security and privacy issues are at the focal point as emerging threats and vulnerabilities continue to grow. In this paper, we present the state-of-the-art security and privacy issues in big data as applied to healthcare industry. **[3]**

Big data create values for business and research, but pose significant challenges in terms of networking, storage, management, analytics and ethics. Multidisciplinary collaborations from engineers, computer scientists, statisticians and social scientists are needed to tackle, discover and understand big data. This survey presents an overview of big data initiatives, technologies and research in industries and academia, and discusses challenges and potential solutions.**[9]**

As we are moving towards the Internet of Things (IoT), the number of sensors deployed around the world is growing at a rapid pace. Market research has shown a significant growth of sensor deployments over the past decade and has predicted a significant increment of the growth rate in the future. These sensors continuously generate enormous amounts of data. However, in order to add value to raw sensor data we need to understand it. Collection, modelling, reasoning, and distribution of context in relation to sensor data plays critical role in this challenge. Context-aware computing has proven to be successful in understanding sensor data. In this paper, we survey context awareness from an IoT perspective. We present the necessary background by introducing the IoT paradigm and context-aware fundamentals at the beginning. Then we provide an in-depth analysis of context life cycle. We evaluate a subset of projects (50) which represent the majority of research and commercial solutions proposed in the field of context-aware computing conducted over the last decade (2001-2011) based on our own taxonomy. Finally, based on our evaluation, we highlight the lessons to be learnt from the past and some possible directions for future research. **[12]**

III. Big Data Privacy Issues in Various Domains

A. Big Data Security and Privacy Issues in Healthcare

The term big data refers to the collection of large and complex data sets, which exceeds existing computational, storage and communication capabilities of conventional methods or systems. In healthcare several factors affects the power of big data. With the increasing cost for healthcare services and increased health insurance premiums, there is a need for proactive healthcare management and wellness. This shift from reactive to proactive healthcare can result in improved quality of care, decrease in healthcare costs, and eventually lead to economic growth.

Reactive healthcare involves reacting to an adverse disease, injury or symptom. For example, if you have a fever and body ache, you may consult the Doctor. Depending on the Doctors diagnosis, he or she may prescribe you with antibiotics to help your body fight the infection. Both you and Doctor are reacting to symptoms.

Proactive healthcare takes actions before symptoms manifest. Prevention is better than cure. For example, rather than until you feel cold, you can take proactive approach in boosting your immune system with Vitamin C, drinking plenty of fluids.

Healthcare is moving towards big data, with patient information residing in multiple locations that must be accessed rapidly. With the increasing cost for healthcare and increased healthinsurance premiums, there is a need for proactive healthcare. Healthcare digitization with integrated analytics is one

of the next big waves in healthcare Information Technology (IT) with Electronic Health Records (EHRs) being a crucial building block. As healthcare experts look for every possible way to lower costs while improving care process, delivery and management, big data emerges as a plausible solution. This results in an overall decrease in healthcare costs and eventually lead to economic growth. Thus big data security and privacy is vital. This paper presents the security and privacy issues in big data as applied to healthcare industry.

Healthcare digitization with integrated analytics is one of the next big waves in healthcare Information Technology (IT) with Electronic Health Records (EHRs) being a crucial building block. EHR stores patients medical history electronically. It facilitates better access to complete, accurate and sharable healthcare data, that eventually lead to improved patient care. This helps

- to reduce the cost overhead,curing diseases,improving profits,predicting epidemics and enhancing the quality of human life by preventing deaths;
- to identify the threats and issues through comparison;
- to identify the problems before they happen.
- doctors to assess likely result of whichever treatment is considering prescribing, backed up by the data from other patients with the same condition, genetic factors and lifestyles.

As the healthcare industry witnesses large volumes of data, the first step will involve governance and linking accurate and actionable data in realtime. In this age of connectivity, integrating health systems with large amounts of clinical, financial, genomic, social and environmental data will be crucial for real-time analytics and patient care. The goal is to understand population health for disease control and predictive analysis. For instance, predictive analysis can help understand aggravating health conditions and could prevent adverse health events from occurring (e.g. chronic diseases such as diabetes).

In addition, with the introduction of Body Sensor Networks (BSN) and their direct application to health care, care providers will be able to monitor vital parameters, medication effectiveness, and predict an epidemic. Wearable blood pressure monitors send data to smartphone app, this helps to predict an epidemic. With the ever-changing risk environment and introduction of new emerging threats and vulnerabilities, security violations are expected to grow in the coming years. Moreover, the Affordable Care Act (Obama care) will lead to more enrollments for health insurance [1], making it an attractive focal point for hackers and opening a floodgate of healthcare breaches in the coming years. Affordable Care Act aims on increasing health insurance quality, expanding insurance coverage and reduce costs of healthcare.

B. Security and Privacy in Healthcare

EHR security must be a high priority to ensure patient safety. Patient information is stored in data centers with varying levels of security. Moreover, most healthcare data centers have HIPAA certification(Health Insurance Portability and Accountability Act), but that certification does not guarantee patient record safety. HIPAA is more focused on ensuring security policies and procedures than on implementing them.

- **Data Governance**
  The goal is to have a common data representation that encompasses industry standards (e.g. LOINC, ICD, SNOMED, CPT, etc.) and local and regional standards. Data generated would require normalization, standardization and governance prior to analysis.
- **Real-time security analytics**
  The challenge is finding new patterns and associations to identify clues about attacks such as phishing.
- **Privacy-preserving analytics**
  Invasion of patient privacy is a growing concern in the domain of big data

C. Big Data Privacy Issues In Public Social Media

Big data is an extremely large data sets that may be analysed computationally to reveal patterns,trends,andassociations,especially relating to human behaviour and interactions processing or analysing the huge amount of data or extracting meaningful information is a challenging task.The privacy crictical big data applications lie in the new domains of the social web.This is very important issues. Since it is entirely up to the controller of the big data sets if the information obtained from various sources if used for criminal purposes or not .In the context of social web there is an increasing awareness of the value potential and risk of the personal data which users voluntarily upload to the web.

Social media privacy issue are very common and very dangerous for the users of these sites.Issues like spamming,hacking,scamming,phishing are effecting people very badly.After the November 2013,security breach where hackers stole usernames and passwords for nearly 2 million accounts at Facebook,Google,Yahoo,Linkedln,Twitter and 93,000 other websites.

- **Location Tracking**
  Location based social networks are part of what is called Location Based Services(LBS).They are made possible by linking Global Positioning System(GPS),which track users locstion,to the capabilities of the World Wide Web, along with other vital features such as instant messaging.
- **Privacy concerns regarding Social Networking**
  The issues of online privacy has been a serious problem for a long time,it has even started to grow rapidely due to technology.Criminals may use social networks to connect with potential victims.5

Big data highest privacy issues in Social sites like Facebook and twitter have the highest privacy issues.

1) Facebook*:*

- Facebook has over a billion active users .
- Teenagers of almost [18 to 35 age]facebook addicted.
- Fake Profile and IDs with fake names is one of the biggest promblems onfacebook.
- Strange people try to get into our profles.

- Accounts can be easily hacked by using different hacking websites.

2) Twitter*:*

- Twitter has mostly dodged privacy concerns.Facebook gets all the bad press,but the bigger threat to your online privacy thse days might be your Twitter account.
- Twitter allows people to share information with their followers .any messages that are not switched from the default privacy setting are public ,ad thus can be viewed by anyone with a twitter account.
- Users can make their timelines privat,but once a tweet is re-tweeted by another user ,it irreversibly becomes public.

D. A Discussion of Privacy Challenges in User Profiling with Big Data Techniques

User profiling is the process of collecting information about a user in order to construct their profile. The information in a user profile may include various attributes of a user such as geographical location, academic and professional background, membership in groups, interests, preferences, opinions, etc. Big data techniques enable collecting accurate and rich information for user profiles, in particular due to their ability to process unstructured as well as structured information in high volumes from multiple sou‘1rces. A user profile is a collection of information that describes the various attributes of a user. These attributes may include geographical location, academic and professional background, membership in groups, interests, preferences, opinions, etc. User profiling is the process of collecting information about a user in order to construct their user profile. User profiles are utilized by a variety of web based services for different purposes. One of the primary uses of user profiles is for recommendation of items, elements or general information that a user has not yet considered but may find useful. General purpose social networks such as Facebook.com use a user profile to find potential friends based on the existing relationships and group memberships of the user. Professional social networks such as LinkedIn.com exploit the skills and professional background information available in a user profile to recommend potential employees. Big data techniques are a collection of various techniques that can be used to discover knowledge in high volume, highly dynamic, and highly heterogeneous data. Big data techniques offer opportunities for user profiling that can result in very comprehensive user profiles.

- **User Profile Contents**
  The information contained in a user profile can be provided explicitly by the user or alternatively it can be either inferred or mined by the service that manages the profile. Gathering accurate, precise, and rich information is clearly the objective when building a user profile. user interests; user knowledge, user background and skills; user goals; user behavior; user individual characteristics
- **Interests**
  The information that can be recorded under this attribute includes a users professional interests, his interests in hobbies, his interests in entertainment such as music, cinema, books, etc., the user.
- **Knowledge, background and skills**
  This attribute can be used to quantify the knowledge of the user in a given domain.
- **Goals**
  The goals and intentions of a user represent what he wishes to achieve in a given context. Goals can be classified as short term and long term. A recommender system can try to predict the needs of a user given his short term and long term goals and intentions.
- **Behavior**
  Users often have repetitive behaviors that can be observed and stored in their user profiles.
- **Individual characteristics**
  The individual characteristics of a user that may be made part of their user profile include personal information such as age, gender, relationship status, address, etc. Knowledge of demographic information is useful information for a recommender system.
- **Context**
  The different types of contexts include environmental contexts, personal contexts, social contexts, and spatiotemporal contexts. Entities that are located in the vicinity of the user form his environmental context, e.g., things, services, temperature, light, humidity, noise, and persons. Personal context comprises of physiological contexts, such as weight, pulse, blood pressure, hair color, etc., as well as mental contexts, such as mood, stress level, etc. Social context can comprise of information such as friends, neighbors, co-workers, and relatives. Spatiotemporal information is a combination of time, location, and the direction of movement.

Big Data Techniques for User Profiling We list below some of the big data techniques that can be used for collecting information about a user and building a user profile. Many existing big data implementations are algorithms adapted for distributed computation platforms such as Hadoop (hadoop.apache.org).

- **Network analysis**
  Network analysis algorithms are used to discover relationships between the nodes in a graph or a network. Network analysis is particularly useful in the context of social networks where important information about the user such as his friends, co-workers, relatives, etc. can be discovered.
- **Sentiment analysis**
  Sentiment analysis is a natural language processing technique that aims to determine the opinion and subjectivity of reviewers. The Internet is replete with reviews, comments and ratings due to the growing popularity of web sites such as Amazon.com, Ebay.com, and Epinion.com where users provide their opinion on others users and items.
- **Trust and reputation management**

Trust and reputation management is a set of algorithms and protocols for determining the trustworthiness of a previously unknown user in the context of his reliability in performing some action. Trust and reputation information can be an important part of a user profile

- **Machine learning**
  Machine learning is a sub-field of artificial intelligence that aims to build algorithms that can make decisions not based on explicit programming but instead based on historical empirical data. An example often cited is the algorithmic classification of email into spam and non-spam messages without user intervention. In the context of user profiling, machine learning can be used for learning user behavior by identifying patterns. Topics in machine learning include: supervised learning approaches, e.g., neural networks, parametric/nonparametric algorithms, support vector machines, etc.; and unsupervised learning approaches, e.g., cluster analysis, reduction of dimensionality, etc.
- **Cluster analysis**
  Cluster analysis is the process of classifying users (or any other objects) into smaller subgroups called clusters given a large single set of users. The clusters are formed based on the similarity of the users in that cluster in some aspect. Cluster analysis can be applied for discovering communities, learning membership of users in groups, etc. Cluster analysis can be considered as a sub-topic of machine learning.

IV. Privacy Related Challenges

- **Providing privacy guarantees**
  At all levels within the system user privacy guarantees must be given. This is most likely one of the hardest tasks. Indeed, as soon as information flows out of a system, sensitive information leaks become a risk. Solutions which may seem trivial, such as anonymization have been shown to be inefficient. A well known example showing that simple anonymization is insufficient to protect privacy is the de-anonymization of the data of the Netflix contest [3]
- **Flexible privacy policies**
  Users are different, in particular with respect to privacy. Some may not have any privacy concerns at all where as others may not want to disclose a single piece of information about themselves.
- **Evaluating trust and reputation**
  What user profile information is disclosed, or at which granularity it is disclosed, may depend on the trust (with respect to privacy concerns) that the user and/or the EEXCESS system has in the content provider. Calculating a content providers reputation and trustworthiness in a privacy preserving manner is thus an issue.
- **Content anonymity**
  To guarantee privacy, the attacker should not be able to identify the user from the provided data. Therefore, the system should ensure that an attacker cannot deduce from the content of a request who it originated from.
- **Request unlinkability**
  If multiple queries can be linked together, even while having content-anonymity for each individual query, the combination of the two could reveal information about the user. Therefore, it should be required that the protocols guarantee that two independent requests originating from the same user are unlinkable.
- **Origin unlinkability**
  This should be feasible by anonymizing 29 the origin of the request but under the condition that the origin is not revealed by the application level protocols. Therefore, we also need to guarantee that the application level protocols are privacy-preserving (i.e. an attacker cannot link a given request to the requesting user).

V. Big Data Privacy in the Internet of Things Era

The Internet of Things (IoT) enables data collection on a large scale, but the extraction of knowledge from this data can lead to user privacy issues. The Internet of Things (IoT)1 is a network of networks in which a massive number of objects, sensors, or devices are connected through the ICT infrastructure to provide value-added services. The IoT connects people and things anytime, anyplace, with anything and anyone, ideally using any path or network and any service.

Big data has no clear definition,3 but it isnt wholly about size. Rather, its defined based on three primary characteristics, also known as the 3Vs: volume, variety, and velocity.4 Volume relates to the datas size (terabytes, petabytes, or zettabytes). Variety refers to the type of data and its source (sensors, devices, social networks, the Web, mobile phones, and so on). Velocity means how frequently the data is generated (for instance, every millisecond, second, minute, hour, day, week, month, or year) Privacy issues in the Internet age have received significant attention over the past few years. For example, allegations of governments spying on their citizens and new laws such as the right to be forgotten11 have opened up a whole range of debate. Compared to the Web era, the IoT is more vulnerable to privacy violations. Therefore, researchers as well as IT professionals will pay more attention to IoT technologies, business models, and potential regulatory efforts to ensure that more secure and privacy-preserving IoT data management techniques are developed.

- **User Consent Acquisition**
  In the IoT, user consent is about acquiring the required level of permission from users and nonusers who are affected by devices or services. In the traditional Web, the method of receiving user consent is through privacy terms and policies presented to users via long paragraphs of text. With the emergence of social media and mobile apps, consent-acquiring mechanisms have changed
  In some cases, developers might not provide accurate information to users for the consenting decision. In other cases, developers might provide accurate information, but users cant understand exactly what the consent entails due to a lack of technical knowledge. One major privacy

challenge in the IoT is to develop technologies that request consent from users in an efficient and effective manner. This is a challenging task because every user has limited time and technical knowledge to engage in the process. Such research will need to combine principles and techniques from humancomputer interaction and cognitive sciences.

- **Control, Customization and Freedom of choice**
  In the IoT, data owners must have full control of data and be able to delete or move data from one service provider to another at any time. Unfortunately, existing IoT solutions in the marketplace provide only limited access to users. Moreover, users should be able to choose hardware devices and software components from different vendors to build their smart environments (for example, a smart home). This gives users full control and freedom of choice. Consequently, users must decide on what kind of data they will share, with what access rights for service providers. Users should also have the ability to withdraw or change previous user consents. Its also important for users to understand that, without having access to some data types, a service provider wont be able to facilitate certain types of services. However, service providers must not unfairly treat consumers, such as by disabling certain features or changing subscription fees to motivate users to provide consent.
- **Promise and Reality**
  Each IoT solution promises to offer a select number of functionalities. Service providers achieve this by requiring certain types of raw data to be processed and analyzed. However, with the development of new technologies, businesses might be able to derive more knowledge from user-acquired data. However, if service providers want to use raw data to derive more knowledge, then they must explicitly request permission by explaining the new possibilities and potential consequences to users. The bottom line is that service providers must not use already collected data for any other purpose without explicit user consent. Both regulations and technology must be developed and put in place to avoid such a misuse.
- **Anonymity Technology**
  Network communication interfaces typically have media access control (MAC) addresses that can be used to trace data communication paths. Combining multiple devices MAC addresses will help create unique fingerprints and a unique profile in which analytics can be used to extract knowledge. Consequently, user location can easily be tracked. Its important to discover new technologies that can anonymize data communication paths to protect user privacy. Due to the usage of large numbers of sensors and services, anonymizing multidimensional data is challenging. In particular, its easy to build fairly unique profiles that might enable knowledge extraction for a specific user. Currently, network communication technologies dont preserve user anonymity. Newer IoT platforms will be required to adopt technologies such as Tor (www.torproject. org), which conceals user location. In essence, a comprehensive anonymization framework is required to facilitate end-to-end anonymity in the IoT. Such a framework must ensure anonymity at different levels, such as data modeling, storage, routing, communication, analytics, and aggregation.

VI. Conclusion

Big Data gradually become an emerging topic of research and business and has been everywhere used in many industries. Big Data security and privacy has been increasingly concerned. However, there is an obvious contradiction between Big Data security and privacy and the widespread use of Big Data. Security and privacy are among the most important requirements in Big Data.Our hope is that this paper will spur action in the research and development community to collaboratively increase focus on the challenges, leading to greater security and privacy in big data platform.

References

[1] Hundepool A, Domingo-Ferrer J, Franconi L, Giessing S, Nordholt ES, Spicer K, de Wolf P-P (2012) Statistical disclosure control. Wiley, New YorkCrossRefGoogle Scholar
[2] Danezis G, Domingo-Ferrer J, Hansen M, Hoepman J-H, Le Mtayer D, Tirtea R, Schiffner S (2015) Privacy and data protection by designfrom policy to engineering. Technical report, ENISA
[3] HK Patil, R Seshadri - Big Data (BigData Congress), 2014 IEEE , 2014 - ieeexplore.ieee.org
[4] "Public Law 111 - 148 - Patient Protection and Affordable Care Act," U.S. Government Printing Office (GPO), 2013.
[5] P. Jessup, Big data and targeted advertising, http://www.unleashedtechnologies.com/blog/2012/06/28/big-data-and-targeted-advertising, June 2012.
[6] J. Yap, User profiling fears real but paranoia unnecessary, http://www.zdnet.com/user-profiling-fears-real-but-paranoiaunnecessary-2062302030/, September 2011.
[7] A. Narayanan and V. Shmatikov, Robust De-anonymization of Large Sparse Datasets, in 2008 IEEE Symposium on Security and Privacy (sp 2008). IEEE, May 2008, pp. 111125.
[8] S. Schiaffino and A. Amandi, Intelligent user profiling, in Artificial Intelligence An International Perspective. Springer, 2009, pp. 193216.
[9] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, Big data: The next frontier for innovation, competition, and productivity, The McKinsey Global Institute, Tech. Rep., May 2011.
[10] E. Damiani, S. D. C. di Vimercati, S. Paraboschi, P. Samarati, and F. Violante, A reputation-based approach for choosing reliable resources in peer-to-peer networks, in Proceedings of the ACM Conference on Computer and Communications Security, 2002, pp. 207216.
[11] L. Atzori, A. Iera, and G. Morabito, The Internet of Things: A Survey, Computer Networking, vol. 54, no. 15, 2010, pp. 27872805.
[12] C. Perera et al., Context Aware Computing for the Internet of Things: A Survey, IEEE Comm. Surveys & Tutorials, vol. 16, no. 1, 2013, pp. 414454.
[13] A. Zaslavsky, C. Perera, and D. Georgakopoulos, Sensing as a Service and Big Data, Proc. Intl Conf. Advances in Cloud Computing (ACC), 2012, pp. 2129.
[14] C. Eaton et al., Understanding Big Data, McGraw-Hill, 2012.
[15] C. Perera et al., Sensing as a Service Model for Smart Cities Supported by Internet oZf Things, Trans. Emerging Telecommunications Technologies, vol. 25, no. 1, 2014, pp. 8193.

Security Issues in Cloud Computing: An Overview

Aswathy V T, Aswani P V, Soya Monson
Dept of Computer applications
Vidya Academy of Science & Technology
Thrissur – 680501

Sajay K R
Associate Professor of Computer applications
Vidya Academy of Science & Technology
Thrissur – 680501

***Abstract*—Cloud computing is an emerging technology for providing computing resources and storage to all kinds of users. This technology is facing lot of challenges. This paper is mainly focused on security issues in cloud computing. It also makes an attempt to describe the security challenges in software as a service model of cloud computing and also provide future security research directions. This paper also reviewed Huang et al.s identity authentication and context privacy preservation in wireless health monitoring system, which is based on identity-based cryptography and identity-based signature. Huang et al. argued that their scheme could provide privacy of the health monitoring system. However, this paper showed that Huang et al.s scheme has the lack of context privacy and provides not enough security for physician.**

***Index Terms*—Cloud computing, Virtualization, Multi-tenancy, Software as a service model, Identity authentication, Context privacy, Body area network**

I. Introduction

Cloud computing is stated as the model for delivering the major resources such as the storage, networks , servers ,services and applications which can be released and provisioned with minimal management effort.Cloud computing contains three service models such as Software as a service model, Platform as a service model, Infrastructure as a service model. The deployment models are public, private, community and hybrid cloud. Public cloud is managed over a third-party location. Hybrid cloud refers to joining the two clouds like private-public cloud. Community Cloud is states to cloud implementation model, it is accessed by a specific community with several organizations infrastructure.Cloud users security contains number of concerns, which includes the data loss, unauthorized access and the Cloud Service Providers not effectively safeguarding the cloud data. The cloud service consumer and Cloud service provider must ensure that the cloud is safe from all kinds of exterior threats so that the user does not face any difficulty such as data theft or data loss. Cloud computing is a type of Internet based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources. Cloud computing has number of challenges, security and privacy is one of them. Cloud computing is vulnerable for attacks and threats because of its openness feature. Security of cloud is an important factor for cloud because without security it is insecure to store the data on cloud. Security concerns in cloud are access control, authentication, data confidentiality, integrity, availability and privacy etc

Here also describes about Identity authentication and context privacy preservation in wireless health monitoring systems. Home healthcare based on ubiquitous technology is emerging as a solution to increasing chronic disease patients with the advent of aging of society. While u-healthcare has the advantage that improves accessibility to medical services, it also increases the probability of the infringement of privacy of personal medical information, the leakage of which can do greater damage than any other information. Recently, Huang et al. proposed the identity authentication and context privacy preservation in wireless health monitoring system, which is based on identity-based cryptography and identity-based signature. This paper is to review Huang et al.s security scheme, analyze it focused on the security and privacy concerns, and provide directions to improve the scheme.

This paper focus on the security issues in cloud computing. The organization of the paper is as follows: Section2 describes the security in cloud computing. Section3 describes the security issues that are posed by the service models of cloud computing. Section4 describes security issues in Software as a service model in detail. Section5 describes Identity authentication and context privacy preservation in wireless health monitoring system. Section6 provides conclusions derived out of the survey undertaken.

II. Cloud Computing Security

Virtualization has been in the IT world for a long time. IBM was the first that introduced the idea in the early 1960s with the term Time Sharing.Virtualization of operating systems, also called server virtualization, is defined as a way of making a physical computer function as if it were two or more computers where each non-physical or virtualized. There are two virtualization types that concern cloud computing:

- Full Virtualization: In this type of virtualization,a complete installation of one machine is run on another.
- Paravirtualization: This type of virtualization allows multiple modified OSs to run on a single hardware device at

the same time by more efficiently using system resources

The CSA goes further and states that multi -tenancy in cloud service models implies a need for policy-driven enforcement, segmentation, isolation, governance, service levels, and chargeback/billing models for different consumer constituencies Multi-tenancy has different definitions and importance depending on the services model and deployment models respectively.

The term multi-tenant means customers by using a common infrastructure and code base.In a multi-tenant environment, tenants would have a private space and a common space shared amongst all tenants. By sharing resources and creating standard offerings, multi-tenancy reduces costs and improves efficiency of operations. Multi-tenancy makes use of virtualization technologies to increase resource utilization, load balancing, scalability, and reliability; and the use of automation reduces complexity, decrease operation costs,and increase provisioning speed. Multi-tenancy can be applied to different levels. According to IBM these levels can include:

- Application level: Multiple tenants use an application which provides logical separation between users, access controls, and customization
- Middleware level : Multiple applications use the same middleware which provides logical separation, access controls, and resources.
- Operating system (OS) level: Multiple middleware runs under the same OS which provides access controls, logical separation, and resources to the middleware.
- Hardware level: The hardware provides logical separation, access control and resources to each OS instance. In this level, each OS is considered a tenant.

The main difference between the Multi-tenancy and virtualization enable an efficient computing model. Multi-tenancy allows multiple tenants to coexist in the same physical machine sharing its resources (CPU, memory, network...) and, at the same time, creates an isolated environment for each one. Virtualization is the means used to obtain multi-tenancy. Virtualization allows multiple operating systems (OS) to run on the same physical device at the same time. This allows several users to execute their applications on the same physical environment, but isolated from each other.

Cloud computing is typically classified based on either their deployment or service models.

1) A Private cloud is owned or rented by an organization. The whole cloud resource is dedicated to that organization for its private use. An example of this model is a cloud built by an enterprise to serve their business critical applications.
2) A Public cloud is owned by a service provider and its resources are sold to the public. End- users can rent parts of the resources and can typically scale their resource consumption up (or down) to their requirements. Amazon, Google, Rackspace, Salesforce, and Microsoft are examples of public cloud providers.
3) A Community cloud is similar to a private cloud, but where the cloud resource is shared among members of a closed community with similar interests.
4) A Hybrid cloud is the combination of two or more cloud infrastructures; these can be either private, public, or community clouds. The main purpose of a hybrid cloud is usually to provide extra resources in cases of high demand, for instance

Cloud service models are typically classified as Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS), though slightly different classifications also exist

1) Cloud SaaS is the use of applications running on a cloud infrastructure to provide services to end-users. SaaS can deliver business applications such as customer relationship management (CRM), enterprise resource planning (ERP), and accounting.
2) Cloud PaaS is the use of tools and resources running on a cloud infrastructure to provide services to end-users. The applications are developed and/or acquired by end-users on top of the tools provided.
3) Cloud IaaS is the use of fundamental computing resources, e.g. storage, networks, servers, to provide services to end users.

A. Service Level Agreements for Cloud Security

In many respects, cloud computing represents outsourcing of computation and storage to an external service provider. Such outsourcing has been governed by Service Level Agreements (SLAs) that specify minimum levels of performance that the customer can expect, e.g., 99.999 % system availability per year. Traditionally, however, SLAs have not covered security aspects such as confidentiality and integrity. Security SLAs will typically follow a lifecycle where they are first published generically by a provider, and when a user wishes to use a cloud service, she will then negotiate a specific SLA to which the provider will commit, and the service will be provisioned. The user may want to monitor the service to ensure that the negotiated SLA is being adhered to by the provider.At any time during the commitment, provisioning and monitoring phases, the cycle may return to the negotiation phase, e.g., if the provider after all cannot commit to the previously negotiated SLA. In federated cloud services, these negotiations will have to be performed at multiple levels.

B. Trusted Data Sharing Over Untrusted Cloud Storage Providers

Cloud computing shifts most of the IT infrastructure and data storage to off-premises third-party providers, with two important consequences: (a) Data owners have only limited control over the IT infrastructure, therefore data owners must establish a mechanism to mandate the enforcement of their security policies to ensure data confidentiality and integrity; (b) Cloud service providers have excessive privileges, allowing them extensive control and ability to modify users IT systems and data. These lead to a low trust level when keeping and sharing data on a cloud, especially in a business model which

requires strict secure data processing in order to safeguard business interests. Hence, a secure system is essential to enable trusted data sharing through untrusted cloud providers.

It is necessary that the cloud storage provider helps to enforce the authorization policy for data access, but the enforcement should not reveal any information to the cloud storage provider or enable the cloud storage provider have excessive privileges to allow unauthorized access. The requirements can be achieved by using either homomorphic encryption or incremental encryption.

Homomorphic encryption is a cryptography scheme where algebraic operations applied on the ciphertext are directly reflected in the corresponding plaintext. Simply put, this allows a third party to compute the sum of two encrypted numbers, and when this encrypted result is returned to the user, it can be decrypted with the original key, and the result is the same as the sum of the two numbers in plaintext form. This allows multiple parties to cooperatively generate a piece of ciphertext without knowing the plaintext that others work on.

Incremental encryption allows the computation of the final ciphertext based on the initial ciphertext and the change of the plaintext. Rong et al. propose an incremental encryption scheme based on elliptic curve cryptography. The mechanism allows users to have trusted data storage and sharing over untrusted cloud storage providers. Being able to implement trusted services on untrusted cloud storage providers allows users to manage their data on any cloud storage provider, eliminating the required trust on the providers.

C. Accountability in the Cloud

While bulletproof confidentiality-preserving solutions for the cloud remain a desirable goal, it is clear that as long as bigdata needs to be processed in the cloud, there are currently no sufficiently efficient mechanisms that can do this without letting the cloud providers have access to cleartext data. Thus, there is a need for other mechanisms that can allay the fears of users that otherwise might be scared away from using the cloud.

The cross-border nature of cloud computing also introduces the challenge of complying with multiple, sometimes conflicting,legal codes, especially when data is of a personal sensitive nature. Pearson states that central components of the notion of accountability are transparency, responsibility, assurance and remediation. She also argues that there is a need to move from only retrospective to also prospective accountability, extending mechanisms for implementing security policies to encompass both preemptive and reactive mechanisms, i.e., both preventing bad things from happening, and establishing that bad things did happen, if they could not be prevent.Achieving accountability in the cloud will require re-engineering many services to incorporate legal mechanisms, procedures and technical measures to support such prospective and retrospective accountability mechanisms.

III. Security Issues in Service Models

Cloud computing contains three service models such as Software as a service model, Platform as a service model, Infrastructure as a service model. Each delivery model has its own security issues.

- Security Issues in SaaS
  In a conventional on-premise application deployment model, the confidential information of each organization persists to reside within the organizational boundary and which subject to its personnel security and access control policies, logical and physical. However, in the Software as a Service model, the organization information is kept beyond the organizational boundary, at the SaaS. Accordingly, the SaaS service provider should take on extra security polices to prevent data security and breaks due to security exposures in the application. Data location, Data Disposal, data Integrity, data Confidentiality, authorization and authentication, network attacks and data availability are challenges of Software as a Service delivery mode.
- Security Issues in PaaS
  The users can use the intermediate equipment to create his program and provide it to the customers over the servers and internets. The users controls the applications that run in cloud environment, but it does not control the hardware or network substructure and operating systems. Lack of validation, anonymous sign ups and service fraud are major issues of PaaS.
- Security Issues in Iaas
  Cloud computing service provider delivers resources to authorized users at Pay-Per- use process it reduces the initial investment in hardware such as processing power and networking devices. IaaS provides additional capabilities like more quickly and cost-effectively data access in an internal data centers. Reliability and physical locations are major issues in IaaS service model. But it does not provide reliability to the customer or user on the physical locations of cloud environment. In Iaas security issues based on cloud deployment model. Issue depends on three kinds of parameters like infrastructure management and ownership, infrastructure location and Access and consumption.

IV. Securing Software as a Service Model

In Software as a Service (SaaS) model, the client has to depend on the service provider for proper security measures. The provider must ensure that the multiple users dont get to see each others data. So, it becomes important to the user to ensure that right security measures are in place and also difficult to get an assurance that the application will be available when needed. While using SaaS model, the cloud customer will, by definition, be substituting new software applications for old ones. Therefore, the focus is not upon portability of applications, but on preserving or enhancing the security functionality provided by the legacy application and achieving a successful data migration. The SaaS software vendor may host the application on his own private server or deploy it on a cloud computing infrastructure service provided by a third-party provider.In the following section, the SaaS

security issues have been categorized as traditional and new cloud specific security challenges, for sake of convenience.

A. Traditional Security Challenges

Although the security concerns in traditional communication systems also apply to the cloud. Some of the traditional security issues which also affect the SaaS model have been described below:

- **Authentication and authorization**
  The authentication and authorization applications for enterprise environments may need to be changed, to work with a safe cloud environment. Forensics tasks may become much more difficult since the investigators may not be able to access system hardware physically. verifies user authenticity using two-step verification, which is based on password, smartcard and out of band (i.e. strong two factors) authentication.
- **Availability**
  The availability ensures the reliable and timely access to cloud data or cloud computing resources by the appropriate personnel. The availability of cloud service providers is also a big concern, since if the cloud service is disrupted; it affects more customers than in the traditional model. For instance, the recent disruption of the Amazon cloud service in the year 2011, took down a number of websites including Reddit, Foursquare, and Quora.
- **Data confidentiality**
  Confidentiality refers to the prevention of intentional or unintentional unauthorized disclosure of information. Confidentiality in cloud system is related to the areas of intellectual property rights, covert channels, traffic analysis, encryption, and inference.
- **Virtual Machine Security**
  Virtualization is one of the main components of a cloud. But this poses major security risks. Ensuring that different instances running on the same physical machine are isolated from each other is a major task of virtualization. Virtual machine monitor should be root secure, meaning that no privilege within the virtualized guest environment permits interference with the host system.

B. Cloud Specific Security Challenges

- **Information Security**
  In a traditional on-premise application deployment model, the sensitive data of each enterprise continues to reside within the enterprise boundary and is subject to its physical, logical and personnel security and access control policies.In the SaaS model, the enterprise data is stored outside the enterprise boundary, at the SaaS vendor end. Consequently, the SaaS vendor must adopt additional security checks. This involves the use of strong encryption techniques for data security and fine-grained authorization to control access to data.
- **Network Security**
  In a SaaS deployment model, sensitive data is obtained from the enterprises, processed by the SaaS application and stored at the SaaS vendor end. All data flow over the network needs to be secured in order to prevent leakage of sensitive information. This involves the use of strong network traffic encryption techniques such as Secure Socket Layer (SSL) and the Transport Layer Security (TLS) for security. In case of Amazon Web Services (AWS), the network layer provides significant protection against traditional network security issues, such as MITM (Man-In-The-Middle) attacks, IP spoofing, port scanning, packet sniffing, etc
- **Resource Localty**
  In a SaaS model of a cloud environment, the end-users use the services provided by the cloud providers without knowing exactly where the resources for such services are locate. This poses a potential problem when disputes happen, which is sometimes beyond the control of cloud providers. Due to compliance and data privacy laws in various countries, locality of data is of utmost importance in much enterprise architecture.
- **Cloud standards**
  To achieve interoperability among clouds and to increase their stability and security, cloud standards are needed across different standard developing organizations. For example, the current storage services by a cloud provider may be incompatible with those of other provider. In order to keep their customers, cloud providers may introduce so called sticky services which create difficulty for the users if they want to migrate from one provider to the other, e.g., Amazons S3 is incompatible with IBMs Blue Cloud or Google storage
- **Data Segregation**
  Multi-tenancy is one of the major characteristics of cloud computing. As a result of multi- tenancy, multiple users can store their data using the applications provided by SaaS. In such a situation, data of various users will reside at the same location. Intrusion of data of one user by another becomes possible in this environment. This intrusion can be done either by hacking through the loop holes in the application or by injecting client code into the SaaS system. A client can write a masked code and inject into the application. If the application executes this code without verification, then there is a high potential of intrusion into others data. A SaaS model should therefore ensure a clear boundary for each users data.
- **Data Access**
  Data access issue is mainly related to security policies provided to the users while accessing the data. In a typical scenario, a small business organization can use a cloud provided by some other provider for carrying out its business processes. This organization will have its own security policies based on which each employee can have access to a particular set of data.
- **Web application security**
  SaaS is software deployed over the internet and/or is deployed to run behind a firewall in local area network or personal computer.SaaS application development may use

various types of software components and frameworks. These tools can reduce time-to-market and the cost of converting a traditional on- premise software product or building and deploying a new SaaS solution.

- **Data breaches**
  Since data from various users and business organizations lie together in a cloud environment, breaching into the cloud environment will potentially attack the data of all the user.
- **Backup**
  The traditional backup methods used with earlier applications and data centers that were primarily designed for web and consumer applications, are not optimally designed for the applications running in the cloud. The SaaS vendor needs to ensure that all sensitive enterprise data is regularly backed up to facilitate quick recovery in case of disasters.
- **Identity management and sign-on process**
  Identity management (IdM) or ID management is an area that deals with identifying individuals in a system and controlling the access to the resources in that system by placing restrictions on the established identities.

V. Current Security Solutions

There are several research works happening in the area of cloud security.The Cloud Security Alliance (CSA) is gathering solution providers, non- profits and individuals to enter into discussion about the current and future best practices for information assurance in the cloud. The best security solution for SaaS applications is to develop a development framework that has tough security architecture.Another approach is resource isolation to ensure security of data during processing, by isolating the processor caches in virtual machines, and isolating those virtual caches from the hypervisor cache.

VI. Security Concerns of Healthcare System

In the 21st century, the healthcare industry has seen the drastic improvements due to the involvement of wireless medical body area networks (BANs) in healthcare applications . A few decades ago BANs were a topic of science/movie fiction for healthcare industries, and now they have become a reality and provide much quality-of-care. In fact the future of modern healthcare in an aging world will need ubiquitous monitoring of health with least actual interaction of doctor and patients. The development of a wireless healthcare application offers many novel challenges, such as, reliable data transmission, node mobility support and fast event detection, timely delivery of data, power management and node computation . Further however, deploying new technologies in healthcare applications without considering security often makes patient privacy vulnerable . Due to the patient health information (PHI) become digitization compared to the traditional medical care, remote medical monitoring system has also brought a series of new challenges. The most important challenge is how to ensure the patient privacy during transmission of data to avoid the threat from the attacker . Transmission for medical information, which uses the public key infrastructure to complete the encryption and authentication of the information . But in the large-scale distributed network, the inconvenience problems from the update and remove of digital certificate need to be considered. Protecting the destination privacy is another alternative to achieve contextual privacy

A. System Configuration

This section reviews healthcare monitoring architecture, adversary model and security requirements for healthcare application for Huang et al.s system.

1) **System Model :** The design of the health monitoring system came with a lot of emerged challenges. The government has established stringent regulations to ensure that the security and privacy of patients PHI are properly protected, To preserve the context privacy, the health monitoring system is organized by a trusted authority (TA). The system model includes the registered patients, physicians, electric health record (EHR) database in the monitoring server and TA

2) **Adversary Model :** There are three types of threats: misuse of patient identities, unauthorized access and modification of PHI in the health monitoring system. Huang et al. in considered three types of adversary: the patient himself/herself, insiders and outsiders.

- **Identity threat**
  There are three concerns here. First, the patient may lose (or share) their identity credentials, enabling others to have access to their PHI in the EHR (or in their mobile node (MN)). Second, insiders may use patient identities for medical fraud.
- **Access threat**
  Huang et al. explored threats related to unauthorized access to PHI, whether in the MN or the EHR. The first threat comes from the patient himself/herself, because the patient has a right to control the collection, use, and disclosure of PHI; if the Patient fails to express their consent consistent with their actual preference, for whatever reason, they may allow broader-than-intended collection, access or disclosure; Insiders may peek at patient data, out of curiosity, or with the intent to harm the patient.
- **Disclosure threats**
  Huang et al. explored threats related to the disclosure of PHI, including data at rest and data in transit.

B. Huang et al.'s Privacy-Preservation System

This section reviews Huang et al.s patient privacy preservation system based on the concept of IBC.

1) **System Parameters Generation :** To set up the system, TA first initializes all required system parameters (q, G1, G2, ł, P0, H1).

2) **Physician and Patient Registration :** When Alice registers in the health monitoring system, she inputs her personal information and then gets her personal identity IDA from the health monitoring server. TA computes her private key SA = SH1(IDA), and transfers the corresponding public

key QA = SP to the health monitoring server. Alice gets the medical equipments, doctor's ID and public key QEHR of the EHR data center.

The doctor Bob gets his identity IDB when he fills in personal information to register the health monitoring server. Bob inputs the personal login id password PWDB, computes the hash value H(PWDB) which is stored in EHR data center. Bob gets his private key SB = S0'H1(IDB) and the corresponding public key QB = S0'P from TA.

3) **Patient Health Information Transmission:** After Alice gets the medical equipments and goes back to home, the BAN constructed by these instruments can collect her health data m. Before the information are sent to the EHR through Internet, it is necessary to take corresponding encryption and put signature to ensure that during the information transmission process it can resist the malicious attacks like decryption, tamper and forging, etc.

4) **Patient Health Information receiving and storing:** Patient Alice's PHI is transferred to the EHR data center. When EHR data center receives the cipher text C, firstly it uses its own private key to decrypt the message and verify the legitimate identity of Alice. The message stored in the EHR data center is the cipher text C = IBCQEHR (IDA||m||IDB) encrypted with the private key of the health monitoring server. Only doctor Bob can access Alice's health message.

5) **Patient Health Information Recovering:** The health monitoring server sends a notice message which shows the receiving information of the patient Alice to the doctor Bob. Registered Bob enters the health monitoring server with the password PWDB. After the health monitoring server authenticates Bobs identity of physician based on role-based access control, Bob enters into the system and accesses the information mby querying.

VII. Security Threats

The medical sensor senses patient sensitive PHI and transmits it over the wireless channels which are more susceptible than wired networks. Thus, patient sensitive PHI must remain in secure and private from security threats.

- Monitoring and eavesdropping on PHI: This is the most common threat to the patient privacy. By patient PHI snooping, an adversary can easily discover the patient information from communication channels. Moreover, if the adversary has a powerful receiver antenna, then he/she can easily pick up the messages from the network.
- Threats to the transmitted data : Wireless communication ranges are not confined, and are easily vulnerable. In wireless healthcare applications, medical devices sense the patient and environmental data, and send it either to the physician or the hospital server. While sending PHI to the system, it may be attacked.
- Masquerade and replay threats : In a home healthcare application, an attacker can easily rogue a wireless rely point while patient data is transmitting to the remote location. In general, wireless rely nodes are unguarded, so it may happen that a rogue rely node can provide unrestricted access to an attacker who can then cause a masquerade. In this threat, an illegal rely node acts as a real node to the network.
- Location threats : Medical BANs support patient mobility, so exact patient location knowledge is necessary since location knowledge allows reaching medical staff in a short time, in case of any emergency . Generally, locationtracking systems are based on radio frequency, ultrasound
- DoS threats : A DoS attack is any event that diminishes or eliminates a networks capacity to perform its expected function. DoS threat could be even more disruptive in healthcare applications because such a network needs always-on patient health monitoring. A list of DoS attacks is categorized depending on layers
- Link/medium access control layer: This layer suffers mainly from collision, exhaustion, and unfairness attacks. In collision attacks, an adversary simultaneously transmits the packets at same frequency, resulting in packet collision and degradation of the network performance. In exhaustion attacks the battery source is self-sacrificing, since wireless nodes most of the time maintains the channel active. In unfairness attacks, network performance degrades because this layer priority is generally disrupted according to the applicationrequirements.
- The network and routing layer: Routing-disturbance attacks lead to DoS threats in multi-hop medical sensor environments. Generally, the routing attacks involve spoofing, altering routing paths or replaying packets, selective forwarding, sinkhole, warm-hole, etc.
- Transport layer: It controls end-to-end links, and suffers mainly from two popular types of attacks, namely, flooding attacks and de-synchronization attacks. Flooding attacks generally are used to drain the memory resources by sending the control signals. In de-synchronized attack, attacker may disturb the established link betweenlegitimate two ends nodes by re-synchronizing their transmission. As a result, it disturbs network communication, and network resources exhaustion.

VIII. Conclusion

Though there are numerous advantages in using a cloud-based system, there are yet many practical issues which have to be sorted. Cloud computing is a disruptive technology with profound implications not only for Internet services but also for the IT sector as a whole. Still, several outstanding issues exist, particularly related to service-level agreements (SLA), security and privacy, and power efficiency.

In this paper an overview of cloud computing service delivery model, challenges faced by cloud computing, SaaS along with the security challenges , including both the traditional and cloud specific security challenges. It also deals with dentity authentication and context privacy preservation in wireless health monitoring system. This paper provided future directions for the secure and privacy preserving wireless health monitoring system. The need for further work on various

security mechanisms has also been highlighted, in order to provide transparent services that can be trusted by all users.

REFERENCES

[1] Amazon EC2, http://aws.amazon.com/ec2/
[2] Google App Engine, http://code.google.com/appengine/
[3] Google Apps, http://docs.google.com/
[4] Amazon Web Services, Zeus Botnet Controller, Accessed on July 2011, http://aws.amazon.com/es/security/zeus-botnetcontroller/
[5] Onankunju, Bibin K., "Access Control in Cloud Computing", International Journal of Scientific and ResearchPublications3,no.9(2013):1.
[6] Dimitrios Zissis and Dimitrios Lekkas, "Addressing Cloud Computing Security Issues", ELESVIER, 2012.pp. 583-592.
[7] Kuyoro S.O., Ibikunllef and Awodele O., "Cloud Computing security issues and challenges", IJCN,2011.
[8] Hashizume et al., "An Analysis of Security Issues for Cloud Computing", Journal of Internet Services and Applications, Springer, 2013.
[9] Y. Zhang, Y. Xu, L. Shang, K. Rao, "An investigation into health informatics and related standards in China," International Journal of Medical Informatics, vol. 76, no. 8, pp. 614-620, 2007.
[10] P. Kumar, H.-J. Lee, "Security Issues in Healthcare Applications Using Wireless Medical Sensor Networks: A Survey," Sensors, vol. 12, pp. 5591, 2012.
[11] M. M. Haque, A.-S. K. Pathan, C. S. Hong, "Securing U-Healthcare Sensor Networks using Public Key Based Scheme," Proc. of ICACT 2008, pp. 17-20, 2008.

Online Social Network: Threats and Solutions

Avinash V, Mobin M M, Samson Peter T
Department of Computer Applications
Vidya Academy of Science and Technology
Thrissur – 680501

Salkala K S
Assistant Professor of Computer Applications
Vidya Academy of Science and Technology
Thrissur – 680501

***Abstract*—In recent years, global Online Social network (OSN) usage has increased very much and it became part of our everyday life as virtual meeting places that facilitate communication. With the increasing usage of OSNs, many users have unknowingly become exposed to threats both to their privacy and security. Fortunately, there are many software solutions and techniques that exist today which can assist users in better defending themselves against these security threats. The proposed approach presents a system of collaborative content management that relies on an extended notion of a content stakeholder. A tool, Collaborative Privacy Management (CoPE), is implemented as an application within a popular social networking site, facebook.com, to ensure the protection of shared images generated by users.**

***Index Terms*—Social networks, K-Anonymity, Non-Linear Dimension Reduction Techniques, Dimensionality Reduction, Quantum cryptography,Quantum-key distribution, Collaborative Privacy Management(CoPE).**

I. Introduction

A social networking site as "A website that provides a virtual community for people interested in a particular subject or just to 'hang out' together." While this is an accurate description, a more detailed definition of online social networking would also encompass all of the ways people can connect. Online networks such as Facebook, Twitter, MySpace, and LinkedIn all offer users a variety of ways to increase their networks, share notes and various types of media, and connect on a variety of levels. Social networking is not a static thing. Networks are growing and changing all the time, with new ones popping up at a fast rate. Many networking websites are geared towards users with specific interests and needs, while others encourage everyone to join. Since this form of networking is always evolving, the definition of a social network will also be fluid. At its heart, however, an online social network is a meeting place for people to extend their reach and stay in contact with their connections.

Some of the ways people can network on these sites include:

- Having current friends or connections suggest other people you may want to network with
- Sharing photos, audio and video files, written works, links, and other media
- Posting a resume or work experience for job searching and recruiting
- a brand or service to those who may need the product or service

Recently, online social network has emerged as a promising area with many products and a huge number of users. With the development of information retrieval and search engine techniques, it becomes very convenient to extract users personal information that is readily available in various social networks. Malicious or curious users take advantage of these techniques to collect others private information. Therefore, it is critical to enable users to control their information disclosure and effectively maintain security over online social networks. One of the challenges in social network is security.

II. What Types of Social Networks Exist?

While Facebook, Twitter and LinkedIn might be the first sites that come to mind when thinking of social networking, these popular websites do not represent the full scope of social networks that exist. Learn more about the different options available for people to interact and collaborate with each other online.

A. Social Connections

Keeping in touch with friends and family members is one of the greatest benefits of social networking. Here is a list of the most widely-used websites for building social connections online.

- Facebook: Arguably the most popular social media utility, Facebook provides a way for users to build connections and share information with people and organizations they choose to interact with online.
- Twitter: Share your thoughts and keep up with others via this real-time information network.
- Google +: This relatively new entrant to the social connection marketplace is designed to allow users to build circles of contacts that they are able to interact with and that is integrated with other Google products
- MySpace: Though it initially began as a general social media site, MySpace has evolved to focus on social entertainment, providing a venue for social connections related to movies, music games and more.

B. Multimedia Sharing

Social networking makes it easy to share video and photography content online. Here are some of the most popular sites for multimedia sharing.

- YouTube: Social media platform that allows users to share and view video content
- Flickr: This site provides a powerful option for managing digital photographs online, as well as for sharing them with others.

C. Professional

Professional social networks are designed to provide opportunities for career-related growth. Some of these types of networks provide a general forum for professionals to connect, while others are focused on specific occupations or interests. A few examples of professional social networks are listed below.

- LinkedIn: As of November of 2011, LinkedIn had more than 135 million members, making it the largest online professional network. Participants have an opportunity to build relationships by making connections and joining relevant groups.
- Classroom 2.0: Social network specifically designed to help teachers connect, share and help each other with profession-specific matters.

D. Informational

Informational communities are made up of people seeking answers to everyday problems. For example, when you are thinking about starting a home improvement project or want to learn how to go green at home, you may perform a web search and discover countless blogs, websites, and forums filled with people who are looking for the same kind of information. A few examples include:

- Super Green Me: Online community where individuals interested in adopting green living practices can interact
- Do-It-Yourself Community: Social media resource to allow do-it-yourself enthusiasts to interact with each other.

E. Educational

Educational networks are where many students go in order to collaborate with other students on academic projects, to conduct research for school, or to interact with professors and teachers via blogs and classroom forums. Educational social networks are becoming extremely popular within the educational system today. Some examples of such educational social networks are listed below.

- The Student Room: UK-based student community featuring a moderated message board and useful resources related to school
- The Math Forum: A large educational network designed to connect students with an interest in math, this site provides interaction opportunities for students by age group.
- ePALS School Blog: This international social network for K-12 students is designed to build international connections to promote world peace.

F. Hobbies

One of the most popular reasons many people use the Internet is to conduct research on their favorite projects or topics of interest related to personal hobbies. When people find a website based on their favorite hobby, they discover a whole community of people from around the world who share the same passion for those interests. This is what lies at the heart of what makes social networks work, and this is why social networks that are focused on hobbies are some of the most popular. A few examples of hobby-focused social networking sites include:

- Oh My Bloom: Social media site specifically for gardening enthusiasts. It features groups, forums, blogs, video content and more.
- My Place at Scrapbook.com: Designed specifically for scrapbooking enthusiasts, users can create profiles, share information, post updates and more.

G. Academic

Academic researchers who want to share their research and review results achieved by colleagues may find academic-specific social networking to be quite valuable. A few of the most popular online communities for academics are:

- Academia.edu: Users of this academic social network can share their own research, as well as follow research submitted by others.
- Connotea Collaborative Research: Online resource for scientists, researchers and clinical practitioners to find, organize and share useful information.[9]

III. Major Characteristics of Social Networks

A. User-based

Before social networks like Facebook or MySpace became the norm, websites were based on content that was updated by one user and read by Internet visitors. The flow of information was in a single direction, and the direction of future updates was determined by the webmaster, or writer. Online social networks, on the other hand, are built and directed by users themselves. Without the users, the network would be an empty space filled with empty forums, applications, and chat rooms. Users populate the network with conversations and content. The direction of that content is determined by anyone who takes part in the discussion. This is what make social networks so much more exciting and dynamic for Internet users

B. Interactive

Another characteristic of modern social networks is the fact that they are so interactive. This means that a social network is not just a collection of chat rooms and forums anymore. Websites like Facebook are filled with network-based gaming applications, where you can play poker together or challenge a friend to a chess tournament. These social networks are quickly becoming a pastime that more people are choosing over television - because it's more than just entertainment, it's a way to connect and have fun with friends.

C. Community-driven

Social networks are built and thrive from community concepts. This means that just like communities or social groups around the world are founded on the fact that members hold common beliefs or hobbies, social networks are based on the same principle. Within most modern online social networks today, you'll find sub-communities of people who share commonalities, such as alumni of a particular high school, or an animal welfare group. Not only can you discover new friends within these interest based communities, but you can also reconnect with old friends that you lost contact with many years ago.

D. Relationships

Unlike the websites of the past, social networks thrive on relationships. The more relationships that you have within the network, the more established you are toward the center of that network. Like the concept most pyramid schemes are focused on, within online social networks, the concept really works in a powerful way. When you have just 20 contacts and you publish a note or an update on that page, that content proliferates out across a network of contacts and sub-contacts that's much larger than you may realize.

E. Emotion over content

Another unique characteristic of social networks is the emotional factor. While websites of the past were focused primarily on providing information to a visitor, the social network actually provides users with emotional security and a sense that no matter what happens, their friends are within easy reach. Whether suffering through divorce, break-up or any other family crisis, people are finding that the ability to jump online and communicate directly with a circle of friends provides a great deal of support in an otherwise unmanageable situation.[7]

IV. Advantages and Disadvantages of Social Networking

A. Advantages

1) Worldwide Connectivity
   No matter if you are searching for a former college roommate, your first grade teacher, or an international friend, no easier or faster way to make a connection exists than social media. Although Facebook, Twitter, LinkedIn and Pinterest are probably the most well-known social networking communities, new websites are popping up regularly that let people connect and interact over the Web. With each of these sites, individuals can make new friends, build business connections or simply extend their personal base by connecting and interacting with friends of friends - which can have a multiplying effect. These connections can help with a variety of things such as:
   - Finding romance
   - Seeking a new job
   - Locating assistance
   - Getting and giving product and service referrals
   - Receiving support from like-minded individuals
   - Making or receiving career or personal advice Sharing political beliefs
   - Accessing news in real time

   In many ways, these social communities are the virtual equivalent of church socials where family and friends gather to exchange news and get updates. Even the age-old custom of connecting with pen pals has been upgraded as private messages can be sent over social media. When it comes to getting information, few methods are faster than social media.
2) Commonality of Interest
   When you opt to participate in a social network community, you can pick and choose individuals whose likes and dislikes are similar to yours and build your network around those commonalities. For instance, if you are a chess aficionado, a book lover or have a particular political leaning, you can find and interact with those who share your interest. It can also be a great way to share tips and ideas. Sites like Pinterest have been very successful due to the ease in which a person can learn - and share - information about hobbies, crafts, cooking, gardening and other do-it-yourself activities. By pinning and sharing, you can attract like-minded individuals into your circle. But, just as these virtual groups can help hobbyists exchange ideas and techniques, other social network groups offer solutions for more vexing, real-world problems. For example, social media groups can be lifelines for individuals suffering from a rare disease. Churches, synagogues and temples also use social media to reach out to members who may be unable to attend services.
3) Real-Time Information Sharing
   Many social networking sites incorporate an instant messaging feature, which lets people exchange information in real-time via a chat. This is a great feature for teachers to use to facilitate classroom discussions because it lets them utilize the vast store of information available on the Web. This can be a great time saver for the teacher - since students no longer need to visit a library to conduct research- and it can be a great way to engage distracted learners. School is not the only setting where this type of real-time information sharing can be beneficial. Social networking can provide a tool for managers to utilize in team meetings, for conference organizers to use to update attendees and for business people to use as a means of interacting with clients or prospects. Some leaders are going so far as to include Tweets or other social media updates during presentations. This approach can make events more interactive and help the presenter reach a larger audience.
4) Targeted Advertising

Whether you are non-profit organization that needs to get the word out about an upcoming fundraiser or a business owner marketing a new product or service, there's no better way than social media to get your message in front of millions of people 24/7. Although social media can be used to spread a company's message for free, fee-based advertising options are also available. One of the best aspects of social networking is the ability to deliver your content only to those users with the most potential interest in your product or service. Each social platform offers an array of tools that enable a business to deliver specific content to a very specific target group. This approach can maximize targeted reach while minimizing waste.

5) Increased News Cycle Speed
Undoubtedly, social networking has revolutionized the speed of the news cycle. Most news organizations now rely on social media sites to collect and share information. Social media - especially Twitter - is steadily becoming a mainstream source for breaking news. Today an individual can know, in real time, what is happening throughout the world. This has led to the development of a nearly instantaneous news cycle as everything from terrorist attacks to local car crashes get shared on social media, quickly alerting their intended audience of the event.[8]

B. Disadvantages

1) Backlash
A joke among friends is one thing but a joke with the world at-large is much different. When potentially offensive content is posted online, the amount of feedback can be excessive and is often brutal. This is particularly true with highly opinionated subjects like politics and religion. This backlash can also have a long-term impact on a person's future, especially in a world that has fallen prey to over-sharing. Even high school students are learning that comments they post on social media can influence whether a college approves their application for admission. In an age where selfies are the norm, the over-sharing may even be altering our worldview by creating a more narcistic mindset.
2) Cyberbullying and Crimes against Children
Use of social networks may expose individuals to other forms of harassment or even inappropriate contact. This can be especially true for teens and younger children. Unless parents diligently filter the Web content their family views, children could be exposed to pornography or other inappropriate content. Besides unleashing age-inappropriate content, the digital age also gave birth to a social phenomenon - cyberbullying. It is often levied more harshly against young females that males and, unlike traditional bullying, it is not limited to physical interaction. Cyberbullying can happen 24 hours a day, every day of the week. Adding to this realm of cyber abuse are the anonymous social media sites which can elevate the severity of the assault - under the false promise of privacy.
3) Risks of Fraud or Identity Theft
Whether you like it or not, the information you post on the Internet is available to almost anyone who is clever enough to access it. Most thieves need just a few vital pieces of personal information to make your life a nightmare. If they successfully steal your identity, it could cost you dearly. A report on Bank rate reveals Millennias are one of the fastest growing groups to be victims. This is linked to the group's comfort with sharing everything online - including personal information.
4) Time Waster
Business Insider reports that social media is the most popular use of the Internet - surpassing email - and smartphones and other mobile devices seem to be the driving force behind this trend since 60 percent of the traffic is from a mobile source. The GlobalWebIndex poll shows that 28 percent of the time spent online is on social networks. With these type of numbers, some of the time spent on social media occurs at work. When these visits are for non-work related activity, it can cost companies money through lost productivity. A report on Forbes states that 89 percent of responders admitted to wasting time on social media while at work.
5) Corporate Invasion of Privacy
Social networking invites major corporations to invade your privacy and sell your personal information. Have you ever posted a comment on Facebook, only to notice an advertisement appear with content related to your post? Last year, Facebook earned an estimated $16 billion in ad revenue. That's not bad for a free site.[8]

V. THREATS

Information that shared in social networks can use against the user to launch cyber-attacks. Increase in the amount of shared information is may be tend to information leakage risks. Hackers are always one step ahead of security specialists. They always misuse human vulnerabilities to launch social engineering attacks. Different types of threats in social networks are:

A. Phishing

Phishing attacks are usually use legitimate looking but fake websites and emails to fool the victim into revealing private information. In this type of attack the attacker tries to clone the legitimate website in a way that the fake copy looks identical to the original one. For instance, first he opens a bank website and then he copies the page into the hard drive. Another page that receive the username and password and store it for the attacker.

B. Malicious Shortened URLs

In here attackers use URL shorter services to hide the address of malicious URL from the victim. In social networks attackers uses URL shorter for obfuscation of malicious URL. In this technique they use a hoax for infecting the victim with

the malware, very similar to the phishing attack. Example of these services is Goo.gl and bit.ly.

C. Identity Theft

Sharing personal information on social networks may allow attackers to collect enough information to capture the identity of the victim. In here attacker to guess passwords, answers to password recovery questions and more. Attacker can make few combinations by using the public information about the users such as name, country; city, birth date, and picture are stored in database of the victim and his interests to guess the password.

D. Malicious Third Party Applications

Some online social networks such as Facebook and Twitter may allow the user to use third party applications. These applications provide more functionality in additional to the basics provided by social network itself. Third party websites to develop and release their own application based on the framework of that specific network. Those applications that are malicious intentionally can spread themselves by posting a fake advertise of themselves on the wall of victim to attract more users. Once victim go to the application and allows it to start working, afterward the application will start sharing itself. Different types of spreading techniques used by malicious applications such as sharing on feed, sharing in message, and tagging friends in the picture of the application.

E. Spam

Spams are advertised messages that are send in bulk amount to the internet users. Spam is flooding the internet with many copies of the message. Spam provide the social relationship between users the Social relationship will bring more trust. Trust in friends can easily convince the victim to read the spam message, and believe in the content of the message.[5]

F. Fake Users

Online social networks try to make the registration process of users easier to attract more users. Process of creating fake accounts much easier because it is simple process that only asks for a name, an email address and a password. Lately Facebook released a statistic that reveals around 83 million of its users are fake. In this type of attack, the attacker makes a fake account with legitimate look information such as a fake name, city, birth date and few fake pictures. Then he tries to connect to the victim by sending his friendship request. By accepting a friend request from a fake account the victim will expose all of his limited privacy personal information to the attacker.

G. Look Redirects

Most of online social networks have a specific page for redirecting the user to another destination. Usually this system used to calculate the statistic of referring to that destination address. However attackers misuse this system to redirect the user to a malicious URL. Example of such service in the Facebook is: `https://www.facebook.com/l/e9bf8;www.maliciouswebsiteaddress.com`.

H. Direct Anonymity Issues

These systems require the user to allow access to his or her social network profile information and at the same time associate that information with the users identity. IN a peer-to-peer context-aware mobile social network system such as Social Aware, we can track a user by logging the date and time that each mobile or stationary device detects the users social network ID. By collecting such logs, we can construct a history of the locations that a user has visited and the times of each visit, compromising the users privacy. We conclude that clear text exchange of social networking IDs in systems such as WhozThat and Social Aware leads to unacceptable security and privacy risks, and allows the users anonymity to be easily compromised. We call such problems that directly compromise a users anonymity direct anonymity attacks. Direct anonymity attacks are also possible in client-server mobile social network systems. [4]

VI. Solutions

1) Trust Networks and Onion Routing
   One way to support privacy in social network applications is to transfer information using a trusted peer-to-peer network .Such a network would require a trust network much like that used by Katz and Goldbeck in which social networks provided trust for default actions on the web. Moreover, in a mobile social network application, nodes could not only share their information directly but could give permission to their trusted network to share their information. This approach was used in the One Swarm system to allow peer-to-peer file sharing with privacy settings that allowed the user to share data publicly, just with friends, or even with a chosen subset of those friends.
2) URL Safety Tools
   These tools by accessing to the online database of phishing attacks and malicious websites can determine that if the target URL is among them or not. Usually these databases are provided by famous antivirus companies or anti phishing websites.
3) Adjust Privacy Settings
   It means readjust the privacy settings by reducing his/her photo albums, videos, posts, comments, friend list, relation status and other personal information details. We strongly suggest minimizing the level of public sharing. This solution will mitigate the risk of Identity theft and spams.
4) Adjust Application access level
   Also user can reduce the application by removing old applications. This solution will mitigate the risk of malicious third party applications.
5) Limited Sharing
   To limit the sharing information do not share personal private information publically. This solution will mitigate the risk of Phishing, Spams, Identity thefts and fake users.

6) Think twice
   Think twice before performing any action in social networks environments.[3]

VII. TECHNIQUES

In this section we describe possible solutions which can assist in protecting the security and privacy of OSN users.

A. Non-Linear Dimension Reduction Techniques

Advances in data collection and storage capabilities during the past decades have led to an information overload in most sciences. Researchers working in domains as diverse as engineering, astronomy, biology, remote sensing, economics, and consumer transactions, face larger and larger observations and simulations on a daily basis. Such data sets, in contrast with smaller, more traditional data sets that have been studied extensively in the past, present new challenges in data analysis. Traditional statistical methods break down partly because of the increase in the number of observations, but mostly because of the increase in the number of variables associated with each observation. The dimension of the data is the number of variables that are measured on each observation. We subdivide techniques for dimensionality reduction into convex and non-convex techniques. Convex techniques optimize an objective function that does not contain any local optima, whereas non-convex techniques optimize objective functions that do contain local optima. The further subdivisions in the taxonomy are discussed in the review in the following two sections[1].

1) Convex Techniques for Dimensionality Reduction: Convex techniques for dimensionality reduction optimize an objective function that does not contain any local optima, i.e., the solution space is convex. Most of the selected dimensionality reduction techniques fall in the class of convex techniques. One technique solves an additional semi definite program using an interior point method. We subdivide convex dimensionality reduction techniques into techniques that perform an eigen decomposition of a full matrix and those that perform an eigendecomposition of a sparse matrix.

1) Full Spectral Techniques:
   Full spectral techniques for dimensionality reduction perform an eigendecomposition of a full matrix that captures the covariances between dimensions or the pairwise similarities between datapoints (possibly in a feature space that is constructed by means of a kernel function). In this subsection, we have five such techniques: a) PCA / classical scaling b) Isomap c) Kernel PCA d) Maximum Variance Unfolding e) diffusion maps.
2) Sparse Spectral Techniques:
   In the previous subsection, we discussed five techniques that construct a low-dimensional representation of the high-dimensional data by performing an eigende composition of a full matrix. In contrast, the four techniques discussed in this subsection solve a sparse (generalized) eigenproblem. All presented sparse spectral techniques only focus on retaining local structure of the data. We discuss the sparse spectral dimensionality reduction techniques:a) LLE b) Laplacian Eigenmaps c)Hessian LLE d) LTSA

2) Non-convex Techniques for Dimensionality Reduction: We have a non-convex techniques for multidimensional scaling that forms an alternative to classical scaling called Sammon mapping , a technique based on training multilayer neural networks ,and two techniques that construct a mixture of local linear models and perform a global alignment of these linear models. So we have following non-convex techniques: a) Sammon Mapping b) MultilayerAutoencoder c) Locally Linear Coordination (LLC) d) Manifold Charting

B. Quantum cryptography

Quantum cryptography was proposed by Bennett and Brassard in 1984, who also defined the first QKD protocol, called BB84. At time of writing, a handful of research teams around the world have succeeded in building and operating quantum cryptographic devices. Fundamental aspects of quantum physics unitarity, the uncertainty principle, and the Einstein-Podolsky-Rosen iolation of Bells inequalities suggest a new paradigm for key distribution: quantum cryptography. Initial experiments seem to confirm the utility of this paradigm. Assuming that the theoretical models continue to be confirmed in the use of actual devices, the fundamental laws of nature can be invoked to assure the confidentiality of transmitted data. Quantum cryptography more properly termed Quantum Key Distribution, QKD employs two distinct channels. One is used for transmission of quantum key material by very dim (single photon) light pulses. The other, public channel carries all message traffic, including the cryptographic protocols, encrypted user traffic, etc. QKD consists of the transmission of raw key material, e.g., as dim pulses of light from Alice to Bob, via the quantum channel, plus processing of this raw material to derive the actual keys (Fig. 2). This processing involves public communication (key agreement protocols) between Alice and Bob, conducted in the public channel, along with specialized QKD algorithms. The resulting keys can then be used for cryptographic purposes, e.g., to protect user traffic. By the laws of quantum physics, any eavesdropper (Eve) that snoops on the quantum channel will cause a measurable disturbance to the flow of single photons. Alice and Bob can detect this, take appropriate steps in response, and hence foil Eves attempt at eavesdropping.

For achieve security in social network, we proposed a model using quantum cryptography to achieve security. In this model first we collect the information regarding the social network data and after that reduce the dimension of the data set using various dimension reduction technique. After reducing the data, we covert the data into digitized form and to encrypt the data we use the MD5 (Message Digest 5) technique. Generally, instead of MD5, we can also use any other encryption techniques, such as DSA, AES and so on, to decrypt the message. For the secure key distribution purpose we can apply the quantum cryptography. Quantum cryptography provides a secure way to distribute the key on

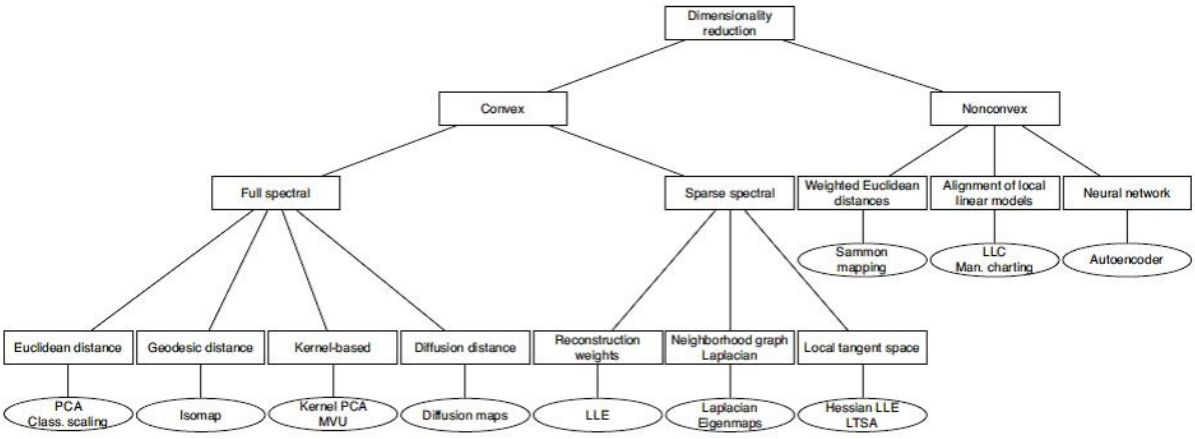


Fig. 1. Taxonomy of Dim. Reduction Techniques

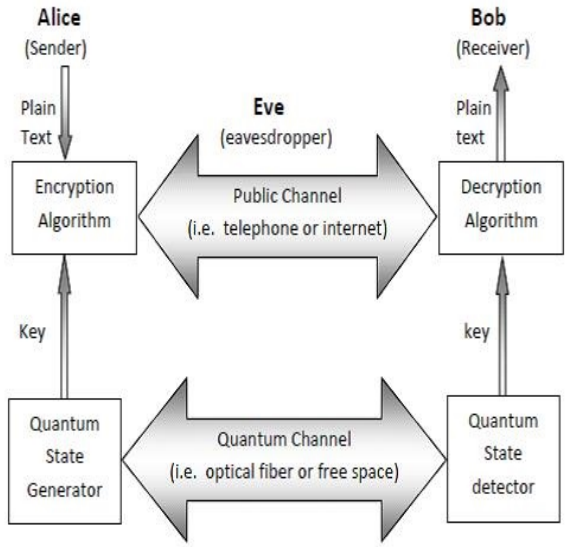


Fig. 2. Quantum-key distribution

the basis of quantum theory. Generally, it is very difficult to implement this key distribution technique because it requires Photon light pulses (PLP) through which photon travel from one user side to another side. So, in this paper we just proposed the QKD (Quantum Key Distribution) for key distribution.[6]

C. Collaborative Privacy Management(CoPE)

We chose online image sharing as the application domain to study collaborative privacy management. This context has a number of characteristics that make it particularly suitable for this research. First, online image sharing has become an increasingly popular feature on OSNs. Consequently, image sharing that increases the opportunities to maintain social relationships to users will become a popular feature globally on OSNs, and its growth trajectory is striking. Second, privacy concerns become particularly salient in this context because online images are often tied to individual profiles either explicitly (through tags on images) or implicitly (through recurrence). As an integral and popular part of personal profiles on OSNs, user-provided digital images constitute a rich data source for those attempting to correlate profiles across multiple services using face recognition. Users are often free to post images regardless of their content, and are usually not required to notify other users prior to publishing their pictures, even if they are explicitly identifiable or tagged.1 Many privacy concerns may arise from this practice because users may be unaware of the fact that a large and potentially unwanted audience could access their personal data or data related to them. While understudied,we believe that issues related to collaborative privacy management for image sharing will rapidly gain attention due to the growing popularity of social media where collaborative action with rich data exchange is the norm. Our method is to design tools that allow users to collaboratively manage their shared images in OSNs. This collaborative privacy management approach considers two major factors: content that need to be protected, and stakeholders who are involved in content-sharing and -privacy management. In this section, we first develop a scenario that demands collaborative privacy management in OSNs, and then describe design requirements for tools to support such collaborative privacy management.[2]

1) User Needs and Requirements of Collaborative Privacy Management:

- A content-owner should be able to invite tagged-users as a co-owner to co-manage privacy-content.
- A co-owner should be aware of the creation of privacy-content that concerns him or her.
- A co-owner should be able to request the control over the privacy-content from the content-owner who created the privacy-content. The control includes the ability to

delete, update, and tag the content.

- A co-owner should be able to specify the accessibility of private-content by content-viewers. Possible access privileges include the ability to view, modify, and comment on given privacy-content.

2) Features of CoPE: Our design of CoPE focuses on supporting the management of the access rights for digital images, and provides the following functions (see the Appendix for the interface design):

- Potential Co-managed Photos Notification: Adding tagged images to the CoPE tool for collaborative privacy management (Tagging can be completed prior to the image being uploaded or as the image is uploaded on CoPE.); Notifying users when they have been tagged by friends who also are using CoPE.
- Stakeholder Request: Allowing users to request co-ownership on images in which they were tagged; Notifying users about the requests on co-ownership;Allowing users to grant co-ownership to others.
- Photo Access Management: CoPE allows a stakeholder to control various privacy-related settings that relate to their photos.That is, a user can set the viewable attribute of any photo to only co-owners, some friends, or public to limit the potential viewers of the photo.
- Track Viewing History: CoPE allows a user to keep track of who has viewed their photos.

VIII. CONCLUSION

OSNs have become part of our everyday life and, on average, most Internet users spend more time on social networks than in any other online activity. We enjoy using OSNs to interact with other people through the sharing of experiences, pictures, and videos. Nevertheless, social networks have a dark side ripe with hackers, fraudsters, and online predators, all of whom are capable of using OSNs as a platform for procuring their future victims. In this paper, we have presented scenarios which threaten OSN users and can jeopardize their identities, privacy, and well-being in both the virtual world as well as the real world. There are remedies to these threats, and we have offered a range of solutions which help protect an OSN users privacy and security. In order to be well protected against the various online threats, users must stay attentive to the information they post online, and they must employ more than one solution. In addition, parents must monitor their childrens activity in these social platforms. As parents, we cannot be nave, we need to recognize the enticements of social networks and be aware of hidden dangers. We are obligated to educate our children to be aware of potential threats, and we must teach them not to engage with strangers either in the real world or in the cyber world.

REFERENCES

[1] Lokesh Jain ,Satbir Jain "A New Approach to Supervise Security in Social Network through Quantum Cryptography and Non-Linear Dimension Reduction Techniques" IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 4, No 6, July 2010.

[2] Anna C. Squicciarini, Heng Xu, and Xiaolong (Luke) Zhang "CoPE: Enabling Collaborative Privacy Management in Online Social Networks", JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 62(3):521534, 2011.

[3] Michael Fire, Roy Goldschmidt, and Yuval Elovici "Online Social Networks: Threats and Solutions" , IEEE COMMUNICATION SURVEYS AND TUTORIALS, VOL. 16, NO. 4, FOURTH QUARTER 2014

[4] Junggab Son , Donghyun Kim and Rahman Tashakkori "A New Mobile Online Social Network based Location Sharing with Enhanced Privacy Protection"

[5] Amirmohammad Sadeghian, Mazdak Zamani and Bharanidharan Shanmugam "Security Threats in Online Social Networks", International Conference on Informatics and Creative Multimedia,2013.

[6] N Gisin et al, "Quantum cryptography", Rev. Mod. Phys., Vol. 74, No. 1, January 2002.

[7] Ryan Dube, (14/04/2017)"Characteristics of Social Networks" [Online].Available:http://socialnetworking.lovetoknow.com /Characteristics_of_Social_Networks

[8] Charlie R Claywell,(14/04/2017)"Advantages and Disadvantages of Social Networking"[online].Available:http://socialnetworking.lovetoknow.com /Advantages_and_Disadvantages_of_Social_Networking

[9] Mary Gormandy White(14/042017)"What Types of Social Networks Exist?"[online].Availble:http://socialnetworking.lovetoknow.com /What_Types_of_Social_Networks_Exist?

A First Introduction to Image Segmentation Techniques

Bindhuja Bhaskaran V B, Nimisha T S, Vijitha K V
Department of Computer Applications,
Vidya Academy of Science & Technology,
Thrissur-680501

Reji C Joy
Associate Professor of Computer Applications,
Vidya Academy of Science & Technology,
Thrissur-680501

***Abstract*—Image segmentation has recently become an essential step in image processing. Image segmentation is a process of partition an image into homogeneous segments based on some criteria. There are new wide varieties of image segmentation technique, some are general purpose and some designed for specific classes of images. In this paper, we discuss three types of image segmentation techniques: Edge-based segmentation, Region based segmentation and thresholding techniques.Here we are also discussing one of the application of image segmentation in fingerprint classification technique.**

***Index Terms*—Image segmentation, edge-based segmentation, region based segmentation, region growing,region merging,region splitting,thresholding technique,fingerprint.**

I. Introduction

Image segmentation is an active research area. Image Segmentation is domain independent partitioning of an image into a set of disjoint regions that are homogeneous and meaningful with respect to some properties such a grey level, color, or texture to enable easy image analysis [1].

Image segmentation approaches are commonly based on one of two fundamental properties of intensity values. In discontinuity based technique image is partitioned by sudden changes in intensity values whereas similarity based technique partition an image by grouping together connected pixels in the regions[6].

In section1 we discussed about the concept of edge based segmentation, the image split by splitting the difference in pixel of the digital image or intensity.

Section2 describes the region based segmentation.Region based segmentation system group pixels together with identical features into regions.

Section 3 describes the thresholding techniques. Thresholding provide an easy an convenient way to perform this segmentation on the basis of the different intensities or colors in the foreground and background region of an image.

Section 4 describes the application of fingerprint classification techniques based on image segmentation.Finger print is emerging as the mot common and trusted biometry for personal identification.

II. Edge Based Segmentation

Edge based segmentation methods are based on the rapid change of intensity value in an image because a single intensity value does not provide good information about edges. Edge detection technique locate the edges where either the first derivative of intensity is greater than a particular threshold or the second derivative has zero crossings. In edge based segmentation methods, first of all the edges are detected and then are connected together to form the object boundaries to segment the required regions[13]. The basic two edge based segmentation methods are: grey histograms and gradients based methods. To detect the edges one of the basic edge detection techniques like sobel operator, prewitt operator and Robert operator etc can be used. Result of these methods is basically a binary image .These is the structural techniques based on discontinuity detection[3].

A. Edge Based Segmentation Methods

1) Roberts Detection: The Roberts Cross operator performs a simple, quick to compute, 2-D spatial gradient measurement on an image. It thus highlights regions of high spatial frequency which often correspond to edges[4]. In its most common usage, the input to the operator is a gray scale image, as is the output. Pixel values at each point in the output represent the estimated absolute magnitude of the spatial gradient of the input image at that point[14].

+1	0
0	-1

Gx

0	+1
-1	0

Gy

Fig.1 robert mask

2) Prewitt Detection: The prewitt edge detector is an appropriate way to estimate the magnitude and orientation of an edge. The prewitt operator is limited to 8 possible orientations, however experience shows that most direct orientation estimates are not much more accurate. This gradient based

edge detector is estimated in the 3x3 neighborhood for eight directions. All the eight convolution masks are calculated. One convolution mask is then selected, namely that with the largest module[14].

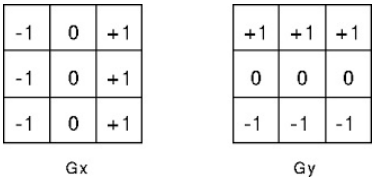


Fig.2 prewitt mask

3) Sobel Detection: The Sobel operator performs a 2-D spatial gradient measurement on an image and so emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient magnitude at each point in an input gray scale image. In theory at least, the operator consists of a pair of 3x3 convolution kernels[14].

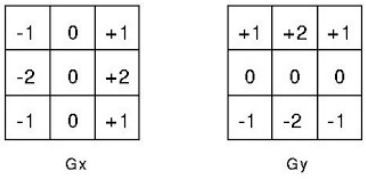


Fig.3 sobel mask

III. Region Based Segmentation

Region based segmentation deals with splitting an image into number of homogeneous regions. Regions in an image are group of connected pixels which have same properties. Each pixel is assigned to any regions. It also require use of appropriate threshold techniques[2].

A. Region Based Segmentation Methods

1) Region Growing: Region growing is a procedure that group pixel or sub regions into large region based on some predefined criterion. Criteria is based on intensity information. In region growing, neighbouring pixels are scanned and joined to a region class of no edges are detected.ths process is iterated for several boundary pixels in the region[9].

First step in region growing is to select a seed point. Initial region create by checking the exact position of seed points. These regions are grownup from seed points to adjacent points depending on some conditions for region membership. When progression of one region completes we can select another seed point which does not belong to any region and start again. The whole process repeated until all pixels fit into a certain region

- **Uniform Blocking** In uniform blocking an image is divided into uniform blocks for processing. we usually use 2x2 blocks if region growing used to work directly or 16x16 if merge split algorithm used.

- **Merge Split Blocking** The merge split is a part of region growing. It need a threshold as input. This threshold defines which blocks can be merged into particular blocks and which block is split into smaller blocks based on the maximum and minimum intensities of every block. The max-min difference is near to zero, then the blocks merged into particular block. A block split into smaller blocks when max-min difference of the block exceeds threshold value.

- **Region Growing By Mean or Max Min** Region growing is completed by detecting properties of each blocks and merging them together based on some condition. In this case, we used two conditions. One condition is to look at the max-min variance and combine to adjacent regions whose max-min variance is near to the seed blocks. These new region is become new seed block and process is repeat. Second condition is using mean value. The mean value of the block can be used to determine which block should be merged.

- **Dissolve** Dissolve algorithm works in combination with the mean-based region growing to merge region that are less than a specified size into adjacent region with the closest mean value. This process reduce number of region reduced by eliminating the less significant region avoiding excessive amount of segmentation.

2) Region merging and Region splitting: Region merging is a an operation eliminate false boundaries by merging adjacent region that belongs to same object.region merging idea is used to combine two regions if the boundaries between them is weak.weak boundary is is one for which the intensity on is either side differ by less than some threshold.

Region splitting operation add missing boundaries by splitting that region that contain part of different objects. splitting or merging might not produce results when applied separately.The better result can be obtained by interleaving merge and split operation[15].

Splitting and merging attempt to divide an image into uniform regions .The basic representation structure is pyramidal ,a square region of size m by m at one level o a pyramid. Usually that algorithm start from the initial assumption that the entire image is a single region, then computes the homogeneity criterion to see if it is TRUE.if FALSE, then square region is split into 4 smaller regions.This process repeated on each of the subregions until no further splitting is necessary.Merging process is used after each split which compare adjacent regions and merge them if necessary[15].

Advantage

- The image can be split progressively according to our demanded resolution because the no of splitting determined bytes
- we can split the image using criteria such as mean or variance of segment pixel.

Disadvantage

- It may produce the blocky segment .Blocky segment problem can be splitting in higher level.

IV. Threshold Based Segmentation

Threshold is one of the widely used methods for image segmentation. It is useful in discriminating foreground from the background. The simplest property that pixels in a region can share is intensity. So, a natural way to segment such regions is through thresholding, the separation of light and dark regions[5].

By selecting an adequate threshold value T, the gray level image can be converted to binary image. The advantage of obtaining first a binary image is that it reduces the complexity of the data and simplifies the process of recognition and classification. The most common way to convert a gray-level image to a binary image is to select a single threshold value T. The binary image should contain all of the essential information about the position and shape of the objects of interest (foreground). Then all the gray level values below this T will be classified as black (0), and those above T will be white (1).

Automatically selected threshold value for each image by the system without human intervention is called an automatic threshold scheme. This is requirement the knowledge about the intensity characteristics of the objects, sizes of the objects, fractions of the image occupied by the objects and the number of different types of objects appearing in the image[7].

A. Threshold Based Segmentation Methods

Threshold technique is one of the important techniques in image segmentation. This technique can be expressed as:

$$T = T[x, y, f(x, y)]$$

where T is the threshold value, x, y are the coordinates of the threshold value point, and $f(x, y)$ are points the gray level image pixels.

Threshold image $g(x, y)$ can be defined as:

$$g(x,y) = \begin{cases} 1 & \text{if} \quad f(x,y) > T \\ 0 & \text{if} \quad f(x,y) \leq T \end{cases}$$

If $g(x, y)$ is a threshold version of $f(x, y)$ at some global threshold $T.g$ is equal to 1 if $f(x, y) \geq T$ and zero otherwise After threshold the image is segmented as follows: Pixels labeled 1 corresponds to object whereas pixels labeled 0 corresponds to the background. This paper applied five threshold techniques.

1) Mean Technique: This technique used the mean value of the pixels as the threshold value and works well in strict cases of the images that have approximately half to the pixels belonging to the objects and the other half to background[8].

2) P-Tile Technique: The p-tile technique uses knowledge about the area size of the desired object to the threshold an image. The P-tile method is one of the earliest threshold methods based on the gray level histogram. It assumes the objects in an image are brighter than the background, and occupy a fixed percentage of the picture area[8]. Let n be the maximum gray level value, $H(i)$ be the histogram of image ($i = 0, \ldots, n$), and P be the object area ratio. The algorithm of the P-tile method is as follows:

Step 1. $S =$ sum $H(i)$
Step 2. Let $f = s$.
Step 3. For $k = 1$ to n :
$f = f - H(k - 1)$
if $f(f/t) < p$ then stop.
Step 4. $T = k$.

where S is total area of image, f is the initialize all area as object area and T is the final threshold value.

3) Histogram Dependent Technique (HDT): The histogram based techniques is dependent on the success of the estimating the threshold value that separates the two homogenous region of the object and background of an image. This required that, the image formation be of two homogeneous and will-separated regions and there exists a threshold value that separated these regions. This technique can be expressed as:

$$C(T) = P_1(T)\sigma_1^2(T) + P_2(T)\sigma_2^2(T)$$

4) EMT Technique : The threshold image by using edge maximization technique (EMT) is used when there is more than one homogeneous region in image or where there is a change on illumination between the object and its background. In this case portion of the object may be merged with the background or portions of the background may as an object[5]. To this reason any of the automatic threshold selection techniques performance becomes much better in images with large homogeneous and well separated regions.

5) Visual Technique: These techniques improve peoples ability to accurately search for target items. These techniques are similar to one another P-Tile technique in that they all use the component segments of original images in novel ways to improve visual search performance but it is different from p-tile don't active when the noise is present in the image[5].

V. Application: Fingerprint Classification Techniques

Fingerprint has been used as the most popular biometric authentication and verification measure because of high acceptability, immutability and uniqueness.The most important stage in automatic fingerprint identification system (AFIS) is a fingerprint classification because it significantly reduces the time taken in identification of fingerprint specifically where the accuracy and speed are critical[11]. Classification allows an input fingerprint to be matched against only by a subset of a database. To reduce the search and space complexity a systematic portioning of database into different class is highly essential[12].

A. The Henry classification System

Sir Francis Galton began the first rigorous study of fingerprint-based identification. Among many contributions to the field, his work contained the first system for fingerprint classification. Galton also provided the first workable fingerprint classification system where fingerprint where classified into 3 classes: the arch, the loop and the whorl[10].

Fig.4 Examples of Galtons three classes.

Edward Henry continued Galtons work on fingerprint classification. Henry subdivided the three main classes into more specific subclasses, namely, arch tented arch, left loop, right loop and whorl. He also introduced the Concept of fingerprint core and delta points.

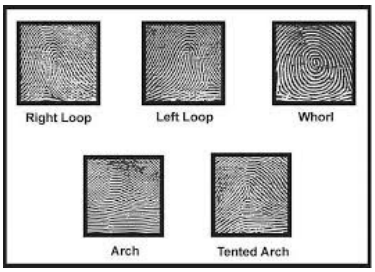


Fig.5 Example of Henrys five classes.

B. Features of Fingerprint Image

Generally a fingerprint image contains two features, viz the global feature and the local feature. The global feature of the fingerprint image is described by structure shape (ridge and valleys) and the singular point as shown in figure. The local feature of the fingerprint consists of the minute details of the ridges. The global feature has the global information which is considered the valid feature used in the design of automatic fingerprint identification system. Another global feature often used by the researcher to distinguish fingerprint classes is the existence and location of singular points.

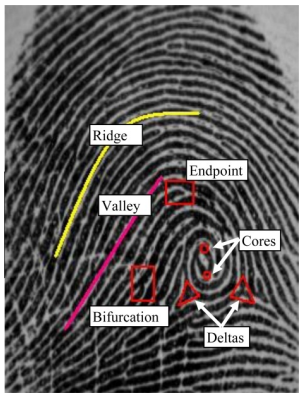


Fig.6 Ridges valleys singular point structure

C. Fingerprint classification process

Before classification can be carried out, the fingerprint pattern has to be transformed into a format which is acceptable for classification.

1) Input fingerprint image:

2) Preprocessing: Fingerprint image is initially preprocessed through consecutive techniques.

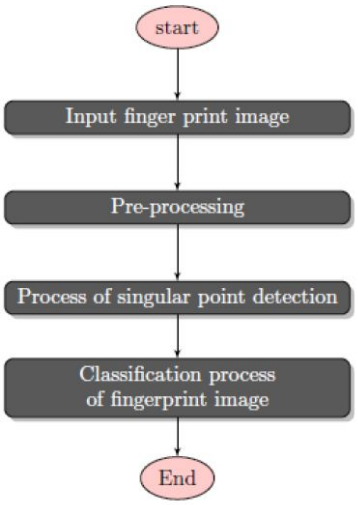


Fig.7 Flow diagram of fingerprint classification process

- **Fingerprint segmentation:** The first step is the segmentation which isolates features that are of similar characteristics. The fingerprint image is split into two regions which are the foreground and the background regions. The foreground region is the area containing ridges and valleys, while the background region corresponds to the fingerprint image borders. Generally, two steps of fingerprint segmentation. They are the block-wise and the bit-wise steps. Block wise step is employed to extract the foreground of fingerprint image from the background. The foreground of the fingerprint extracted is normally corrupted by noise. Bit-wise step is used to remove noise and other unwanted interference by operating in the domains associated with image gray-scale statistical features, local directional features or coherence features.

- **Fingerprint enhancement:** The second step is the enhancement algorithm. Applying enhancement algorithm to fingerprint images are necessary steps towards recovering the quality of fingerprint image. For good quality there must be high contrast between ridges and valleys. There must also be clear continuity in the ridge structures.

3) Intermediate process:

- **Orientation field estimation:** The final step in preprocessing is the orientation field estimation, in which involve the process to convert the fingerprint image to the vector form and improve the smoothing quality of the fingerprint ridges. The accuracy of fingerprint orientation field estimation consider as most important step to detect the singular point as well as to get high accuracy in fingerprint classification system.
- **Singular points detection:** A process of singular point detection is applied on the preprocessed fingerprint image. In this process two kinds of singular points can

be detected, namely core point and delta point. A core point is the turning point of an innermost ridge and delta point is a place where two ridges running side-by-side.

4) Classification process: Finally fingerprints are categorized with classification process based on global features. The scheme classifies fingerprints into five classes namely; plain arch(A), tented arch(T), left loop(L), right loop(R) and whorl(W).

VI. Conclusion

In this paper, we discussed about methods for various applications. There are suitable segmentation methods that can be applied.

Edge based segmentation is the most suitable methods. When the edge of image is clear in edge based segmentation, all the edges are detected and then grouped together to form a boundaries. The main reason for using Robert cross operator is that it is very quick to compute 2D spatial gradient but disadvantage is that since it uses such a small kernal, it is very sensitive to noise. sobel operator is better than robert operator. It is used to find approximate absolute gradient magnitude at each point in a gray scale image. But prewitt operator is mainly used to estimate magnitude and orientation of an edge.

Region based segmentation is suitable when the edge of the image is not clear because of noise. Region growing, group pixel or sub region into large regions based on predefined criteria. Region growing is an efficient method in region based segmentation, thresholding is useful in separating foreground from the background.

The main goal of splitting and merging is to distinguish homogeneity of an image.Region splitting operations add missing boundaries by splitting regions that contain parts of different objects.Region merging operation eliminate false boundaries by merging adjacent regions that belong to the same object.

Thresholding is a common way to convert a gray scale image into binary image. Different technique are used to find threshold value. P-tile technique assume the object in an image are brighter than background.

Different segmentation technique are suitable for different type of images.

References

[1] M.S. Sonawane and C.A. Dhawale,*"A Brief Survey on Image Segmentation Methods",*International Journal of Computer Applications (0975 8887),National conference on Digital Image and Signal Processing, DISP 2015

[2] S.Karthick and Dr.K.Sathiyasekar*"A Survey Based on Region Based Segmentation",*International Journal of Engineering Trends and Technology (IJETT) Volume 7 Number 3- Jan 2014

[3] N. Senthilkumaran and R. Rajesh*"Edge Detection Techniques for Image Segmentation A Survey of Soft Computing Approaches",* International Journal of Recent Trends in Engineering, Vol. 1, No. 2, May 2009

[4] Rafael C. Gonzalez and Richard E.Woods, *"Digital Image Processing(2nd Edition)"*

[5] Salem Saleh Al-Amri,N.V. Kalyankar and Khmitkar S.D, *"Image Segmentation By Using Threshold",* Journl Of Computing, Volume 2,May 2010

[6] Muhammad Waseem Khan *"A Suvey:Image Segmentation Techniques",* Interntional Journal of Future Computer and Communication, Vol 3,No.2,April 2014

[7] Francis H.Y.Chan, F.K Lam, and Hui Zhu, *"Adaptive Thresholding by Variational Method",* IEEE Transaction On Image Processing,Vol 7,No. 3, March 1998

[8] K.H. Liang and J.J.W Mao, *Image Thresholding by Minimizing the Measures of Fuzziness,*Pattern Recognition, Vol.28, No.1,PP.41-51, 1995.

[9] H.G.Kaganami,Z.Beij, *Region Based Detection versus Edge Detection,*IEEE Transactions on Intelligent information hiding and multimedia signal processing, pp. 1217-1221, 2009.

[10] Alaa Ahmed Abbood and Ghazali Sulong, *"Fingerprint Classification Techniques: A Review",*IJCSI International Journal of Computer Science Issues, Vol. 11, Issue 1, No 1, January 2014 ISSN (Print): 1694-0814 — ISSN (Online): 1694-0784

[11] D. Maltoni, *"A Tutorial on Fingerprint Recognition",* pp. 4368, 2005

[12] A. K. Jain, S. Prabhakar an S. Member, and L. Hong, *"A Multichannel Approach to Fingerprint Classification",* vol. 21, no. 4, pp. 348359, 1999.

[13] N. Senthilkumaran and R. Rajesh, *"A Study on Edge Detection Methods for Image Segmentation",* Proceedings of the International Conference on Mathematics and Computer Science (ICMCS-2009), 2009,Vol.I, pp.255-259.

[14] N. Senthilkumaran and R. Rajesh, *"Edge Detection Techniques for Image Segmentation - A Survey",* Proceedings of the International Conference on Managing Next Generation Software Applications (MNGSA-08), 2008, pp.749-760.

[15] Waseem Khan *"Image Segmentation Techniques: A Survey"* Journal of image and graphics Vol 1,No. 4,December 2013

Security Issues in E-Commerce: A Study

Chowgule Kavita Kashinath, Devika Chandrasekharan, Thadathil Riya Ravindran
Department of Computer Applications
Vidya Academy of Science and Technology
Thrissur 680501

Dijesh P
Associate Professor of Computer Applications
Vidya Academy of Science and Technology
Thrissur 680501

***Abstract*—E-commerce is an important term which revolutionised the entire buying and selling aspects of customers.in e-commerce scenario customers can utilize the powerful internet as a medium for fulfilling their needs. A person can understand the importance of the term ecommerce while glancing through the advantages it can offer . This seminar paper is an attempt to cover various aspects like security issues of ecommerce, solutions for security issues, legal and policy aspects of technology solutions in India, issues in electronic payment systems, bio metric security etc.**

***Index Terms*—E-commerce, payments, security, threats, guidelines, policy aspect, biometrics.**

I. Introduction

E-commerce had brought a lot of changes in the world of business. This enable customers to perform their selling and buying needs from any part of the world and the e-commerce brought the entire world into the finger tips. This E-commerce has enabled customers to save their precious time by avoiding them to visit shops as in traditional commerce and allow them to purchase commodities while performing other jobs also. It also has played a very good role in upbringing the small scale industries into the limelight by enabling their promotional activities reach any corner of the world.Even though the E-commerce offers a lot of advantages it also has disadvantages.

A. What is E-commerce?

Trading products by using internet-commerce often come in wide variety of forms like B2B, B2C, C2C, C2B, non-business and intra business EC. E-commerce also offers great advantages like reachability, low cost, interactiveness, fast and convenient. But it also faces some dis advantages like insufficient band width, need of special web servers, in compatibleness.

B. Common Modes of Payment in E-commerce

The following are the common modes of payment in e-commerce.

1) **Credit card**
   This is the most used mode of payment.it is a plastic card which is of rectangular shape contain unique card number, account number and magnetic strip.
2) **Debit card**
   Thisis similar to credit card in which amount is deducted when user make transaction.
3) **Mobile Payments**
   A popular payment method in countries with low credit card and banking penetration, mobile payments offer a quick solution for customers to purchase on ecommerce websites. Mobile payments are also commonly used on donation portals, browser games, and social media networks such as dating sites, where customer can pay with SMS. Majority of mobile payments are done in the Asia Pacific, with 64 million users expected by the year 2016.
4) **Bank Transfers**
   Customers enrolled in an internet banking facility can do a bank transfer to pay for online purchases. A bank

Fig. 1. E-commerce

Fig. 2. Modes of Payments

transfer assures customers that their funds are safely used, since each transaction needs to be authenticated and approved first by the customers internet banking credentials before a purchase happens.

5) **E-wallets**
An e-wallet stores a customers personal data and funds, which are then used to purchase from online stores. Signing up for an ewallet is fast and easy, with customers required just to submit their information once for purchases. Additionally, ewallets can also function in combination with mobile wallets through the use of smart technology such as NFC (near field communication) devices. By tapping on an NFC terminal, mobile phones can instantly transfer funds stored in the phone.
6) **Prepaid Cards**
An alternative payment method, commonly used by minors or customers with no bank accounts. Prepaid cards come in different stored values for customers to choose from. Online gaming companies usually make use of prepaid cards as their preferred payment method, with virtual currency stored in prepaid cards for a player to use for in-game transactions. Some examples of prepaid cards are Mint, Ticket surf, Paysafecard, and Telco Card. It appears that age rather than income is the trait that affects the adoption of prepaid cards, according to Troy Lands research.
7) **Direct Deposit**
Direct deposits are when customers instruct their banks to pull funds out of their accounts to complete online payments. Customers usually inform their banks on when funds should be pulled out of their accounts, by setting a schedule through them. A direct deposit is a common payment method for subscription-type services such as online classes or purchases made with high prices.
8) **Cash**
Fiat, or physical cash, is a payment method often used for physical goods and cash-on-delivery transactions. Paying with cash does come with several risks, such as no guarantee of an actual sale during a delivery, and theft. Even though e-commerce is powerful scenario security is a major concern. Now first we will look security issues of ecommerce.

II. Security Issues in E-commerce

The wide spreading use of the e-commerce has led to variety of security attacks. An e-commerce site should satisfy four criteria listed below to ensure its safety.

1) **Authenticity**
To safe guard the data from unauthorised parties. In TCP/IP the basic means of verifying the identity of a user is a password, but passwords can be guessed and intercepted. Internet Protocol (IP) addresses can also be screened to prevent unauthorized access. An authentication service must be able to guard against masquerade and replay attacks.
2) **Confidentiality**
To protect data from unwanted tampering. In some cases, it may also be necessary to keep the sender and receiver informed as breaches to confidentiality can occur both during and after transmission. Once a message is received, sender must be assured that its contents remain confidential. Encryption and injecting dummy data into the network are the methods to ensure confidentiality
3) **Integrity**
Ensures the validity of a persons identification. The TCP/IP transmits data packets in plain text. Because the packets associated with a given message often transverse a number of routers and lines as they move from client to server and back again, they are susceptible to capture and modification while en route. Integrity is achieved by systematically adding some redundancy to the data through hash functions. The resulting hash value is encrypted to ensure data integrity
4) **Non-repudiation**
It act as a proof so that the sender cannot deny that message is not from him. The key to nonrepudiation is digital signature that proves that sender and receiver were involved in the exchange. A complete transaction involves the sender sending a digitally signed message. The receiver then sends the sender a digitally signed acknowledgement, containing either the whole original message or its hash value.

III. Privacy, Authentication, Integrity and Non-repudiation

Privacy is an important issue in e-commerce. It is the responsibility of e-commerce companies to protect the credentials of the user. Without ensuring the privacy the Companies would not be able to achieve the loyalty and trust worthiness of customers.

Authentication is an important measure which enables the customers to ensure that the person to whom they are sending the information is valid. Integrity make sure that data is not being subjected to any unwanted changes. Non-repudaition is a technique to ensure the message is from correct sender.

Digital signature is a very powerful security measure. An electronic signature may be defined as “any letters, characters, or symbols manifested by electronic or similar means and executed or adopted by a party with the intent to authenticate a writing” (Blythe, 2006). It involves the use of private key and public key for security.

Technical attacks possesses a major threat for e-commerce. Denial of security attack is an important problem. The denial of security attack can be identified by the several symptoms like poor network performance, website unavailability, problems with the access and the wide increase in spam mails. There are wide variety of way in which the dos can be carried out like ICMP Flood, tear drop attack and plashing.

There is also distributed denial of service (ddos) attacks. The ddos attack mainly focuses on attacking the bandwidth of network. There are also brute force attacks Brute force

attacks relies on various possibilities.In non-technical attacks Phishing is a severe concern. Phishing attacks starts when a victim receive a fraudulent mail. The link which is embedded the email lead to the fake site. Social engineering Social engineering is the art of manipulating people into performing actions or divulging confidential information (2). it can carried out in several technique like pretexting, IVR and the Trojan horses.

IV. Threats in E-Commerce

There are several threats to e-commerce which make malicious insiders can cause harm to security and there are 2 type of attacks namely active and passive. Passive attack involve change of actual data while active attack is to eaves drop the data. The attacks involve malicious code attack which include worms [replicate through internet] and viruses [need a host file to attack and result in loss of resources]. Denial of Service (it deny customer from obtaining resources. Now we have gone through security issues. The next topics this paper is going to focus on:

V. Guidelines for Secure E-commerce

The following are some guidelines for secure e-commerce.

1) **Shop at secure website**
   Secure sites use encryption technology to transfer information from your computer to the online merchant's computer. Encryption scrambles the information you send, such as your credit card number, in order to prevent computer hackers from obtaining it en route. The only people who can unscramble the code are those with legitimate access privileges.
2) **Research the website before you order**
   Do business with companies you already know. If the company is unfamiliar, do your homework before buying their products. If you decide to buy something from an unknown company, start out with an inexpensive order to learn if the company is trustworthy.
3) **Read shopping website privacy and security policy**
   Every reputable online Web site offers information about how it processes your order. It is usually listed in the section entitled Privacy Policy. You can find out if the merchant intends to share your information with a third party or affiliate company. Do they require these companies to refrain from marketing to their customers? If not, you can expect to receive spam (unsolicited email) and even mail or phone solicitations from these companies.
4) **Decide which mode is safest, and disclose the bare facts when we order**
   The safest way to shop on the Internet is with a credit card. In the event something goes wrong, you are protected under the federal Fair Credit Billing Act. You have the right to dispute charges on your credit card, and you can withhold payments during a creditor investigation. When placing an order, there is certain information that you must provide to the web merchant such as your name and address. Often, a merchant will try to obtain more information about you. They may ask questions about your leisure lifestyle or annual income. This information is used to target you for marketing purposes. It can lead to "spam" or even direct mail and telephone solicitations.
5) **Be aware of the cookies and behavioural marketing**
   Online merchants as well as other sites watch our shopping and surfing habits by using "cookies," an online tracking system that attaches pieces of code to our Internet browsers to track which sites we visit as we search the Web. "Persistent" cookies remain stored on your computer while "session" cookies expire when you turn the browser off. Online merchants use cookies to recognize you and speed up the shopping process the next time you visit. You may be able to set your browser to disable or refuse cookies but the tradeoff may limit the functions you can perform online, and possibly prevent you from ordering online. Generally, you will need to enable session cookies to place an order.
6) **Never give out social security number**
   Providing your Social Security number is not a requirement for placing an order at an online shopping site. There is no need for the merchant to ask for it. Giving out your Social Security number could lead to having your identity stolen.
7) **Keep your passwords safe**
   Many online shopping sites require the shopper to log-in before placing or viewing an order. The shopper is usually required to provide a username and a password. Never reveal your password to anyone. When selecting a password, do not use commonly known information, such as your birthdate, mother's maiden name, or numbers from your driver's license or Social Security number. Do not reuse the same password for other sites, particularly sites associated with sensitive information. The best password has at least eight characters and includes numbers and letters.
8) **Beware of phishing messages**
   Identity thieves send massive numbers of emails to Internet users that ask them to update the account information for their banks, credit cards, online payment service, or popular shopping sites. The email may state that your account information has expired, been compromised or lost and that you need to immediately resend it to the company. Some emails sent as part of such phishing expeditions often contain links to official-looking Web pages. Other times the emails ask the consumer to download and submit an electronic form.
9) **Check website address**
   The address bar at the top of your device's screen contains the web site address (also called the URL, or Uniform Resource Locator). By checking that address, you can make sure that you are dealing with the correct company. Dont click on any link embedded within a potentially suspicious email. Instead, start a new Internet

session by typing in the links URL into the address bar and pressing Enter to be sure you are directed to a legitimate Web site.

10) **Always print or save copies of orders**
After placing an order online, you should receive a confirmation page that reviews your entire order. It should include the costs of the order, your customer information, product information, and the confirmation number. We recommend you print out or save a copy of the Web page(s) describing the item you ordered as well as the page showing company name, postal address, phone number, and legal terms, including return policy. Keep it for your own records for at least the period covered by the return/warranty policy.
11) **Pay attention to shipping facts**
Under the law, a company must ship your order within the time stated in its ad. If no time frame is stated, the merchant must ship the product in 30 days or give you an "Option Notice." This gives you an opportunity to cancel the order and receive a prompt refund, or agree to the delay.
Here are key shipping questions to ask:
    - Does the site tell you if there are geographic or other restrictions for delivery?
    - Are there choices for shipping?
    - Who pays the shipping cost?
    - What does the site say about shipping insurance?
    - What are the shipping and handling fees, and are they reasonable?
12) **Learn the Merchant's Cancellation, Return and Complaint-Handling Policies**
Even under the best of circumstances, shoppers sometimes need to return merchandise. Check the Web site for cancellation and return policies.
13) **Beware of identity theft**
As online shopping becomes more common, there will be more cases of identity theft committed over the Internet. Imposters are likely to obtain their victims' identifying information using low-tech means like dumpster diving, mail theft, or workplace access to SSNs. But they are increasingly using the Web to apply for new credit cards and to purchase goods and services in their victims' names.
14) **Know how online auction operate**
Online auctions connect buyers and sellers, allowing them to communicate in a bidding process over items for sale. Many people are drawn to online auction sites because they allow you to buy items at discounted prices. And they offer a chance to sell some of your unneeded or unwanted possessions to raise extra money. For the most part, online auction sites are a safe way to exchange goods. But it makes sense to be cautious and aware.

VI. Cryptography Based Security

The cryptography based security involve the secure socket layer (in which confidential information like credit card number, debit card number can be transmitted safely), secure commerce protocol(both customer and merchant can issue certificate to initiate transaction),security certification(it deals with confirming characteristics of person, organisation. (SAP, Seimens and Verizon are the famous certificate providers). Digital signatures(enhance security. it is to prove that message id from original sender)and public key infrastructure(allow distribution and use of public key and digital certificates to provide security).

VII. Legal and Policy Aspects of Technology Solutions in India

Before moving to the legal aspects there is need to gain a knowledge on the solutions of e-commerce.

1) **Encryption**
Technique of changing the plain text into the unintelligible cipher text by utilizing a key to secure the information from falling into wrong hands. Ensure the validity and confidentiality and privacy of messages. The encryption technique is not limited to the binary data even video and audio can be encrypted.
2) **Digital signature**
Ensure that message is from correct source and involve the usage of private key and public key for encryption and decryption.
3) **Fire walls**
There are mainly two types of firewalls namely dual homed gateway and screen host. Dual homed gateway involve the bastation gateway and network cards while screen host relies on the usage of the network routers.
4) **Intrusion detectors**
De-activate intruders.
5) **Virtual Private Networks**
Combines techniques like encryption, authentication and protocol tunnelling to ensure better communication.

VIII. Policy Aspects in India

The e-commerce technologies and related business developments are growing at a rapid pace but the relevant laws and policies lag behind them. Admissibility of electronic records in courts poses a challenge to e-commerce. However, there are certain considerable developments in legal and policy aspects of e-commerce. Some of them are the Bill of Lading Act, 1856 the Indian Contract Act, 1872 the Negotiable Instruments Act, 1881 the Bankers Books Evidence Act, 1891 the Sale of Goods Act, 1930; the Banking Regulation Act, 1949 the Hire-Purchase Act, 1972 UNCITRAL, 1996 the Information Technology Act, 2000 the Insurance Regulation and Development Authority Act, 2000 Uniform Customs and Practices 500 and Electronic Uniform Customs and Practices.

The legal and policy aspects of technology involve several rules and policies which is listed below.

1) **Indian contract act**-binds parties involved and specifies for valid contract.
2) **Negotiable instrument act**-person to whom money is transferred become eligible for further transfer.

3) **Bankers book act**-which enables the bank entries to select as evidence.
4) **Sales of goods act**-act which deals with goods instead of labour.
5) **Banking regulation act**-regulation of the banking services.
6) **Hire purchase act**-a person can either pay money to buy goods or can hire them.Incase of hiring they have to return goods and end liability.
7) **UNICITRAL act**-deals with international business.

IX. Electronic Customs and Practices

The e-UCP was proposed in 2002 as a supplement to the UCP 500 to deal with the presentation of all electronic or part electronic documents. It was developed due to the strong sense of feeling at the time that banks and corporates together with the transport and insurance industries were ready to make that leap into the electronic world of document and data delivery. As a supplement, the e-UCP does not alter the existing articles of UCP 500. It includes rules regarding such matters as the format for such documents, how they are presented and what happens if they are corrupted. The e-UCP deals only with electronic presentation of documents and not with electronic creation or delivery of LCs. Even if a LC is subject to e-UCP, the LC may stipulate that the beneficiary present some or all of the documents on paper. The e-UCP anticipates that there will be mixed presentations of paper and electronic documents. Further, it anticipates that presentation of electronic documents will happen individually rather than all the documents in a batch at the same time. Even a LC that calls entirely for paper documents may be made subject to the e-UCP. The only change will be that the LC must specify acceptable formats for each document. Like UCP 500, the e-UCP does not specify any specific document formats. It is left to the parties to decide this bi-laterally. The e-UCP also requires all electronic documents to follow the pillars of e-security However, standardization of e-UCP has been found to be difficult

There are also acts like information technology act (enable electronic information as alternative to paper evidence), insurance regulation act and uniform custom and practise act. The legal aspects which mentioned above provide security. But for better security e commerce is going for more scientific technologies.

X. Payment Gateways

There are several steps in conducting a transaction through a payment gateway. Figure 3 shows diagramatically the various steps in conducting a transaction through a payment gateway.

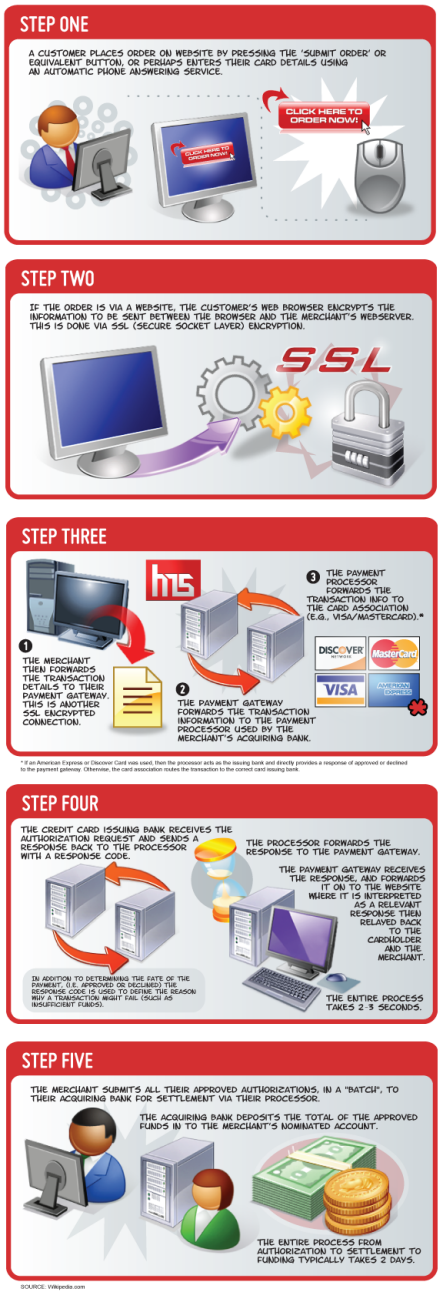


Fig. 3. Different steps in payment gateway

XI. Biometric Security

The methodology include biometric authentication(including identifying various features like face,finger,hand. It can operate and iris etc.It acquire pattern recognition system operate by acquiring biometric data and extract feature set and compare it. It can operate in both verification mode and identification mode.

A. Types of Biometric Security

1) **Dna**
The recognition sequence, sometimes also referred to as recognition site, of any DNA-binding protein motif that exhibits binding specificity, refers to the DNA sequence (or subset thereof), to which the domain is specific. Recognition sequences are palindromes. The

transcription factor Sp1 for example, binds the sequences 5'-(G/T)GGGCGG(G/A)(G/A)(C/T)-3', where (G/T) indicates that the domain will bind a guanine or thymine at this position. The restriction endonuclease PstI recognizes, binds, and cleaves the sequence 5'-CTGCAG-3'. However, a recognition sequence refers to a different aspect from that of recognition site. A given recognition sequence can occur one or more times, or not at all on a specific DNA fragment. A recognition site is specified by the position of the site

2) **Ear**
The outer ear is an emerging biometric trait that has drawn the attention of the research community for more than a decade. The unique structure of the auricle is long known among forensic scientists and has been used for the identification of suspects in many cases. The next logical step towards a broader application of ear biometrics is to create automatic ear recognition systems.
3) **Face recognition**
A face recognition system is a computer application capable of identifying or verifying a person from a digital image or a video frame from a video source. One of the ways to do this is by comparing selected facial features from the image and a face database. It is typically used in security systems and can be compared to other biometrics such as fingerprint or eye iris recognition systems.Recently, it has also become popular as a commercial identification and marketing tool.
4) **Gait**
Gait analysis is the systematic study of animal locomotion, more specifically the study of human motion, using the eye and the brain of observers, augmented by instrumentation for measuring body movements, body mechanics, and the activity of the muscles. Gait analysis is used to assess, plan, and treat individuals with conditions affecting their ability to walk. It is also commonly used in sports biomechanics to help athletes run more efficiently and to identify posture-related or movement-related problems in people with injuries.
5) **Iris**
Iris recognition is an automated method of biometric identification that uses mathematical pattern-recognition techniques on video images of one or both of the irises of an individual's eyes, whose complex patterns are unique, stable, and can be seen from some distance. Retinal scanning is a different, ocular-based biometric technology that uses the unique patterns on a person's retina blood vessels and is often confused with iris recognition. Iris recognition uses video camera technology with subtle near infrared illumination to acquire images of the detail-rich, intricate structures of the iris which are visible externally. Digital templates encoded from these patterns by mathematical and statistical algorithms allow the identification of an individual or someone pretending to be that individual. Databases of enrolled templates are searched by matcher engines at speeds measured in the millions of templates per second per (single-core) CPU, and with remarkably low false match rates.
6) Hand and finger geometry

B. Fingerprint Technology

The fingerprints are unique because it is unique.(4) This biometric technology uses the pattern of friction ridges and valleys on an individual's fingertips. These patterns are very much unique to a specific individual. There is no need to type password a touch will do its job and it also provide instant access (typically less than 1 sec.). Poor ridges are just the problem. But capturing by 3 pixel camera produce distortion free images and solve the problems like cleanliness maintenance etc.

Pre-processing is an important step prior to fingerprint feature extraction. The generic process of pre-processing encompasses segmentation, enhancement, and core point detection. Here the captured fingerprint image is in RG B format is first converted to gray scale. This gray scale image is input to the normalization process. Finger print segmentation is necessary to eliminate the undesired background and reduce the size of the input data. (4)

In the secured transaction android is important area. It includes an operating system, application framework, and core applications. The Android software stack is built on the Linux kernel, which is used for its device drivers, memory management, process management, and networking. The next level up is formed by Android native libraries. A fuzzy logic controller consists of three main operations: Fuzzification, Inference Engine and Defuzzification. The input sensory (crisp or biometric) data are fed into fuzzy logic rule based system here physical quantities are represented into biometric variables. (5)

C. Fingerprint Identification

The steps involved in the finger print identification are

Finger print matching: include minutae based and co-relation based methods

Finger print classification.

Finger print image enhancement.

D. How Fingerprint Scanner Works

A row of LEDs scans bright light onto the glass (or plastic) surface on which your finger is pressing (sometimes called the platen).

The quality of the image will vary according to how you're pressing, how clean or greasy your fingers are, how clean the scanning surface is, the light level in the room, and so on.

Reflected light bounces back from your finger, through the glass, onto a CCD or CMOS image sensor.

The longer this image-capture process takes, the brighter the image formed on the image sensor.

If the image is too bright, areas of the fingerprint (including important details) may be washed out completelylike an indoor

digital photo where the flash is too close or too bright. If it's too dark, the whole image will look black and details will be invisible for the opposite reason.

An algorithm tests whether the image is too light or too dark; if so, an audible beep or LED indicator alerts the operator and we go back to step 1 to try again.

If the image is roughly acceptable, another algorithm tests the level of detail, typically by counting the number of ridges and making sure there are alternate light and dark areas (as you'd expect to find in a decent fingerprint image). If the image fails this test, we go back to step 1 and try again.

Providing the image passes these two tests, the scanner signals that the image is OK to the operator (again, either by beeping or with a different LED indicator). The image is stored as an acceptable scan in flash memory, ready to be transmitted (by USB cable, wireless, Bluetooth, or some similar method) to a "host" computer where it can be processed further. Typically, images captured this way are 512512 pixels (the dimensions used by the FBI), and the standard image is 2.5cm (1 inch) square, 500 dots per inch, and 256 shades of gray.

The host computer can either store the image on a database (temporarily or indefinitely) or automatically compare it against one or many other fingerprints to find a match.

XII. Conclusion

This paper is an attempt to cover some of the major aspects of e-commerce and e-commerce payment mechanism. The e-commerce is a term that brought entire change in this digital world. Like 2 sides of coin it also suffer from advantage as well as disadvantage and security issues. If some security measures were followed e commerce will surely bring a new trading experience.

References

[1] Eamonn O Raghalliagh, "Major Security Issues in E-Commerce". (Online) Available: https:// www.slideshare.net/ EamonnORagh/

[2] Niranjana Murthy and D Dharmendra Chahar, "The Study of E-Commerce Security Issues and Solutions". (Conference Paper Available Online)

[3] Mangala Belkhade,Veena Gulhane,Dr.Preeti Bajaj, "Biometric Mechanism For Enhanced Security of Online Transaction Android System: A Design Approach". (Conference Paper Available Online)

[4] V Prasanth Kumar and Mrs Evelyn Brindha, "Personal authentication using finger print biometric". (Online) Available: https://www.rroij.com

[5] Paymentwall Team, "Types of payment methods for ecommerce",(online) Available: http://blog.paymentwall.com

[6] "How Payment Gateways Work", (online) Available: https://www.hostmerchantservices.com

[7] Biometric fingerprint scanners,(Online) Available:http://www.explainthatstuff.com

[8] Dr. Harman Preet Singh, "E-Commerce Security: Legal and Policy Aspects of Technology Solutions in India". (Online) Available: https://www.researchgate.net>publication

Cross Platform Mobile Application Development Tools : A Comparative Study

Fayisa P H, Nayana Venugopal, Tansey T C
Department of Computer Applications
Vidya Academy of Science and Technology
Thrissur - 680501

Siji K B
Assistant Professor in Computer Applications
Vidya Academy of Science and Technology
Thrissur - 680501

***Abstract*—Mobiles are an integral part of daily life. Mobile devices and mobile computing have made tremendous advances and become ubiquitous in the last few years. The existence of different mobile operating systems with different programming languages and development tools can be a problem when someone wants to release an application in as many platforms as possible. A solution that could generate an application for several platforms (multi or cross-platform) without compromising the overall quality of the application would decrease the time to market of the application and increase enormously the number of potential users. Cross-platform mobile development is a new area of software engineering that allows to reduce development time and cost. This paper presents a pragmatic comparison among six popular cross platform tools, which include payed and freeware tools.**

***Index Terms*—cross platform,mobile application, cross-platform development tool, operating system, native api,smartphone**

I. Introduction

Today mobile devices are everywhere. Nearly everyone has a smartphone, and often a tablet as well. Besides making calls there are many other features which are gaining popularity like Camera, Music, Global Positioning System (GPS), Accelerometer, etc. These kinds of built-in features are provided by all major available mobile Operating Systems (OSs), such as Android, BlackBerry, iPhone Operating System (iOS), Symbian, Windows Mobile/Phone.

Developing mobile applications for several platforms with the native development tools typically means that the development cycle must be repeated for each individual platform. Each platform has its own family of devices, programming languages, APIs and distribution markets, and requires a specific set of skills to support it.

Native applications are developed for a specific target platform, using that platform's SDK and frameworks, and the app is tied to that specific environment. For example, an Android application is developed with Java using the Android SDK and the APIs provided by Android and uses platformprovided elements for rendering the UI. If the developer wants to support multiple platforms with pure native applications, the applications need to be developed separately for each platform.

TABLE I
Major mobile platforms.

Vendor	Operating System	Programming Language	Development Environment	Application Store
Google	Android	Java	Eclipse	Google Play
Apple	iOS	Objective-C	Xcode	iPhone App store
Microsoft	Windows iphone	Visual c#	VisualStudio	Windows Phone Marketplace
RIM	Blackberry OS	Java	Eclipse	Blackberry App World

Applications developed for one platform with traditional development methods only work on that platform, and supporting multiple platforms requires developing the application separately for each of the platforms. Different cross-platform methods have been introduced as a solution to this problem. They allow deploying the application for multiple platforms from a single code base. Cross platform mobile development tools are gaining popularity in the world due to their characteristic to compile the application source code for multiple supported OSs. Such tools are mainly depending on web programming languages like HyperText Markup Language (HTML), JavaScript and Cascading Style Sheets (CSS) with some native wrapper code for accessing native Application Program Interfaces (API) like Camera, Contacts, etc. The application development is very easy and time saving with these tools. For example, DragonRad is providing Drag and Drop (D&D) features, which require reduced programming skills to develop applications. There are plenty of such tools available now on the market, thus creating confusion among developers on which one to use and which one to skip. In the future, crossplatform tools can bring a drastic change in the business model of mobile OSs, especially due to the fact of reduction of development costs of new applications. Therefore, this paper proposes a survey on six major available cross-

platform development tools on the market which are PhoneGap,senchaTec,Titanium,xemarin DragonRad and JQuery.

II. Benefits and Some Properties of Cross Platform Tools

To meet the needs of developers, cross platform mobile development tools have been developed with the purpose to give them the possibility to write the application source code once and run it on different OSs. Benefits that these tools have brought are:

- Reduction of required skills for developers to develop applications due to the use of common programming languages.
- Reduction of coding, because the source code is written once and it is compiled for each supported OS.
- Reduction of development time and long term maintenance costs.
- Decrement of API knowledge, because with these tools is not needed to know the APIs of each OS, but only the APIs provided by the selected tool.
- Greater ease of development compared to building native applications for each OS; and
- Increment of market share for the corresponding business model with the advantage to raise the Return On Investment (ROI).

Some of the properties provided by each tool that have been taken in consideration are:

- Mobile Operating Systems supported to understand possible effects on respective business of use;
- Tool licences offered to evaluate the terms and conditions of use;
- Programming languages offered to developers for building applications;
- Availability of APIs provided with the aim to get an idea of different hardware parts accessible in the OS;
- Accessibility to native APIs to compare how it is possible to access them from each tool;
- Architecture provided for the development process ofthe application; and
- Integrated Development Environments available for developing applications.

III. PhoneGap

PhoneGap was originally created by Nitobi Software, which was acquired by Adobe Systems in 2011 and named as Apache Cordova. It is an open source cross-platform mobile development framework under Apache License Version 2.0. PhoneGap can be used for developing free, open-source or commercial applications. PhoneGap supports a wide range of mobile platforms, including Android, iOS, BlackBerry and Windows Phone.

PhoneGap is a useful solution for building mobile applications using modern web programming languages, such as HTML, HTML5, CSS, CSS3 and JavaScript, and the functionality of SDKs instead to use less-known languages such as Objective-C or other languages . It has the benefit to bring many advantages to skilled developers and especially to attract web developers . Essentially, PhoneGap is a wrapper that allows developers to enclose applications written in known programming languages into native applications. Moreover, as each valid open-source software it is composed by many components and extensions. PhoneGap applications are hybrid, which means that they are not purely native or web-based. The meaning of not purely native comes from the layout rendering that is done via web-view instead of the native language of the OS, whereas not purely web-based comes from the lack support of HTML in some functions . Besides, PhoneGap also offers the possibility to extend the tool by developing own plug-ins.

Adopting a cross-platform approach the applications building and maintenance can be enhanced because developers have to write a single source code for any mobile OS supported by the tool. PhoneGap does not provide an IDE to develop applications, but developers have to write the source code with an IDE and port the it on other IDEs (e.g. Eclipse for Android, XCode for iOS, etc.). This approach does not allow developers to have a centralized development environment, so the effort required to compile the source code and produce the executable application (final product) is high. Thanks to the use of different IDEs for the development, PhoneGap can be performed on different PC OSs such as Mac, Linux and Microsoft Windows. Unfortunately, sometimes there are some exceptions because not all IDEs are compatible with all PC OSs.

A. PhoneGap architecture

The PhoneGaps architecture is composed mainly of 3layers: Web Application, PhoneGap, and OS and native APIs.

In Fig. 1 the top layer represents the application source code.The central layer is composed by JavaScript and native APIs. Mainly, this layer is responsible for the interfacing between web application and PhoneGap layers. Furthermore, it also takes care of the interfacing between JavaScript APIs that are used by the application with native APIs that are used by mobile OSs. The functionality of this layer is to maintain the relationship between JavaScript APIs and native APIs of each mobile OS. PhoneGap provides JavaScript APIs to developers that allow the access to advanced device functionality, such as Accelerometer, Barcode, Bluetooth, Calendar, Camera, Compass, Connection, Contacts, File, GPS, Menu, NFC, etc.

IV. DragonRad

DragonRad 5.0 is a cross platform mobile application development platform by Seregon Solutions Inc. and distributed under a commercial license. It allows developers to design, manage and deploy mobile applications once and use it across iOS, Android,BlackBerry and Windows Mobile . The tool focuses on database driven mobile enterprise applications with easy and wide range of databases support. It provides the D&D environment which help developers to save programming time and to create logics. DragonRad provides their own built IDE, that can be configured for different simulators like iOS,

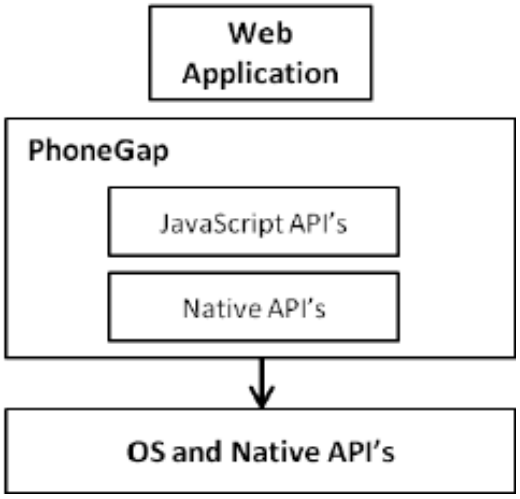


Fig. 1. Interfacing layers of the PhoneGap architecture

Android, BlackBerry, Windows Mobile etc. As DragonRad has host-client architecture, it is required to setup server and database based on the needs of developers but it also comes incomplete package with all prerequisites of server and database like Tomcat, MySql etc. DragonRad is commercial tool with the support to its own language D&D, the possibilities ofextension in terms of adding plugins and other support to the framework are quite limited.

DragonRad facilitates the integration and synchronization of database system with native functions of above defined mobile OSs, such as Contacts, Calendar, Geolocation, Menu and Storage. The Architecture of DragonRad mainly composed of three major components.

A. DragonRad Designer

It is a D&D visual environment or GUI for developers to design, develop and install mobile applications. Features of D&D are not only helping developers to design applications, but also reduce the efforts for maintenance and coding .

B. DragonRad Host

DragonRad host component could be run on either Linux or windows server which fill the gap between database of enterprise and mobile applications. It helps to maintain the communication with mobile device, which also includes query of transaction during network unavailability. It also plays the role to establish problem free connection with database access and updates with synchronization. The following list summarizes the most relevant features of DragonRad host.

- Taking data query from the device
- Executing data query on specific target
- Sending data back to device based on request
- Handling data updates from the devices and updating databases and
- Compression and data checking of data packets

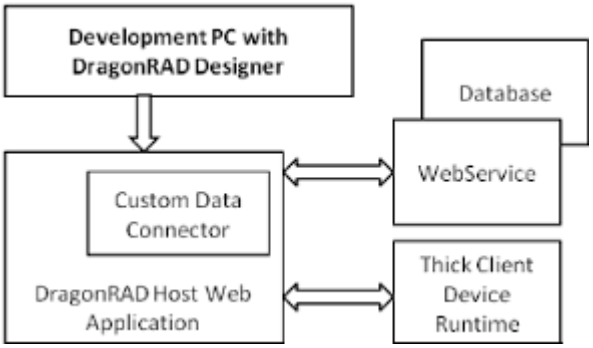


Fig. 2. DragonRad Components

C. DragonRad Client

This component behaves like a native application on device which helps to run and interpret code of the created application by designer. DragonRad has the emulator to run and debug the application. This component also has the feature to customize application like change icon, application name, DragonRad project and link for installation of DragonRad host. By changing the link to DragonRad host, the application automatically connects to the given host when it started. It would also help in updating the project when required. One advanced function provided by the DragonRad company is to compile the application online .

In the area containing your network provides all backend support such as a custom data connector that could fit for any web service available, while Tomcat/MapDataServer is to support for the database. This full network is connected to mobile phone with different OSs with the help of the Wi-Fi. The other sub-part BlackBerry Enterprise Server (BES) is specifically for BlackBerry products, such as PlayBook.This tool allows creating application in very easy manner with having only three step, which are connect, build and deploy. In DragonRad, it needs one connector that could be used to connect data source to DragonRad host web application. Therefore, it provides custom data connector to meet features. To handle the data transmission between back-end and device this connector is useful. AsDragonRad designer and Responsible for transferring data request and updates form device to database with keeping synchronization;

- Receiving data queries from device
- Processing received data queries
- Re-query data or resend stored data based on the query
- Receiving notification form device and
- Sending data to device

V. XAMARIN

Xamarin is a commercial application platform based on the Mono open source project, launched in 2001 by Ximian.Ximian was acquired by Novell in 2003 and then by Attachmate in 2011. After the Attachmate acquisition, part of the original team that worked on Mono formed Xamarin, which would take over the Mono project. Xamarin currently

supports iOS, Android and Windows Phone.Xamarin's main programming language is C# and with the release of Xamarin 3 also supports F#. Most of the code can be shared between the platforms, but the UI code is generally made separately for each platform using the standard native functions and following the platform's UI conventions. The UI can either be done programmatically or by using Xamarin's graphical UI designer for the target platform. The latest version, Xamarin 3, also introduced the Xamarin.Forms that allows part of the UI design to be done platform-independently. This way the data on the screen is separated from the code that renders it, and only the renderers vary by platform. Xamarin offers developers its own standalone IDE Xamarin Studio. Alternatively, Xamarin can be integrated into Microsoft's Visual Studio. Building the applications requires the native SDKs.

When considering how to build iOS and Android applications, many people think that the native languages, Objective-C, Swift, and Java, are the only choice. However, over the past few years, an entire new ecosystem of platforms for building mobile applications has emerged.Xamarin is unique in this space by offering a single language C#, class library, and runtime that works across all three mobile platforms of iOS, Android, and Windows Phone (Windows Phones native language is already C#), while still compiling native (non-interpreted) applications that are performant enough even for demanding games. Each of these platforms has a different feature set and each varies in its ability to write native applications that is, applications that compile down to native code and that interop fluently with the underlying Java subsystem. For example, some platforms only allow apps to be built in HTML and JavaScript, whereas some are very low-level and only allow C or CPP code. Some platforms dont even utilize the native control toolkit. Xamarin is unique in that it combines all of the power of the native platforms and adds a number of powerful features of its own, including:

1) Complete Binding for the underlying SDKs Xamarin contains bindings for nearly the entire underlying platform SDKs in both iOS and Android. Additionally, these bindings are strongly-typed, which means that theyre easy to navigate and use, and provide robust compile-time type checking and during development. This leads to fewer runtime errors and higher quality applications:

2) Objective-C, Java, C, and CPP Interop Xamarin provides facilities for directly invoking Objective-C, Java, C, and CPP libraries, giving you the power to use a wide array of 3rd party code that has already been created. This lets you take advantage of existing iOS and Android libraries written in Objective-C, Java or C or CPP. Additionally, Xamarin offers binding projects that allow you to easily bind native Objective-C and Java libraries using a declarative syntax.:

3) Modern Language Constructs Xamarin applications are written in C#, a modern language that includes significant improvements over Objective-C and Java such as Dynamic Language Features , Functional Constructs such as Lambdas , LINQ , Parallel Programming features, sophisticated Generics , and more:

4) Amazing Base Class Library (BCL) Xamarin applications use the .NET BCL, a massive collection of classes that have comprehensive and streamlined features such as powerful XML, Database, Serialization, IO, String, and Networking support, just to name a few. Additionally, existing C# code can be compiled for use in an applications, which provides access to thousands upon thousands of libraries that will let you do things that arent already covered in the BCL:

5) Modern Integrated Development Environment (IDE) Xamarin uses Xamarin Studio on Mac OS X and Visual Studio on Windows. These are both modern IDEs that include features such as code auto completion, a sophisticated Project and Solution management system, a comprehensive project template library, integrated source control, and many others:

6) Mobile Cross Platform Support Xamarin offers sophisticated cross-platform support for the three major mobile platforms of iOS, Android, and Windows Phone. Applications can be written to share up to 90% of their code, and our Xamarin.Mobile library offers a unified API to access common resources across all three platforms. This can significantly reduce both development costs and time to market for mobile developers that target the three most popular mobile platforms: Because of Xamarins powerful and comprehensive feature set, it fills a void for application developers that want to use a modern language and platform to develop cross-platform mobile applications.

A. How Does Xamarin Work?

Xamarin offers two commercial products: Xamarin.iOS and Xamarin.Android. Theyre both built on top of Mono, an open-source version of the .NET Framework based on the published .NET ECMA standards. Mono has been around almost as long as the .NET framework itself, and runs on nearly every imaginable platform including Linux, Unix, FreeBSD, and Mac OS X. On iOS, Xamarins Ahead-of-Time (AOT) Compiler compiles Xamarin.iOS applications directly to native ARM assembly code. On Android, Xamarins compiler compiles down to Intermediate Language (IL), which is then Just-in-Time (JIT) compiled to native assembly when the application launches.

In both cases, Xamarin applications utilize a runtime that automatically handel.

B. Application Output

When Xamarin applications are compiled, the result is an Application Package, either an .app file in iOS, or .apk file in Android. These files are indistinguishable from application packages built with the platform's default IDEs and are deployable in the exact same way Developing mobile applications is a challenge for every professional developer. They need to focus on each elemental aspect, while coding applications that are responsive, functional and compatible. Covering the bases for every feature is not exactly an easy task, even for the most talented developers. Thats why the Titanium application framework has emerged as one of the best for extensive mobile applications.

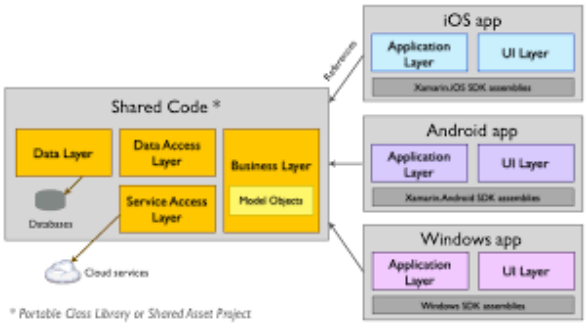


Fig. 3. Interfacing layers of the PhoneGap architecture

VI. Titanium

Titanium is an open-source application framework developed by Appcelerator and licensed under Apache Public License Version 2.0. It supports all the major mobile platforms, including Android, iOS, Windows Phone and BlackBerry.

Titanium offers a different approach to cross-platform applications than most hybrid application tools. Instead of HTML and CSS, applications are written completely in JavaScript using the Titanium API. The code is packaged with Titanium's engine, which interprets the code at runtime and renders the UI. This allows the applications to have the look typical to native applications, as the UI is made up of native elements through the API. Appcelerator offers a free IDE for developing applications with Titanium, called Titanium Studio. The IDE works on all the major operating systems: Windows, Mac OS and Linux. Building the applications requires installing the native SDK on the system, so building iOS applications requires a Mac OS device.

A. Features of Titanium Framework Development

The Titanium framework is loaded with exhilarating features, which is why it has become one of the most preferred frameworks among developers. Lets have a glimpse at the features of the Titanium mobile framework:

1) Enabled with JavaScript SDK: This open-source platform is embedded with JavaScript SDK, which aids developers in building functional-rich applications.:

2) Entails Trending Web Technology: It utilizes web technologies that are both trendy and powerful, including AJAX, HTML5, CSS3, and jQuery.:

3) Supports Multiple Platforms: Titanium has exclusive features that can be used to deploy applications without compatibility issues on various devices such as tablets, desktop and mobiles.:

4) Ensures Native Support: This framework comes with platform-independent API that ensures full access for native-supporting features.:

B. Advantages of Titanium App Development

1) Easy to Learn and Deploy: Titanium App Framework comprises HTML5, CSS3, jQuery, Ajax and JavaScript. Hence, it will be easy for expert developers to learn every nitty-gritty aspect of Titanium, and hence empower them in their design work.:

2) Easily Available: Titanium framework is free and open-sourced, so it is readily accessible to developers.:

3) Easy-to-Access Native Features: Titanium framework comes with a platformindependent API that can make applications feature-rich because it can access advanced features such as touchscreens, cameras, GPS, navigation, contacts, storage, and much more.:

4) Simplified Coding Structure: Because it supports HTML5 and other robust web technologies, developers can build apps that are compatible with iOS, Android and other powerful platforms. In short, there is no need to code one application multiple times for different platforms.:

5) Superb performance: The Titanium framework is packed with various functional entities that will definitely deliver a high-performance application.:

6) Deliver Quick Prototype: It comes with an integrated environment that will help developers build quick prototypes so they can get user feedback as quickly as possible. :

VII. Sencha Touch

Sencha Touch is a user interface JavaScript library, or web framework, specifically built for the Mobile Web. It can be used by Web developers to develop user interfaces for mobile web applications that look and feel like native applications on supported mobile devices. It is based on web standards such as HTML5, CSS3 and JavaScript. The goal of Sencha Touch is to facilitate quick and easy development of HTML5 based mobile apps which run on Android, iOS, Windows, Tizen and BlackBerry devices, Sencha Touch is a product of Sencha, which was formed after popular JavaScript library projects Ext JS, jQTouch and Raphal were combined. The first release of Sencha Touch, version 0.90 beta, was made available on July 17, 2010. This beta release supported devices running Android, and iOS. Subsequently, the first stable version, 1.0, was released in November 2010. Version 1.1.0 added support for devices runningBlackBerry OS 6.0. The Sencha portfolio of products and services forms an integrated, modular platform for designing, developing, and testing your cross-platform web applications. Sencha products can be deployed separately or together to form an end-to-end solution. With the Sencha Platform you can prototype, develop, theme, debug, and test your web applications on any device running any browser.

A. Design

The Sencha platform helps you accelerate your web application development efforts with out-of-the box theming capabilities across all applications. We can help teams:

1) Design visually compelling web applications using pre-built and pre-tested components:

2) Improve agility and the overall design process with tools and resources, including Ex t JS Stencils, that make it faster and easier to mock up, style, prototype, and evaluate interface designs.:

3) Enhance collaboration between line-of-business and IT teams to move apps quickly to the development phase.:

4) Create customized themes in a visual environment, using Sencha Themer, without writing a single line of code:

B. Develop

The Sencha platform offers powerful JavaScript and Java frameworks to help developers do their best work. We can help teams:

1) Build better applications faster with an enterprise-ready framework.:

2) Improve development efficiency by automatically handling critical issues such as rendering and cross-browser testing:

3) Build higher quality applications with tightly integrated and well-tested libraries and components.:

4) Develop once for multiple platforms and devices.:

5) Improve productivity with a rich set of web application tools including JetBrains, Visual Studio, and Eclipse IDE plugins.:

6) Code and debug web applications just once for multiple device types desktops, tablets, and smartphones.:

C. Test

Sencha Test provides the most comprehensive solution for testing Ext JS applications. We help developers and QA teams:

1) Write unit and end-to-end functional tests in JavaScript.:

2) Leverage the power of WebDriver in Sencha Test to create end-to-end tests.:

3) Run tests on any or all browsers on a local machine, connected mobile device, or on a browser farm.:

4) Maximize testing efficiency through automated test runs:

.

5) Review results from automated and manual test runs.:

6) Identify and rectify code coverage gaps.:

7) Perform visual screen comparisons across test runs.:

8) Create robust tests by leveraging the deep integration between Sencha Test and Ext JS.:

D. Developer Tools and Plugins

1) Sencha Touch provides IDE plugins for ease of development. Benefits of these plugins include code generation and auto-completion, code refactoring and ease of navigation to framework codebase and custom classes. Sencha Touch has plugins for popular IDEs like JetBrains, Visual Studio and Eclipse.:

2) There is a visual app builder, Sencha Architect , for building cross platform HTML5 apps. It provides addition features like theming and command line integration:

VIII. JQUERY

jQuery is a fast and concise JavaScript Library created by John Resig in 2006 with a nice motto Write less, do more. jQuery simplifies HTML document traversing, event handling, animating, and Ajax interactions for rapid web development. jQuery is a JavaScript toolkit designed to simplify various tasks by writing less code. Here is the list of important core features supported by jQuery

A. features

1) DOM manipulation The jQuery made it easy to select DOM elements, traverse them and modifying their content by using cross-browser open source selector engine called SizzleWhat jQuery doesThe jQuery library provides a general-purpose abstraction layer for common web scripting, and is therefore useful in almost every scripting situation. Its extensible nature means that we could never cover all the possible uses and functions in a single book, as plugins are constantly being developed to add new abilities. The core features, though, assist us in accomplishing the following tasks::

2) Access elements in a document: Without a JavaScript library, web developers often need to write many lines of code to traverse the Document Object Model(DOM) tree and locate specific portions of an HTML document's structure. With jQuery, developers have a robust and efficient selector mechanism at their disposal, making it easy to retrieve the exact piece of the document that needs to be inspected or manipulated.:

3) Modify the appearance of a web page: CSS offers a powerful method of influencing the way a document is rendered, but it falls short when not all web browsers support the same standards. With jQuery, developers can bridge this gap, relying on the same standards support across all browsers In addition, jQuery can change the classes or individual style properties applied to a portion of the document even after the page has been rendered :

4) Alter the content of a document: Not limited to mere cosmetic changes, jQuery can modify the content of a document itself with a few keystrokes. Text can be changed, images can be inserted or swapped, lists can be reordered, or the entire structure of the HTML can be rewritten and extendedall with a single easy-to-use Application Programming Interface (API).:

5) Respond to a user's interaction: Even the most elaborate and powerful behaviors are not useful if we can't control when they take place. The jQuery library offers an elegant way to intercept a wide variety of events, such as a user clicking on a link, without the need to clutter the HTML code itself with event handlers. At the same time, its event-handling API removes browser inconsistencies that often plague web developers.:

6) Animate changes being made to a document: To effectively implement such interactive behaviors, a designer must also provide visual feedback to the user. The jQuery library facilitates this by providing an array of effects such as fades and wipes, as well as a toolkit for crafting new ones.:

7) Retrieve information from a server without refreshing a page: This code pattern is known as Ajax, which originally stood for Asynchronous JavaScript and XML, but has since come to represent a much greater set of technologies for communicating between the client and the server. The jQuery

TABLE II
COMPARISON OF TOOLS BASED ON SOME FEATURES

Tool	**Language**	**Technology Approach**	**Mobile OS Support**	**Open Source**	**Result ing App**
PhoneGap	HTML,CSS andJavaScript	Web-to-native wrapper	Android,BlackBerry, iOS, WebOS,Windows Phone	Yes	Hybrid
DragonRad	WYSIWY Gand Lua	App Factory	Android, BlackBerry, iOS, Mobile Windows	No	Native
Titanium	HTML, CSS and JavaScript	Runtime	Android, iOS, Mobile Windows	Yes	Native
Xamarin	c#,Objective-C and Java	Runtime	Android, iOS, Mobile Windows	No	Native
Sencha Touch	, CSS3 and JavaScript.	Runtime	iOS, Android, BlackBerry, Windows Phone, and more.	Yes	Web based, Hybrid
jQuery Mobile	HTML5, CSS	Web-to-native	iOS, Android, BlackBerry, Windows Phone	Yes	Web based, Native

library removes the browser-specific complexity from this responsive, feature-rich process, allowing developers to focus on the server-end functionality.:

8) Simplify common JavaScript tasks: In addition to all of the documentspecific features of jQuery, the library provides enhancements to basic JavaScript constructs such as iteration and array manipulation.:

B. Why jQuery works well

1) Event handling The jQuery offers an elegant way to capture a wide variety of events, such as a user clicking on a link, without the need to clutter the HTML code itself with event handlers.:

2) AJAX Support The jQuery helps you a lot to develop a responsive and feature-rich site using AJAX technology.:

3) Animations The jQuery comes with plenty of built-in animation effects which you can use in your websites.:

4) Lightweight The jQuery is very lightweight library - about 19KB in size (Minified and gzipped).:

5) Latest Technology The jQuery supports CSS3 selectors and basic XPath syntax.:

IX. CONCLUSION

This paper provides the details about the cross-platform and six different cross-platform tools to develop applications on different mobile OSs. Currently, these tools are mainly used in companies with the aim to create applications designated to be sold on different market places, such as Apple Store, Play Store, etc. Cross-platform mobile development is a recent and unexplored area of software engineering. The possibility to develop for mobile platforms at once is a benefit appreciated by most of the mobile application developers. Lots of cross-platform tools are available online nowadays being the major challenge understand which one is the best to achieve the goals of a certain user or company. Moreover, cross-platform tools are still evolving and just like other software tools have flaws and limitations, but represent a straightforward solution to solve the platform fragmentation problem.The current state of the cross-platform mobile development tools market is dynamic, which means that the researches similar to this have to be conducted several times a year. The situation changes every day and there is no guaranty that a similar research conducted six months in the future will have the same results.

The future work for this research, other than updating the existing data and adding possible new tools, may be improving the collected data by creating prototypes using each tool. The researcher will be able to experience the strengths and shortcomings of each tool while testing them, and therefore provide the research with more accurate and reliable data.

REFERENCES

[1] Andr Ribeiro,Alberto Rodrigues da Silva,*"Survey on Cross-Platforms and Languages for Mobile Apps","Eighth International Conference on the Quality of Information and Communications Technology"*,2012.

[2] Manuel Palmieri,Inderjeet Sing,Antonio Cicchetti,*"Comparison of Cross-Platform Mobile Development tools"*, *"16th International Conference on Intelligence in Next Generation Networks"*,2012.

[3] Suyesh Amatya,*Cross-Platform Mobile Development:An Alternative to Native Mobile Development*,2013-10-29.

[4] Alireza Pazirandeh, Evelina Vorobyeva,*"Evaluation of Cross-Platform Tools for Mobile Development"*,"Chalmers University of Technology",June 2013.

[5] Allan Hammershj, Antonio Sapuppo, Reza Tadayoni,*"Challenges for Mobile Application Development"*,Center for Communication, Media and Information Technologies (CMI) Aalborg University.

[6] Nabil Litayem, Bhawna Dhupia, Sadia Rubab,*"Review of Cross-Platforms for Mobile Learning Application Development","IJACSA International Journal of Advanced Computer Science and Applications"*,2015

[7] Jesus Garcia and Anthony De Moss, Mitchell Simoens, *Sencha Touch in Action*, ISBN 978-1-61729-037-4

[8] Ajit Kumar, *jquery Cookbook - Second Edition*, ISBN 978-1-78216-918-5.

Cyber Security : Network Attacks and Countermeasures

Jency Jose, Nimisha C, Sreelakshmi K
Department of Computer Applications
Vidya Academy of Science and Technology
Thrissur-680501

Manesh D
Assistant Professor of Computer Applications
Vidya Academy of Science and Technology
Thrissur-680501

***Abstract*—Cyber security is the body of technologies, processes and practices designed to protect networks, computers, program and data from attack, damage or unauthorized access. This paper describes an overview of cyber security and it various fields such as cyber security for home user, wireless sensor network and also provides emerging threats of cybersecurity. It also gives emphasis on attacks and its countermeasures.**

***Index Terms*—Wireless Sensor Network,Symmetric and Asymetric cryptography,Home users,Emerging Threats,Malware,Holistic security**

I. Introduction

Cyber security and its underlying infrastructure are vulnerable to a wide range of risk from both physical and cyber threats and hazards. Active and passive attackers exploit vulnerabilities to steal information and are developing capabilities to destroy or threaten the delivery of essential services. A range of traditional crimes are now being committed through cyber space. This includes the production,banking and financial fraud.

The growth of the Internet has led to a significant growth of cyber attack incidents often with disastrous and grievous consequences. Malware is the primary choice of weapon to carry out malicious intents in the cyber space, either by exploitation in to existing vulnerabilities or utilization of unique characteristics of emerging technologies. The development of more innovative and effective malware defense mechanisms has been regarded as an urgent requirement in the cyber security community. This paper tells about an overview of the most exploited vulnerabilities in existing hardware, software, and network layers. Then discuss new attack patterns in emerging technologies such as social media, cloud computing, smart phone technology, and critical infrastructure. Finally, describe our speculative observations on future research directions.

II. Cyber security for Home users

Personal internet users are increasingly exposed to security threats while using their home PC systems. Such personal internet users in to two categories: Home Users (HUs) and Non Home Users (NHUs).

A. Non-home Users and Information Security Awareness

A lot of research has already been published on how to protect information properly within the NHU domain (academic, industry and government). This has led to the development and implementation of numerous information security awareness programmes within these domains. Within these domains, users are forced by their organizations to make themselves aware of information security and to apply a wide range of information security awareness tools. These include information security policies, procedures, guidelines and awareness courses. Two aspects of this NHU approach is clear information security awareness and enforcement. And to ensure safe practices when accessing cyberspace. This forces NHUs to gain access to the internet and web via a secured route. Relevant corporate policies, procedures, guidelines and best practices enforce this.

B. Home Users and Information Security Awareness

In the case of HU, the situation is totally different. Although research has been done on making home users aware of the importance of security their own information, the enforcement to do so does not exist. HUs there for in many cases venture onto the internet without any idea of what the risk are and what they must do to protect themselves. Growing numbers of HUs accessing the Internet for social networking, Internet banking and many other reasons, the big problem and worry is that in many cases. Such HUs are not information security aware, and are there for pottentially exposing themselves in a big way. The HU can get access without being exposed to the relevant information security awareness tools and support that are essential. These tools and support may be options to the HU, but in the most cases the HU does not make use of them because they are not enforced. So the amount of information security awareness programmes available for HU is far less than that for NHU. The main problem that if the HU does not know that he/she search for these awareness programmes online. From the investigation above two challenges were derived. The first is to create a framework for the design and implementation of information

of information security tools. This addresses the challenge to ensure that HUs obtain the relevant information security awareness to safely use the internet. The second challenge is to investigate ways in which HUs can be ways in which HUs can be forced to get exposed to such awareness tools to prepare themselves for the possible risks when obtaining access to the web. This challenge addresses the enforcement of information security awareness. These two challenges will be addressed by proposing a model. The proposed model pvovides one way of addressing these challenges. This paper specifically investigates the position of the home user, and proposes a new model, the E- Awareness Model (EAM), in which home users can be forced to acquaint themselves with the risks involved in venturing into cyber space. The EAM consists of two components: the awareness component housed in the E-Awareness Portal, and the enforcement component. This paper proposes an E-Awareness Model that can empower users by giving them a better understanding of security issues, possible threats and how to avoid them. The main difference between the presented model and other existing information security awareness models is that in the presented model the acquiring/ absorption of the awareness content is compulsory the user is forced to proceed via the E-Awareness Portal without the option of by passing it. The model as proposed is still very abstract, future research will concentrate on actually implementing the model in terms of a prototype, and then experimenting with the prototype to try to answer many open questions including those mentioned throughout this paper.

III. Cyber security challenges And Direction

This paper describes a data driven approach to studying the science of cyber security (SoS). It then describes issues and approaches towards the following three aspects: (i) Data Driven Science for Attack Detection and Mitigation, (ii) Foundations for Data Trustworthiness and Policy based Sharing, (iii)A Risk-based Approach to Security Metrics.

A. Data Driven Science for Detecting and Mitigating Attacks

Intrusion detection and prevention systems (IDPSs) perform signature based monitoring to identify malicious activities and generate alerts. Present system share two key limitations: they are unable to identify attacks whose signatures are not known and they are point based solutions geared to defend a single target. we need to create the foundations of the systems - that dynamically analyze heterogeneous streams of information. To extract facts that populate and maintain a semantically rich knowledge base (KB) with information about the resources being protected . These facts will be used to deduce context of the system, the possibility of attacks, and potential mitigations. The Cyber Kill Chain idea that tries to capture the offensive actions that an adversary is likely to take in attacking a system. Our research will move from this state of the practices to a context aware system. That largely automates this sophiscated analysis.

1) Context Representation and Sharing: Our context model requires information on components of the system, network and host data, attacks. Knowledge sharing is an effective mechanism to help agents to build contextual Knowledge, but requires that independently developed components and agents must share a common ontology and communication language. In most case contextual information has been considered as data structures or objects. But such representations lack expressiveness and extensibility. So encoding information in a Meta language like XML. So provide strong, community and stds based points of departure.

2) Detecting Attacks by reasoning: One approach to detect attacks is for the analyst to define rules based on the system context and incoming data. A combination of description logic and rule based logical inference is a feasible and scalable approach to allow this decoupling.

3) Detecting Attacks using Graph Grammars: We have developed efficient methods for learning graph grammars (both structure and parameter) from data. The grammar can be used to (1) recognize new attacks (2) partial parsing (3) expose compositional structure in the domain for human consumption as production rules extracted from data.

4) Detecting Attacks using stream based classification: Data streams have a dynamic nature, problem of concept-drift and concept evolution, the first concept occurs as results of a change in the underlying concept of data and second concept refer to the emergence of novel classes over time. Data stream is divided into equalized chunks.

5) Text Analysis to detect emerging Cyber Threats: The Web is often our first source of information about new software vulnerabilities exploits and cyber attacks.

6) Handling Inconsistency: We need to address the many problems involving representing using both data and knowledge that is uncertain. Inference rules may themselves be heuristics that are accurate most, but not all of the time.

B. Foundations for Trustworthy Data

Our objective is to explore the foundation of data trustworthiness as well as assured information sharing. So that accurate and trustworthy data is provided to carry out cyber security experiments. And we need to introduce a novel idea of the repeatability of cyber security experiments.

1) Trustworthiness of Data Based on Provenance: To correctly identify which data can or can not be trusted, we are investigating a novel method, based on the iterative filtering technique and provenance.

2) Formal Policy Analysis: We have developed multiple assured information sharing systems under the common MURI project. One such system shares data in the cloud. In this approach, the data and policies are represented in RDF and stored in the cloud. We developed a SPARQL query engine to query the data and a RDF policy engine to process the information sharing policies.

3) Reproducibility of Security Experiments: To tackle the problem of reproducibility of security experiments, we need to

to then extend research community. We need to Trustworthiness of Data Based on province, formal policy Analysis and Reproducibility of Security Experiments are the techniques used investigating problem. Analyze scientific literature in the area of security to determine which specific security topics involve experimental analysis and the types of experiment that are carried out. We need to also determine commonly used datasets by the security available reproducibility tools or develop new tools for security experiments and apply these tools to specific security attacks and defenses, such as return-on- programming attacks and the corresponding defense techniques (e.g. randomization) and intrusions and the corresponding intrusion detection techniques.

C. Security Metrics: A Risk Analysis Based Approach

One of the main challenges that arise in establishing a science for cyber security is to develop metrics to measure the security of the system. And any defensive mechanism employed by the defender may cause attackers to adapt and change their strategy. So further invoke mean field game theories to study our unified risk framework. That are:

1) Modeling Attackers and Their capabilities: Different types of attackers will have varying goals. So we plan to capture these varying goals in terms of different utility functions. Another important function is capture and representation of his/her capabilities. It hard to know the precise technical capabilities of an attacker in advance, so we plan to represent attackers capabilities as possible systems .

2) User Risk Modeling: The different type of users is malicious users and normal users, malicious users because the inside threat problem. So we need to model the malicious user and has as an attacker who has already gained control of the system. This will enable us to use our attack model with slight modifications.

3) Modeling Network Risk under Attacks: In order to estimate the risks associated with certain attacks, we need to understand how such attacks could affect the computer networks to be protected. Our model will be composed of three assessment models: a local risk assessment model, a relational assessment model, and a collective inference model.

4) Modeling Defensive Strategies: We need to model defensive strategies based on their cost, effectiveness and applicability to specific attack type. So we create profiles that can represent their properties. After this process we need to specially focus on important aspects related to networks. Due to the diverse properties of the different network and systems, it is important to understand the computational and power needs of the different defensive strategies.

5) Game Theoretic Models for Holistic Risk Assessment: All the previous components need to be combined in an extensive game. The attacker tries to optimize his attack by taking into account the defenders existing defensive strategies and network information. This game theoretical model will be able to answer questions by approximately analyzing the equilibrium behavior under different scenarios obtained by varying parameters in our models. So we believe that the three aspects addressed in this paper will form the basis for studying the Science of Cyber Security.

IV. Techniques In Cyber Security

A. Cryptography

The encryption decryption techniques devised for the traditional wired networks are not feasible to be applied directly for the wireless network and in particular for wireless sensor networks.WSNs consist of tiny sensors. Some questions arise when applying encryption schemes to WSNs like how the keys aregenerated or disseminated,How keys are managed or revoked. For secure transmission of various types of information over sensor networks, several cryptographic techniques are used: symmetric key ciphers and asymmetric key ciphers. The idea of the symmetric cryptography is to load secret information in the sensor nodes. This secret information may be the secret key itself or auxiliary information that helps the sensor nodes to derive the real secret key. By using this secret key, nodes can securely communicate . The main disadvantage of this solution is that compromising one node it lead to compromise the entire network. To overcome this limitation, researchers propose pairwise keys rather than a unique global key. Asymmetric key cryptography uses two separate keys , for encryption and decryption but those two keys are interconnected with complex mathematical algorithm. Since it is using complex mathematical algorithms it will induce huge overhead on power , computation and memory Asymmetric key cryptography is not preferred in WSN. Asymmetric key cryptography is more secured and efficient when compared to symmetric key cryptography.

B. Steganography

Steganography hiding the existence of the message. Itconverts messages in to multimedia data. If we want to send a secret data without sender information it is very useful.

V. Attacks In Cyber security

Attacks[1] against wireless sensor network could be broadly considered from different vies.one is attack against security mechanism and other is against basic mechanism.

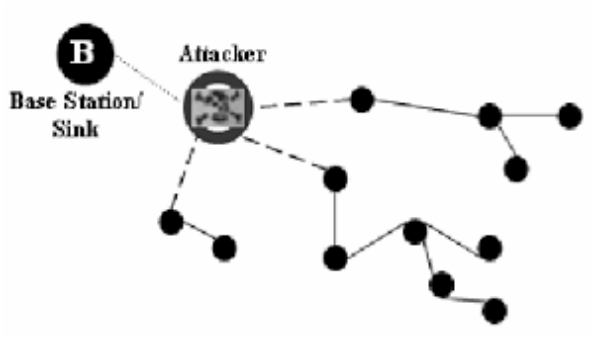
A. Denial of Service

It is produced by the failure of nodes or malicious action. In different layers several type attacks are performed. That are Jamming, Collision ,pushback, strong authentication etc. pushback, strong authentication are the mechanism to prevents DOS[2] attacks.

B. Sybil attack

In case of Sybil attack a node forges the identities of more than one node. It affect the identity of data security and resource utilization .Sybil attack can attack the distributed storage, routing mechanism, resource allocation.

C. Black hole/sink hole attack



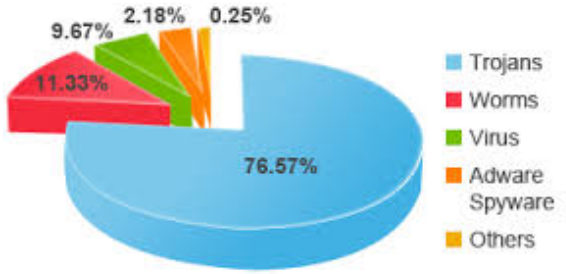
A malicious node act as a black hole which can attack all the traffic in the sensor network. Once the malicious device can able to insert between communicating nodes, it can able to do anything with the packet passing between them.

D. Hello flood attack

Hello packets are used as a weapon to convince the sensors in WSN. Here attacker sends Hello packets to number of nodes which are dispersed in WSN.

VI. EMERGING THREATS

A. Malware



Malware[3] is a general term for any type of unwanted software that does mischief or permanent damage to your computer. Malware authors use a number of different intermediaries to spread malware to infect a victims system. Traditionally, spam, phishing and web download have been the most commonly used mediums for the purpose.

1) Spam:: Spam is junk email that you did not ask for .To reduce spam, Delete junk email messages without opening them Dont reply to spam Dont give personal information in an email or instant messages. Dont forward chain email messages.

2) Phishing:: This is another way hackers try to get personal information from you. They use email to do this.They send a fake request acting like it is from a bank, or other institution asking for personal information.You click a link in the email and it takes you to a fake website imitating the institution.

3) Drive-by Downloads:: It concerns the unintended downloads of malware from the Internet and have been increasingly used by the attackers to spread malware fast. Drive-by downloads happen in a variety of situations; for example, when a user visits a website, while viewing an email message by user or when users click on a deceptive pop-up window.

B. Social media

Attackers are taking advantage of the social media craze as a new medium for launching insidious attacks. By the end of 2008, the Kaspersky Lab collection contained more than 43,000 malicious files relating to social media sites. Koop face worm that spreads through social media sites in 2009 is notably the best known malware case that utilizes the proliferation of social media sites.

C. Cloud computing

The efficiencies of moving data and applications to the cloud continue to attract consumers who store their data in Drop Box and Cloud, use Gmail and Live mail to handle email, and track their lives using services such as Ever note and Mint.com. Cloud computing is arguably one of the most significant technological shifts in recent times. Cloud computing provides unique characteristics that are different from the traditional approaches. The five key characteristics of cloud computing include on-demand self-service, ubiquitous network access, location-independent resource pooling, rapid elasticity, and measured service, all of which are geared towards using clouds seamlessly and transparently.

D. Smart phones

Smart phones, coupled with improvement in wireless technologies, have become an increasingly sophisticated computer and communication device that is readily carried by individuals throughout the day. The convergence of increasing computting power, personalization and mobility makes the man attractive means of planning and organizing work and private life of individuals. Mulliner listed the following features unique to mobile computing such as Mobility, Strong Personalization Strong Connectivity, Technology Convergence, Limited Resources and Reduced Capabilities. Another type of attacks is derived from the vulnerabilities in mobile software applications especially exploiting mobile web browser.

E. Critical infrastructure

The critical infrastructure systems that form the life line of a modern society and their reliable and secure operation are of paramount importance to national security and economic vitality. In most sense, the cyber system forms the back bone of a nations critical infrastructures, which means that a major security incident on cyber systems could have significant impacts on there liable and safe operations of the physical systems that rely on it. Critical infrastructure protection is harder to address than information and communication technology (ICT) protection because of these infrastructures interconnection complexity, which can lead to different kinds of problems.

F. Botnet

It is a network of compromised computers under the control of a remote attacker.

G. Social Engineering

It is the art of prying information out of someone else to obtain access or gain important details about a particular system through the use of deception.

H. Viruses

Malicious computer programs that are often sent as an email attachment or a download with the intent of infecting your computer, as well as computers of everyone in your contact list. eg: send spam.

I. Worm

A worm is computer program that moves itself from one machine to another often keeping record of the last environment it has entered.

VII. EXPLOITING EXISTING VULNERABILITIES

A. Hardware

A hardware is vulnerability is an exploitable weakness in a computer system that enable attack through remote or physical access to system hardware. Compared to software level attacks where many security patches, intrusion detection tools, and anti-virus scanners exist to detect malicious attacks periodically, many of the hardware-based attacks have the ability to escape such detection. Among different types of hardware misuse, hardware Trojan is the most hideous and common hardware exploits . The hardware Trojans are malicious and deliberately stealthy modification made to electronic devices such as Integrity Circuits(IC) in the hardware.

B. Software Defects

A software bug is the common term used to describe an error, flaw, mistake, or fault in a computer program such as internal OS, external I/O interface drivers, and applications. Cyber attacks utilize the software bugs in their benefits to cause the systems to behave unintended ways that are different from their original intent. The majority of cyber attacks today still occur as a result of exploiting software vulnerabilities caused by software bug and design flaws. Software-based exploitation occurs when certain features of software stack and interface is exploited. Most common software vulnerabilities happen as a result of exploiting software bugs in the memory, user input validation, race conditions and user access privileges . Memory safety violations are performed by attackers to modify the contents of a memory location.

C. Network Infrastructure and Protocol Vulnerabilities

The early network protocol was developed to support entirely different environment we have today in a much smaller scale and often doesnot work properly in many situations it is used today. Weaknesses in network protocols are complicated when both system administrators and users have limited knowledge of the networking infrastructure. One of the most common network attacks occurs by exploiting the limitations of the commonly used network protocols Internet Protocol(IP), Transmission Control Protocol(TCP) or Domain Name System(DNS).

VIII. FUTURE RESEARCH DIRECTION

A. Focus on Privacy

The goal of privacy aware security is to enable users and organizations to better express, protect, and control the confidentiality of their private information, even when they choose to (or require to) share it with others. Building techniques for data policy for data collection, data sharing and transmission, and dealing with privacy violations are other active are as of research in this category.

B. Next Generation Secure Internet

One of the main reasons for these security vulnerabilities is that the Internet architecture and its supporting protocols were primarily designed for a benign and trustworthy environment, with little or no consideration for security issues. A new paradigm of architectural design described as clean-slate design has been suggested. The theme of clean slate design is to design the system from scratch without being restrained by the existing system, providing a chance to have an unbiased look at the problem space.

C. Towards Trustworthy Systems

The term trustworthy systems have been defined by the Department of Homeland Security (DHS) in US as a long-term goal to indicate a computing system that is inherently secure, available, and reliable, despite environmental disruption, human user and operator errors, and attacks by hostile parties. Towards this goal, the author advocates the requirement for secure hardware and software combinations as essential building block towards trustworthy system.

D. Global-scale Identity Management and Traceback Techniques

Global-scale identity management concerns identifying and authenticating entities suchas people, hardware devices, distributed sensors and software applications when accessing critical information technology systems from anywhere. The term global-scale is intended to emphasize the pervasive nature of identities, due to increasing use of mobile phones and embedded sensors in every where of our daily life. Provenance technique is another notable one that has been emerging and provides an ability to trace the life time changes and transformation of computer related resources such as hardware, software, documents, database, data, and other entities.

E. Usable Security

Many security technologies have tried to improve the usability aspects; most of which fall short in terms of usability. Password schemes have been believed to be one important parts of usable security. Therefore, several elaborate procedures have been progressed such as frequency of changing, inclusion of non alphabetic characters, or visual and biometric based passwords that users do not have to remember. Despite these attempts, security pitfalls of poorly implemented password schemes have been extensively documented over the years. Users resort to writing them on slips of paper or storing them unencrypted on hand held devices.

IX. Conclusion

This paper focused on two aspects of information system: understanding vulnerabilities in exiting technologies, emerging threats and attacks in cyber security. Growing threats have been found in emerging technologies, such as social media, cloud computing, smartphone technology and critical infrastructure. We described cyber security for home users through awareness enforcement. Then, we discuss common set of general attack patterns found in the cyber secrity. We also illustrated future research directions. As more and more people are connected over the Internet, understanding all levels of users including both experts and non-experts in computing system.

References

[1] Choong Seon Hong , Hyung Woo Lee ,Al-Sakib Khan Pathan, Security in Wireless Sensor Network,2015
[2] Kahina CHELLY, Wireless Sensor Network issues and countermeasures, 2015
[3] S.H.von Solms , Survey on Cybersecurity, 2014
[4] Surya Nepal, Julian Jang-Jaccard, Survey of emerging threads in Cyber security, 2014
[5] Alhaji idi Babate, State security of cyber : Security in Home Users,2012
[6] E.Kritzinger Cybersecurity, Challenges and Direction, 2015
[7] J Liu, Y Xiao, S Li, W Liang, Surveys and Tutorials, 2012 - ieeexplore.ieee.org
[8] Yan, Y Qian, H Sharif, D Tipper Surveys and tutorials, 2012
[9] DG Padmavathi, M Shanmugapriya - arXiv preprint arXiv:0909.0576, 2009 arxiv.org
[10] C Karlof, D Wagner - Ad hoc networks, 2003 Elsevier
[11] P N Raj, P B Swadas - arXiv preprint arXiv 0909.2371, 2009 arxiv.org blackhole
[12] J R Douceur - International Workshop on Peer-to-Peer Systems, 2002 Springe sybil attack
[13] J Yick, B Mukherjee, D Ghosal - Computer networks, 2008 - Elsevier Wireless sensor network survey
[14] C F Huang, Y C Tseng - Mobile Networks and Applications, 2005 - dl.acm.org The coverage problem in a wireless sensor network

A Study of Cyber Forensic in the Context of Digital Evidence and Emerging Forensics

Juhy Prabha M P, Sruthy N T
Department of Computer Applications
Vidya Acadamy of Science and Technology
Thrissur - 680501

Reji C Joy
Associate Professor of Computer Applications
Vidya Acadamy of Science and Technology
Thrissur - 680501

***Abstract*—Nowadays cyber crimes have been increased rapidly. This also leads to improve the relevence of cyber forensic department. Cyber forensic focuses through the digital evidence and various emerging forensics are used of protecting our files we can add watermark on it . Software can be coyright protected. Owner has every moral right to ensure the property right to the software. The duty of forensic professionals to explain the forensic role of various different watemarks separately and then genaralize these differents roles to form a single forensic philosophy which can be ultimately used by the judiciary for the effective decision making in any software copyright infringement litigation. Cyber Laws are described in each and every crimes related to computer or the communication devices .**

***Index Terms*—Cyber Crime, Digital Evidence, Cyber Laws, Watremark, Emerging Forensics**

I. Introduction

This paper briefly explain about cyber forensic, cyber crime, evidence, different laws, importance of watermark and new developed technologies. Computer forensic is a branch of forensic science pertaining to evidence found in computer and digital storage media. cyber crime is a fast growing area of communication technology. Define cyber crime as any criminal activity that dealing with computer and network, examples are hacking, virus, denial of the services of network etc. The involvement of computer in crime has resulted in a abundance of digital evidence that can be used to applied and prosecute offenders. Digital evidence can be useful in a wide range of criminal investigation including homcidies, civil cases etc. Cyber law encompasses relating to Cyber Crime. Two basic levels of cyber crime laws are federal cyber crime law and state cyber crime law. Watermark is widely used mechanism to protect the ownership of digital file and it provide direct evidence to establish copyright infringement more than other programming blunders. It also help to make dicision in judiciary world. There are many technologies are used to help forensic they are handeld devices, mac forensic, linux and unix forensic, live forensic, remote acquisition, digital conservation and digital archaeology

II. Cyber Crimes

The cyber crimes are mainly computer targeted. Cyber financial fraud, identity theft, hacking, virus, threat through email, vulgar messages, use of computer for anti -national activities, use of computers for personal gains, violation of company acceptable policies, launching of denial of service attacks on computer network servers, software piracy also criminal activities like misuse of telephone technology, pornography, unauthorized disclosure of internal and confidential information, theft or trade of intellectual property etc, these all are fall under cyber crime . In this scenario cyber crimes create investigative and prosecution related issues. The issues are solved based on electronic evidence. The categorization Computer in cyber crime:

1) Computer can be the subject of a crime: These are used to being stolen or damaged.
2) The site of the crime: These are important, for example, when a child is solicited for sex in a chat room.
3) The instrument of the crime: It is used to store information illegally.

III. Classification of Cyber Crimes

There are different types of cyber crimes. A few of the important ones are listed below:

1) **Fraud and Financial Crimes**
   This is behave misrepresentation of fact intentended to let another to do or refrain from the doing something which causes loss. These crimes are happen on unauthorized way, the storing data are deleted, destroying the data.
2) **Cyberterrorism**
   In generally say that the terrorism committed through the use of computer resources.
3) **Computer virus**
   Computer virus is type of intented to harmful software program.These can make system failure, wasting computer source, destroying data, often perfoming access private information etc.
4) **Denial of service attacks**

These attacks are characterized by an direct attempt by attackers to prevent authorized users of a service from that using the service. These can typically used for the network became slowing process.

5) **Malware**
Malware is any software is used to destroy computer or mobile operations. The malicious software was reffered to as computer viruses. The malware sometime known as computer contaminant.
6) **Intellectual property threat**
These includes following:
   a) Copyright violation : Without permission the files or data canbe copied thorugh the computer or mobile-phones etc.
   b) Cyber squatting : Registering a domain that is the trademark of another person or company
   c) Name changing : Misspelled variation of the domain name.
   d) Name stealing : Changing the ownership in the registrations database.
7) **Software piracy**
Illegal copying , distribution or use of software.
8) **Pronography**
This is based of one of the cybercrime because the childrens are use solicited sex sites.
9) **Threats through Email**
The malicious software or virus are spread through the email.
10) **Hacking**
Hacking is unauthorized intrusion in to computer or network.

In the cybercrimes the investigation department focuses mainly the digital evidence.

IV. Digital Evidence

Digital forensics is a branch of forensic science the recovery and investigation of material found in digital devices often in relation to computer crime. The technical aspects of a investigation divided in to several sub branches relating type of digital devices involved computer foerensics, network forensic, data analysis and mobile device forensics.

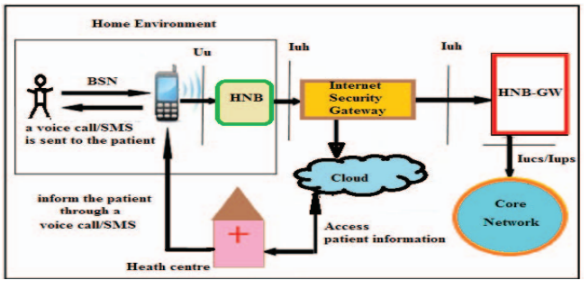


Fig. 1. Digital evidence storage devices

The involvement of computer crime has resulted in a abundance of digital evidence that can be used to apprehend and prosecute offender. Digital evidence can be useful in a wide range of criminal investigation including homcidies, sex offences, missing person , child abuse, drug dealing , fraud and theft of personal information also civil cases can hinge on digital evidence or electronic discovery is becoming a routine part of civil disputes. Digital evidence can reveal how a crime was committed, provide investigative leads, disprove or support wictness statemnets and identify likely suspects.

Fig. 2. Digital evidence storage devices

Digital evidence is defined as any data stored, the data reffered to in this definition are essentially combination of numbers that represent information of various kinds , including text, images, audio and video. When considering digital evidence it is useful to categorize computer system in to three groups

1) **Open Computer System**
It consist of hard drive key etc. these systems with their ever increasing amount of storage space can be rich source of digital evidence.
2) **Communication Systems**
Traditional telephone system,wireless communication system . These digital investigation access to communications (message, text, attachments, telephone conversations).
3) **Embedded Computer Systems**
Mobile devices, smart cards and many other systems with embedded computers may contain digital evidence.

Attorneys and police are encountering progressively more digital evidence in their work. In the following aspects digital evidence processed using various steps i.e.,

1) **Assessment**
The examiner asses the digital evidence throught the case. These are the scope of the case.
2) **Acquisition**
The digital evidence by very nature, the evidence may be altered, destroyed by improper handling or examination so the original evidence should be aquire in proper manner and protect and preserve.

3) **Examination**
   The main process of the examination process is to extract and analyse the digital evidence.
4) **Documenting and Reporting**
   The observation and action related to the forensic evidence.

The digital evidence said to be forensically sound, it was collected, analysed, handled and stored in a manner that is acceptable by the law authentication means satisfying the court the contents of the record have remained unchanged.

V. Cyber Law

(1) **Federal Cybercrime Law**: Federal cybercrime law focuses on the computer fraud and abuse act as well as the identity theft, child pornography and copyright and trademark offenses.
(2) **State Cybercrime Law**: These includes the acess-crimes(simple hacking,aggregated hacking), malware, deniel of services.

A. Information Technology Act 2008

Information technology act 2008 has been notified and enforced on 27th October 2009. This act punishes various cybercrimes including cyber terrorism. The following are some of the important sections related cybercrimes.

1) **Sec.66**
   Computer related offences: If any person fraudently or illegal does any act refferd to section 43, he/she shall be punnishable with arrested and extended to three years or with fine which may be exteded to five lakh rupees or both.
2) **66 A**
   A punishment for sending offensive message through communication service: It include electronic mail and electronic mail messages,computer resources or recieved on a communication device including attachments in text, image, audio, video, and any other electronic reffered punnishable with imprisonment for three year with fine one lakh rupees.
3) **66 B**
   Punishment for dishonestly receiving stolen: Computer resource or communication device. These also punishment imprisonment with fine one lakh or both.
4) **66 C**
   Punishment of identity theft: Here we use the electronic signature or password or any other unique identification feature of any other person. Punished with three year imprissonment and fine may be extended one lakh rupees.
5) **66 D**
   Punishment for cheating by personation by using computer resource: These means any communication device or computer resource cheats by personation. Punished with three year imprisonment may extended the three year and fine extended one lakh rupees.
6) **66 E**
   Punishment for violation privacy: These related transmit, published, the image of private area of any person without his/her consonent, under circumstances violating the privacy of that person. Punished the person imprisonment upto three year and fine will be two lakh rupees or both.
7) **66 F**
   Punishment for cyber terrorism : Access computer resouces without auothorization or exceeding authorised access and by means of such conduct obtain access information or any restricted information, data or computer database, the security of state group of individual commit offense cyber terrorism. Punishable with imprisonment which may extended to imprisonment for life.
8) **67 A**
   Punishment for publishing and transmitting of material containing sexually explicit act etc. In electronic form.
9) **67 B**
   Punishment for publishing and transmitting of material depict in children in sexually explicit acts in the electronic form- These include facilitate abusing children online. They punished imprisonment of either descrption for a term which may extended with fine ten lakh rupees.
10) **67 C**
    Preservation and retension of information by intermediates punished with a imprissonment for a term which may extended to three years and with fine.
11) **Sec.69**
    Powers to issue direction for interception or monitoring or decription of any information through any computer resource. Punished with an imprisonment for seven years with fine.
12) **69 B**
    Power to authorised to monitor and collect traffic data and informations.
13) **Sec.70**
    Protected system

B. Offenses Covered Under IPC and Special Laws

1) **Sec.503 IPC** – Sending threatning messages by email
2) **Sec.499 IPC** – Sending defamatory messages by email
3) **Sec.463 IPC** – Forgery
4) **Sec.464 IPC** – False document
5) **Sec.468** – Forgery purchase of cheating
6) **Sec.383** – Webjacking
7) **Sec.500** – Email abuse
8) **Sec.51** – Copyright infringed
9) **Sec.379** – Punishment for theft
10) **Sec.378** – Theft on computer hardware[3]

VI. Digital Water Mark in the Context of Copyright Law

Watermark is the process of embedding some kind of information into the source code of the file. It used to protect the ownership of a digital file that include audio, video, image etc. Software can be protected for this protection we can add watermark on our software. When copyright have been

violated some time its for commercial profit. The injuired party or copyright owner has a moral and legal right to ensure the protection of their property to the software. In this situation attach a watermark with the file will helps to the forensic department. It make a evidence to establish criminal activity behind the infringement allegation[1].

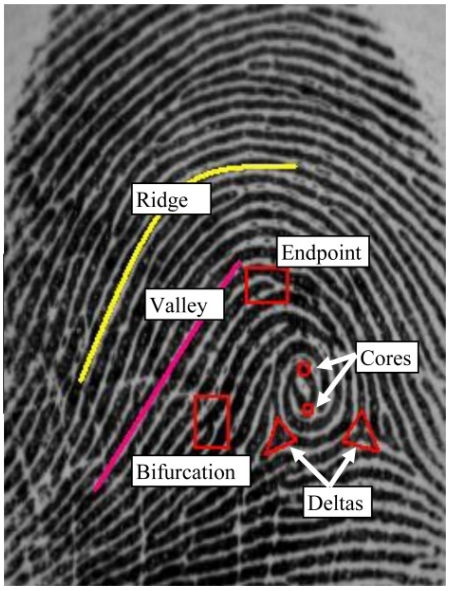


Fig. 3. Sample Watermark

Watermark can be classified in to two, they are static and dynamic watermarking. Static watermarks are embedded as code segments with in the source code of the digital file. Dynamic watermarks are genertated during the runtime with the help of code segments embedded with in the souce code of the digital file. Every watermark has certain features they are Effectiveness or correctness and aptness of the intended purpose of watermark. Integrity it is the ability to not to interfere with the performance of the source code. Fedelity(how closely the watermark accurately or truthfully helps to identify the owner of the software) . Robustness is the ability to withstand any kind of alteration of the contenct of the file in which the the watermark is embedded. Watermark can be directly linked to copyright. Watermark need to be considered as sepperate program segment. The judiciary system world wide also need to be encouraged to the study and legalise evidencial aspect of digital watermark in the context of copyright law. The software forensic research community incorporate water mark in the AFC test(Abstracton Filtration Comparison) in [1]. Thus, most watermarks can be provide direct evidence to establish copyright infringement more than other programming blunders and play a major role in forensic community this helps to make decision in judiciary world

VII. Emerging Forensics

A. Handheld Devices

Most forensic has been inclined towards the desktop, laptop and associated media like hard drive floppy disk , optical disk access and anlaysed through file system by using opertating system . Two of the leaders in mobile forensics dedicated to this field are CelleBrite with its UFED physical analyser and micro systemation which produces XRY complete . with in the broader forensic field, Parabens Device Seizure has established a reoutation in hand held forensics

B. Mac Forensic

Mac forensic is the another emerging forensic. If tools such encase which are meant to use with windows are capable of capturing information and to interpret but face a lot of limitations. Then Apple introduced Macintoish environment . A distict functionality that can be used to acquire Apple Mac system is the Target Disk Mode. However, care is needed, because there are times when it is unavailable due to a firmware password , resulting in the operating system being engaged, and consequent write protection failure

C. Live Forensics

There is increasing interest in the acquisition of live date, of the physical memory. In part this is in order to bypass encryption. While it may not be an immediate priority for the digital curators, it may become useful in scenarios such as cloud computing. It is commonly used in seeking to understand and target malicious processes in windows.

D. Remote Acquisition

If the activities were conducted using internet the scientists may save their time. Given the support and knowledge of the donor, standard means might include secure email(akin to services such as Voltage), uploading by the donor through a secure fprm of FTP or a use of a remote access service such as LogMein Forensic technologies offer several key advantages :

(i) Forensically sound inspection and acquisition
(ii) Security and control of access
(iii) Detailed auditing and logging of the activities of the examiner

For example , the Field Intelligence Model(FIM) of encase illustrates the use of a forensically sound and accountable authentication administration server,linked via secure connections over a network using a public key AES encryption system explained in [2].

E. Digital Forensic,Digital Archaelogy

Digital forensic does not encompass the investigation of damaged media and objects. But it is necessary to check them. Digital conservation for those situations where the storage media and other hardware are significantly degraded or damaged , or where the media and handware are being investigated at fundamental levels with a view to enhancing or expanding the recovery and preservation of the digital information [2]. Archaeology (digital media or otherwise) for the situation

where fragments of the information are not only recovered but used to investigate and interepret social circumstances, such as online communities or early computer game player, using the phrase digital conservation for the recovery and care of the information itself.

VIII. Conclusion

This paper elaborates the new emerging forensics and digital evidence. These two terms are more helpful for the forensics department . Another section describes watermark, we can add watermark to protect our files and software, also explain diffrent types of watermark techniques and their security relevance. then the next section about various laws related to the cyber crimes . Cybercrime laws are not aware of the common peoples so, here the importance of laws are exibited. The cyber crime investigation digital evidence are important evidences, these electronic evidence are more supported elements of the forensics department. Through the analysis of digital evidence forensic expert can detect behaviour of the crime and capture information using the appropriate forensic tools.

References

[1] Vinod.P.Battathiripad,Snehasudhakaran,Roshna.k.thalayaniyil,*"Conference on Digital Forensics"*,Security and law ,2015
[2] Jeremy Leighton john, digital forensic and preservation(book)
[3] N.NR.Sharavanan, N.Balu, "Cybercrime and Related Laws", *International Journal of Emerging Technology And Advanced Engineering*, 11/11/2012
[4] kk.Sindhu, BB.Mesharam, "Digital Forensics and Cyber Crime Datamining", *Journal of Information Security*, July 2012
[5] Dr. S Santhosh baboo & S. Mani Megalai, "Cyber Forensic Investigation and Exploration on Cloud Computing Enviornment", *Global Journal of Computer Science And Technology : B Cloud And Distribute*,2015
[6] Battathiripad,P.V[2014],Judiciary friendly forensics of software infringgement,IGI-Global
[7] Collberg,C.& Thomborson,C[1999],Software watermarking: models and dynamic embeddings,proceedings of the 26th ACM SIGPLAN-SIGACT sympossium on principales of programming langugages()pp.311-324)ACM
[8] Eoghan Casey, *Digital Evidence And Computer Crime Forensic Science Computers And The Internet*, Elsever Inc., 12/4/2014

A Survey of Internet of Things Based on Security and Privacy

Nidheesh S, Nimex Nedumparambil, Rohit Raveendran
Department of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501

Manesh D
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501

***Abstract*—Internet of Things is a new era of computer technology. IoT (Internet of Things) refers to the standard internet protocol for the human-to-human, human-to-thing communication. IoT (Internet of Things) is the network of physical objects-devices, vehicles, buildings and other items embedded with electronics, software, sensors, and network connectivity-that enables these objects to collect and exchange data. This paper focus on the major security and privacy concerns and smart home where an IoT applications are implemented.**

***Index Terms*—IoT Security, IoT Privacy, IoT Applications**

I. Introduction

The Internet of things (IoT) is the inter-networking of physical devices, vehicles (also referred to as "connected devices" and "smart devices"), buildings, and other itemsembedded with electronics, software, sensors, actuators, and network connectivity that enable these objects to collect and exchange data. IoT allows objects to be sensed and controlled remotely across existing network, creating opportunities in integration of the physical world into computer-based systems. The Internet of Things (IoT) denotes the interconnection of highly heterogeneous networked entities and networks following a number of communication patterns such as: human-to-human (H2H), human-to-thing (H2T), thing-to-thing (T2T), or thing-to-things (T2Ts). Aside from these security issues, the average consumer is concerned about his or her privacy. The Internet of Things (IoT) has particular security and privacy problems. The Internet Engineering Task Force is designing authentication and authorization mechanisms for the most constrained devices which are part of the Internet of Things. In many of the applications of Internet of Things (IoT), sensors measure variables such as speed, pressure, consumption, temperature or heart rate, and actuators control physical systems, such as brakes, valves, lights, power circuits, or automated drug dispensers. What makes these scenarios interesting from a security and privacy perspective, is that they all affect the physical world, sometimes controlling critical infrastructure, and sometimes gathering very private information about individuals. This paper studies the main security and privacy issues that comes in IoT (Internet of Things). And also look into one of the application of IoT, that is smart home and its security challenges

II. Security and Privacy needs

Internet of Things (IoT) is facing many security and privacy problems in the current technology. Since it is using wireless network it has many security concerns in connection mechanism.

A. Requirements related to IoT technology

Privacy includes the concealment of personal information as well as the ability to control what happens with this information. The attribution of tags to objects may not be known to users, and there may not be an acoustic or visual signal to draw the attention of the objects user. Thereby, individuals can be followed without them even knowing about it and would leave their data or at least traces thereof in cyberspace [1]. Since business processes are concerned, a high degree of reliability is needed. In the literature, the following security and privacy requirements are described:

- Resilience to attacks: System has to avoid single points of failure and should adjust itself to node failures.
- Data authentication: As a principle, retrieved address and object information must be authenticated.
- Access control: Information providers must be able to implement access control on the data provided.
- Client privacy: Measures need to be taken that only the information provider is able to infer from observing the use of the lookup system related to a specific customer; at least, inference should be very hard to conduct.

B. Privacy enhancing technologies (PET)

The fulfillment of customer privacy requirements is quite difficult. A number of technologies have been developed in order to achieve information privacy goals. These Privacy Enhancing Technologies (PET) can be described in short as follows:

- Virtual Private Networks (VPN) are extranets established by close groups of business partners. As only partners have access, they promise to be confidential and have integrity. However, this solution does not allow for a dynamic global information exchange and is impractical with regard to third parties beyond the borders of the extranet.
- Transport Layer Security (TLS), based on an appropriate global trust structure, could also improve confidentiality and integrity of the IoT. However, as each ONS delegation step requires a new TLS connection, the search of information would be negatively affected by many additional layers.
- DNS Security Extensions (DNSSEC) make use of public-key cryptography to sign resource records in order to guarantee origin authenticity and integrity of delivered information. However, DNSSEC could only assure global ONS information authenticity if the entire Internet community adopts it.
- Onion Routing encrypts and mixes Internet traffic from many different sources, i.e. data is wrapped into multiple encryption layers, using the public keys of the onion routers on the transmission path. This process would impede matching a particular Internet Protocol packet to a particular source. However, onion routing increases waiting times and thereby results in performance issues
- Private Information Retrieval (PIR) systems conceal which customer is interested in which information, once the EPCIS have been located. However, problems of scalability and key management, as well as performance issues would arise in a globally accessible system such as the ONS, which makes this method impractical.

A further method to increase security and privacy are Peerto- Peer (P2P) systems, which generally show good scalability and performance in the applications

III. Security and Privacy concerns in IoT

A. Security Concerns in IoTs

Internet of Things virtually is a network of real world systems with real-time interactions. The development of the initial stage of IoT, is M2M (Machine to Machine), having unique characteristics, deployment contexts and subscription. Unattended operation without human intervention is possible for long periods of time by the wireless area network (WAN) or WLAN. Though providing improvements in social efficiency it creates an array of new problems concerning breach of privacy and that information security.

1) Front-end Sensors and Equipment:
   Front-end sensors and equipment receives data via the built-in sensors. They then transmit the data using modules or M2M device, thus achieving networking services of multiple sensors. This methodology involves the security of machines with business implementation and node connectivity . Machine or perception nodes are mostly distributed in the absence of monitoring scenarios. An intruder can easily access these devices which imply damage or illegal actions on these nodes can be done. Possible threats are analyzed and are categorized to unauthorized access to data, threats to the Internet and denial of service attack.
2) Network
   Network plays an important role providing a more comprehensive interconnection capability, effectualness and thriftiness of connection, as well as authentic quality of service in IoTs. Since a large number of machines sending data to network congestion, large number of nodes and groups exist in lOTs may be resulted in denial of service attacks.
3) Back-end of it systems
   Back-end IT systems form the gateway, middleware, which has high security requirements, and gathering, examining sensor data in real time or pseudo real-time to increase business intelligence. The security of IoT system has seven major standards viz; privacy protection, access control, user authentication, communication layer security, data integrity, data confidentiality and availability at any time.

B. Privacy Concerns in IoTs

Privacy should be protected in the device, in storage during communication and at processing which helps to disclose the sensitive information .The privacy of users and their data protection have been identified as one of the important challenges which need to be addressed in the IoTs [2].

1) Privacy in Device
   The sensitive information may be leaked out in case of unauthorized manipulation or handling of hardware and software in these devices. For example, an intruder can re-programme a surveillance camera could such that it sends data not only to the legitimate server, but also to the intruder. Thus, for devices that gather sensitive data robustness and tamper-resistance are especially important. To ensure IoTs security trusted computing technologies including device integrity validations, tamper-resistant modules and trusted execution environments are useful.
2) Privacy during Communication
   To assure data confidentiality during the transmission of the data, the most common approach is encryption. Encryption on certain occasions adds data to packets which provides a way for tracing, e.g. sequence number, IPsec- SecurityParameterIndex, etc. These data may be victimized for linking packets to the analysis of same flow traffic. Secure Communication Protocol could be the suitable approach. During the communication

Pseudonyms can be replaced for encryption in case it is not feasible to the devices identity or users in order to decrease the vulnerability. One of the long-familiar examples is Temporary Mobile Subscriber Identity (TMSI). Devices should communicate if and only if when there is a need, to derogate privacy disclosure induced by communication. In 3GPP machine type communications, in order to avoid unnecessary collection of location information by the network after a certain period of inactivity the devices will detach from the network.

3) Privacy in Storage
For protecting privacy of information storage, following principals should be considered.
   - Only the least possible amount of information should be stored that is needed.
   - In case of mandatory then only personal information retained.
   - Information is brought out on the basis of need-to-know.

4) Privacy at Processing
It is mainly of two folds. Firstly, personal data must be treated in a way that it should be simpatico with the intended purpose. Secondly, without explicit acceptance and the knowledge of the data owner, their personal data should not be disclosed or retained to third parties.

IV. Applications of IoTs

A survey done by the IoT-I project in 2010 identified IoTs application scenarios which are grouped in 14 domains viz; Transportation, Smart Home, Smart City, Lifestyle, Retail, Agriculture, Smart Factory, Supply chain, Emergency, Health care, 3User interaction, Culture and tourism, Environment and Energy [10].

A. IoTs in Medical Application

Due to population growth, rural urbanization, declining birthrate, population aging, economic growth and social unbalanced resource utilization, some social problems have become increasingly apparent in the healthcare field.

- The health management level and the incapability of responding to emergency is a pressing social problem.
- There is a serious shortage in medical staffs, institutional facilities especially in rural areas, lack of medical facilities, low level of treatment, inadequate healthcare system
- The imperfect diseases prevention system cannot meet the national strategy requirements to safeguard the health of the citizen becoming heavy burden on economy, individuals, families and state.
- Inadequate disease prevention and early detection capability.

B. IoT in Smart Home

Now a days, smart homes are becoming more and more costeffective and intellectualized with continued progress and cost reduction in communication technology, information technology, and electronics, which connects the Internet with everyday devices and sensors for connecting virtual and physical objects through the data capture and communication development. Reading of remote meters can be attained through these smart home systems. That implies, the data related with home power, telecommunications, gas and water can be sent automatically to their corresponding utility company to enhance the efficiency of the work. In addition, by virtue of smart home systems, windows, home ventilation, doors, lighting, air conditioning etc., can be controlled by remotely

C. Intelligent community security system (ICSS)

The intelligent community security system (ICSS) holds several subsystems, such as Vehicle Management Subsystem (VMS), Surrounding Security Subsystem (SSS), Central Information Processing System (CIPS), Property Management Subsystem (PMS), Fire and Theft

1) Vehicle Management Subsystem of the ICSS: The Vehicle Management Subsystem in ICSS adopts IPR, sensor network technologies and RFID. Image registration can be taken by RFID card and video camera which is given to the vehicles, as shown in Figure. The vehicle license information will be messaged to the CIPS, when it enters the communities. The visitors are allocated with the temporary parking places. The record data and the information of the driver RFID card must be coherent, when the car leaves. This guarantees the security of cars and prevents theft occurrences. In the garages video monitoring devices will prevent stealing or damage to assure the vehicles safety. Through the Human-Computer interface system CIPS can controls the garages to facilitate and observe the vehicle management.

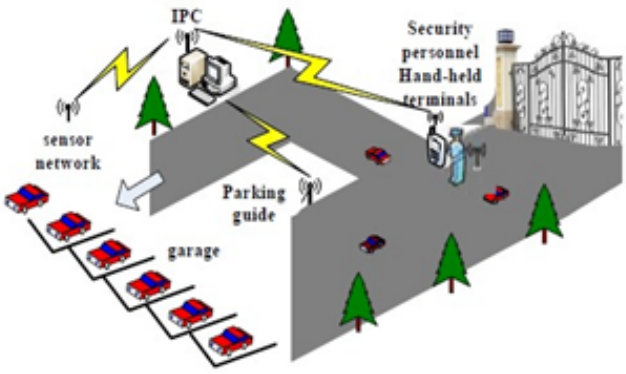


Fig. 1. Vehicle Management Subsystem

2) Property Management Subsystem of the ICSS: humanized and efficient property management system provides more convenience and happiness to the residents. As shown in the IoT technology can get better residential property management which is more standardized and scientific. Public Facilities Monitoring System use the unified coding sensor network technology which provides real-time monitoring of the public facilities such as the public

transportations, swimming pools, emergency exits, residential elevators, community basketball courts and so on.

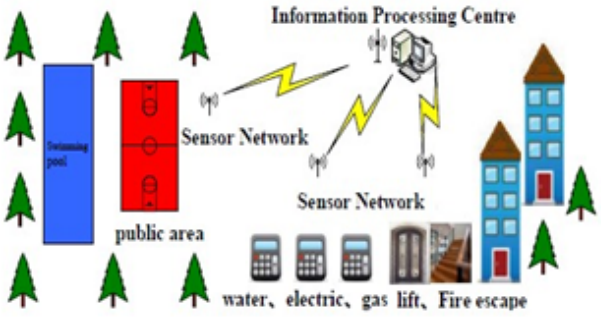


Fig. 2. Property Management Subsystem of the ICSS

3) Fire and Theft Prevention Subsystem (FTPS) of the ICSS: Electrical equipments and appliances may induce huge potential dangers. The FTPS can be used for the indoor security. As shown in Figure, it contains anti-theft and anti-fire alarm system, video monitors and emergency alarm functions, etc. The system primarily use the uniform coded of sensing window fences, monitor cameras, entrance guard devices, emergency calling devices , temperature sensors, and smart detectors of smoker combustible gas. To form the network of this subsystem home network, sensor network and the CIPS were used.

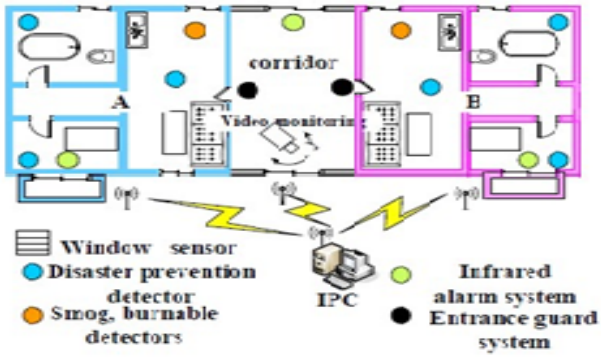


Fig. 3. Fire and Theft Prevention Subsystem of the ICSS

V. Security Challenges in IoT

A. Security Aspects

The term security subsumes a wide range of different concepts. In the first place, it refers to the basic provision of security services including confidentiality, authentication, integrity, authorization, non-repudiation, and availability. These security services can be implemented by means of different cryptography mechanisms. In the context of the IoT, however, security must not only focus on the required security services, but also on how these are realized in the overall system and how the security functionalities are executed [4].

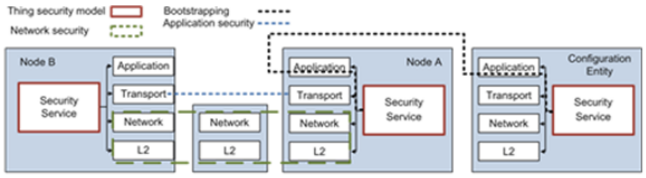


Fig. 4. Overview of Security Mechanism

1) The security architecture refers to the system elements involved in the management of the security relationships between things and the way these security interactions are handled (e.g., centralized or distributed) during the lifecycle of a thing.
2) The security model of a node describes how the security parameters, processes, and applications are managed in a thing. This includes aspects such as process separation, secure storage of keying materials, etc.
3) Security bootstrapping denotes the process by which a thing securely joins the IoT at a given location and point in time. Bootstrapping includes the authentication and authorization of a device as well as the transfer of security parameters allowing for trusted operation
4) Network security describes the mechanisms applied within a network to ensure trusted operation of the IoT. Specifically, it prevents attackers from endangering or modifying the expected operation of networked things. Network security can include a number of mechanisms ranging from secure routing to data link layer and network layer security.
5) Application security guarantees that only trusted instances of an application running in the IoT can communicate with each other, while illegitimate instances cannot interfere

B. Challenges for Secure Internet of Things

1) Constraints and Heterogeneous Communication: The IoT is a resource-constrained network that relies on lossy and low-bandwidth channels for communication between small nodes, regarding CPU, memory, and energy budget. These characteristics directly impact the threats to and the design of security protocols for the IoT domain. Scarce CPU and memory resources limit the use of resource-demanding crypto primitives, such as public-key cryptography as used in most Internet security standards. A further fundamental need refers to the limited energy budget available to IoT nodes. Careful protocol (re)design and usage is required to reduce not only the energy consumption during normal operation, but also under DoS attacks.

The tight memory and processing constraints of things naturally alleviate resource exhaustion attacks. Especially in unattended T2T communication, such attacks are difficult to notice before the service becomes unavailable. As a DoS

countermeasure, we implement return routability checks based on a cookie mechanism to delay the establishment of state at the responding host until the address of the initiating host is verified. However, they are less effective in broadcast media or when attackers can influence the routing and addressing of hosts.

Cryptographic payload processing applies message authentication codes or encryption to packets. These protection methods render the protected parts of the packets immutable as rewriting is either not possible because (a) the relevant information is encrypted and inaccessible to the gateway or (b) rewriting integrity-protected parts of the packet would invalidate the end-to-end integrity protection. The drawback of this approach, however, lies in its high signaling traffic volume compared to other approaches. Hence, future work is required to ensure security, performance and interoperability between IoT and the Internet.

2) Bootstrapping of a Security Domain: Most things might be required to support both centralized and distributed operation patterns. Distributed thing-to-thing communication might happen on demand. In todays IoT, most common architectures are fully centralized in the sense that all the security relationships within a segment are handled by a central party. A centralized architecture allows for central management of devices and keying materials as well as for the backup of cryptographic keys. Decentralized architectures, on the other hand, allow to create ad-hoc security domains that might not require an online management entity and are operative in a stand-alone manner.

Bootstrapping refers to the process by which a device is associated to another one, to a network, or to a system. The way it is performed depends upon the architecture: centralized or distributed. In a distributed approach, a DiffieHellman type of handshake can allow two peers to agree on a common secret. In a centralized architecture, preconfigured keys or certificates held by a thing can be used for the distribution of operational keys in a given security domain.

As the IoT involves not only passive devices, but also includes active and sensing devices, the IoT might irrupt even deeper in peoples privacy spheres. Thus, IoT protocols should be designed to avoid these privacy threats during bootstrapping and operation. Authentication can be used to prove membership of a group without revealing unnecessary individual information

3) Operation: After the bootstrapping phase, the system enters the operational phase. During the operational phase, things can relate to the state information created during the bootstrapping phase in order to exchange information securely and in an authenticated fashion.

Providing end-to-end security is of great importance to address and secure individual T2T or H2T communication within one IoT domain. Moreover, end-to-end security associations are an important measure to bridge the gap between the IoT and the Internet. End-to end security services includes peer entity authentication, end-to-end encryption and integrity protection above the network layer and the transport layer respectively. In addition to end-to-end security, group key negotiation is an important security service for the T2Ts and Ts2T communication patterns in the IoT as efficient local broadcast and multicast relies on symmetric group keys. Solutions that provide secure group communication at the network layer may have an advantage regarding the cryptographic overhead compared to application-focused security solutions.

It is expected that many things (e.g., wearable sensors, and user devices) will be mobile in the sense that they are attached to different networks during the lifetime of a security association. The specific need for IP-layer mobility mainly depends on the scenario in which nodes operate.

VI. Security Issues of the Internet of Thing

There are many problems in Internet of Things such as RFID tag security, wireless security, network transmission security, privacy protection, information processing security. The Internet of Things is integration of product of Wireless Sensor Network (WSN). The main feature of Internet of Things are comprehensive perception, using RFID, sensor, two-dimensional code to access to the information of the object anytime, anywhere.

A. Internet of Things Security Requirement

The Internet of Things safe issue is classified into: First is the, Physical Network. The important is Sensor Security including Sensor Interference. The Second is the Operation safe, it related to the normal operation of the sensor. The Third is the Data Security, it demands the information in the sensor, the transmission system, and the processing system. The security problems faced by the sensor and the sensor network (WSN) is more complex than the traditional information security. If these issues are not handled properly and efficiently the countrys economic and security will be threatened [5].

1) RFID Tag Information Security: Radio Frequency Identification (RFID) is an automatic identification technology through radio frequency. RFID security defects mainly in the following three areas:

- RFID identification itself to access security issues
- Security issues of the communication channel
- Security issues of the RFID reader

2) Wireless Communications and Information Security: The factors endanger the security of the information on Internet can also cause harm on the Internet of Things. Malicious intrusion of Things may result in violation of user privacy and user's actual loss. Cloud computing can collect the global hacker attacks node address, the host computer and other information, to share this information on the Internet, and global ASA firewall can real-time synchronous these libraries to prevent network attacks,

3) Privacy Protection: Privacy is another important part of the Internet of Things. It must provide confidentiality of IoT information, access control mechanisms to control IOT in the information collection, transmission. Privacy protection mainly involves the following questions:

- Data Privacy Protection:- Data privacy is confidentiality of the data itself. Data integration, aggregation, storage and other operations make the wireless sensor networks present a unique data privacy protection features based on data management.
- Location Privacy Protection:- The node location information often played the role of identity, and location privacy has a specialand crucial role in wireless sensor networks.
- Identity Privacy Protection:- Sensors and collected relevant information are often associated with the user entity, resulting in the user's identity leakage. This is one of the key issues that must be addressed to achieve large-scale deployment of wireless sensor network applications.

4) Information Processing Security: Internet has a complex security protection, but a huge number of nodes in IOT exist in the way of the cluster, thus will lead to large amounts of data sent simultaneously, the network congestion, resulting in denial of service attacks. Aggregation node and a large number of sensor nodes provide the information that have been monitored, perceived and collected. If these security issues are not handled, there will be a big risk on the application of IOT. Therefore, IOT security issues is bound to rise to the national level, and it is great significant to promote IOT security.

VII. Security Issues on Smart home in IoT Environment

Smart home is the smart life environment based on human that enables the convenience for people, promotion of welfare, and safety of life. In Smart home environment, all home devices are connected via Internet. By connecting the Smart home components to internet many security issues are arising. Smart home devices are tightly coupled with humans real life, so it is important to make sure that the Smart home are secure.

A. Smart home

All electronic devices in a home are connected through internet is called Smart home. When internet take over the control of Smart home there should a possibility of security attacks. Many of the Smart home device providers do not consider the security aspects. Therefore, there exists home gateway to control the information flows among smart devices and connect external network [7].

B. Differenet Attack Scenarios and Countermeasures

1) Trespass: Attacker could trespass on his/her home without destroying the Smart home devices. This threat could cause the loss of life and property. To protect this attack, password of smart devices could change frequently and hard to infer. Authentication and access control also have to be applied.

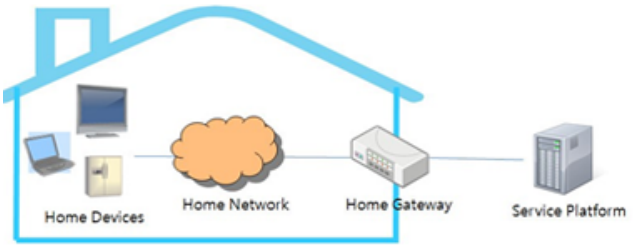


Fig. 5. Constitution of smart home

2) Monitoring and Personal Information Leakage: One of the purposes to use Smart home is safety. Therefore, there are many sensors for fire watch, housebreaking, baby monitoring, etc. If these are hacked by malicious codes, attackers could monitor inside the home. To protect this attack, data encryption between sensors and gateway has to be applied. It is also important to use authentication for detecting and blocking unauthorized devices and adopt anti-virus product.

3) Distribute denial of services: Attackers access smart home network illegally and send messages such as RTS(Request to Send)/CTS(Clear to Send) to smart devices in bulk. They also infect a target device using malicious codes and perform dos attack to a target device or other devices in Smart home network. To protect this attack, it is important to use authentication for detecting and blocking unauthorized devices. Security techniques such as rate limiting, null0 routing have be applied to home gateway.

4) Falsification: When smart devices communicate application server, attacker could gather packets by manipulating routing table in gateway. By doing this, they could falsify the contents or leak confidential information. To protect this attack, SSL technique with proper authentication has to be applied to Smart home components. It is also important for blocking unauthorized devices to access Smart home network.

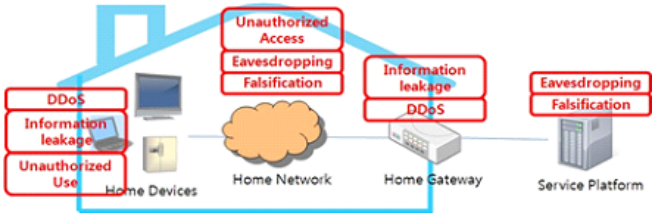


Fig. 6. Possible security threats in smart home

VIII. Technologies used in IoT environment and its security

There are four major technologies are used in the deployment of IoT [9]. They are:

1) Radio frequency identification (RFID)
2) Wireless sensor networks (WSN)
3) Middleware
4) Cloud computing

A. Radio Frequency Identification

Radio frequency identification (RFID) allows automatic identification and data capture using radio waves, a tag, and a reader. The tag can store more data than traditional barcodes. Applications of these can be found in supply chains, passports, and electronic tolls etc Active tags can contain external sensors to monitor temperature, pressure, chemicals, and other conditions.

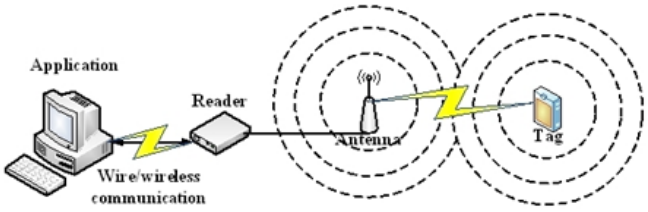


Fig. 7. RFID concept

B. Wireless Sensor Networks

Wireless sensor networks (WSN) consist of spatially distributed autonomous sensor-equipped devices to monitor physical or environmental conditions and can cooperate with RFID systems to better track the status of things such as their location, temperature, and movements. WSN allow different network topologies communication.

C. Middleware

Middleware is a software layer interposed between software applications to make it easier for software developers to perform communication and input/ output. Its feature of hiding the details of different technologies is fundamental to free IoT developers from software services that are not directly relevant to the specific IoT application.

D. Cloud computing

Cloud computing is the practice of using a network of remote servers hosted on the Internet to store, manage, and process data, rather than a local server or a personal computer. One of the most important outcomes of the IoT is an enormous amount of data generated from devices connected to the Internet. Many IoT applications require massive data storage, huge processing speed to enable real-time decision making, and high-speed broadband networks to stream data, audio, or video. Cloud computing provides an ideal back-end solution for handling huge data streams and processing them for the unprecedented number of IoT devices and humans in real time.

IX. Conclusion

The paper reviewed major security concerns, privacy issues, and smart home security issues and problems are all viewed. Securing only the application layer leaves the network open to attacks, while security focused only at the network and link Layer might introduce possible inter-application security threats. The IoT has done a drastic change in everyone's life. The connections between people and communications of people will grow and between objects to objects at anytime, in any location. The privacy and security implications of such an evolution should be carefully considered.

References

[1] Rolf H. Weber, "Internet of Things New security and privacy challenges", Computer Law & Security Review,2010
[2] J. Sathish Kumar, "A Survey on Internet of Things: Security and Privacy Issues", International journal of Computer Applications, 2014
[3] G. Gang, L. Zeyong, and J. Jun, "Internet of Things Security Analysis", International Conference on Internet Technology and Application 2011
[4] Tobias Heer, Ren Hummen and Klaus Wehrle, "Security Challenges in the IP-based Internet of Things", (2011)
[5] Chen Qiang, Bai Yu and Liu Yang, "Research on Security Issues of the Internet of Things", International Journal of Future Generation Communication and Networking, 2013
[6] L. Atzori, A. Iera and G. Morabito "The Internet of things: a survey", Computer Networks, 2010
[7] Seokung Yoon, Haeryong Park, and Hyeong Seon Yoo "Security Issues on Smarthome in IoT Environment"
[8] Erzen R, "Review of main security threats in Smart Home networks", 2012
[9] YANG Jin-cui, "Security model and key technologies for the Internet of things", 2011
[10] In Lee, Kyoochun Lee, "The Internet of Things (IoT): Applications, Investments, and Challenges for enterprises",

Combating Learning Disability with New Technologies

Raichel Sunny, Reshma P M, Sneha Theresa
Dept of Computer Applications
Vidya Academy of Science and Technology
Thrissur – 680501

Salkala K S
Assistant Professor of Computer Application
Vidya Academy of Science and Technology
Thrissur – 680501

***Abstract*—Learning Disabilities refer to a number of disorders which may affect the acquisition, organization, retention, understanding or use of verbal or nonverbal information. These disorders affect learning in individuals who otherwise demonstrate at least average abilities essential for thinking and/or reasoning. As such, learning disabilities are distinct from global intellectual deficiency. They results from impairments in one or more processes related to perceiving, thinking, remembering or learning. These include, but are not limited to: language processing; phonological processing; visual spatial processing; processing speed; memory and attention; and executive functions.**

***Index Terms*—Learning disability, Deficit Hyperactive Disorder, electroencephalograms, Digital Signal Processing, Soft Computing Methodologies, Intellectual disability, cognitive disability, Autism spectrum disorder, Auditory Processing Disorder, Visual Processing Disorder, Assistive technology, Inclusive classrooms.**

I. Introduction

Learning disability is a general term that describes specific kinds of learning problems. A learning disability can cause a person to have trouble learning and using certain skills. The skills most often affected are reading, writing, listening, speaking, reasoning, and doing math. Learning disabilities vary from person to person. One person with LD may not have the same kind of learning problems as another person with LD. One person may have trouble with reading and writing. Another person with LD may have problems understanding math. Still another person may have trouble in each of these areas, as well as with understanding what people are saying. It is a neurological disorder. In simple terms, a learning disability results from a difference in the way a person's brain is "wired." So that children with learning disabilities are as smart as or smarter than their peers. But they may have difficulty reading, writing, spelling, and reasoning, recalling and/or organizing information if left to figure things out by them or if taught in conventional ways.

LD is a group of disorders that affects peoples ability to either interpret what they see and hear or to link information from different parts of the brain. LD cannot be cured completely by medication and there is also no standard method for diagnosing it.Hence a computational approach to diagnose LD is suggested.This paper introduces different types of LD among primary level school children, some applications used in different level of people with learning disability. They are some kind of genetic disorder, which occurs during pregnancy and birth. They may be caused by illness or injury during or before birth. It is not caused by economic disadvantage, environmental factors or cultural differences. In fact, there is frequently no apparent cause for learning disabilities. It affects about 15% of the population, and can have a profound impact on individuals and families.

II. Learning Disabilities

Learning disabilities have nothing to do with how smart a person is. Rather, a person with a learning disability may just see, hear, or understand things differently. That can make everyday tasks, such as studying for a test or staying focused in class, much more difficult. There are strategies a person can learn to make it easier to cope with these differences.

To help a student with learning disability, first, accurate diagnosis for him/her, after that proper instruction methods for it are required. One of the effective instruction methods for students with learning ability is to check precedent knowledge of students before learning and to supplement it when want of precedent knowledge is found. It is a problem that affects how a person receives and processes information. Not every person with a particular disability will have all of the signs of that disability.

A. Specific Types of Learning Disabilities

A specific learning disability is unique to the individual and can appear in a variety of ways. It may be difficult to diagnose, to determine impact, and to accommodate.

Generally speaking, someone may be diagnosed with a learning disability if he or she is of average or above-average intelligence and there is a lack of achievement at age and ability level, or a large discrepancy between achievement and intellectual ability.

An untrained observer may conclude that a person with a learning disability is lazy or just not trying hard enough. He may have a difficult time understanding the large discrepancy between reading comprehension and proficiency in verbal ability. The observer sees only the input and output, not the processing of the information. Deficiencies in the processing of information make learning and expressing ideas difficult or impossible tasks. Learning disabilities usually fall within four broad categories:

A person with a learning disability may have discrepancies in one or all of these categories. The effects of an LD are manifested differently for different individuals and range from mild to severe. Learning disabilities may also be present along with other disabilities such as mobility or sensory impairments. Often people with Attention Deficit Disorder/Attention Deficit Hyperactive Disorder (ADD/ADHD) also have learning disabilities. Specific types of learning disabilities include:

Dysgraphia: A person with dysgraphia has a difficult time with the physical task of forming letters and words using a pen and paper and has difficulty producing legible handwriting.it can be characterised by problems with writing. This disorder may cause a child to be tense and awkward when holding a pen or pencil, even to the extent of contorting his or her body. A child with very poor handwriting that he or she does not outgrow may have dysgraphia

Dyscalculia: A person with dyscalculia has difficulty understanding and using math concepts and symbols. It include problems understanding basic arithmetic concepts, such as fractions, number lines, and positive and negative numbers.

Dyspraxia: Language comprehension of a person with dyspraxia does not match language production. She may mix up words and sentences while talking. It dyspraxia has problems with motor tasks, such as hand-eye coordination, that can interfere with learning.

Dyslexia: A person with dyslexia may mix up letters within words and words within sentences while reading.Usually have trouble making the connections between letters and sounds and with spelling and recognizing words He may also have difficulty spelling words correctly while writing; letter reversals are common. Some individuals with dyslexia may also have a difficult time with navigating and route finding using right and left or compass directions.

Nonverbal Learning Disorder: A neurological disorder which originates in the right hemisphere of the brain, causing problems with visual-spatial, intuitive, organizational, evaluative and holistic processing functions. It is a nonverbal learning disorder is demonstrated by below-average motor coordination, visual-spatial organization, and social skills.

B. Detecting Learning Disabilities

Learning disabilities can be hard to diagnose, because there is no definitive list of symptoms that fits every child. Also, many children try to hide the problem. You may not notice anything more obvious than frequent complaints about homework or a child who doesn't want to go to school.

Auditory Processing Disorder - is a problem with the way the brain processes the sounds a person takes in. It is not caused by hearing impairment. People with this disorder may have troubles

- Learning to read
- Distinguishing sounds from background noise
- Following spoken directions
- telling the difference between similar-sounding words
- Remembering things they've heard.

Visual Processing Disorder - Someone with a visual processing disorder has trouble interpreting visual information. He or she may have a hard time with reading or telling the difference between two objects that look similar. People with a visual processing disorder often have trouble with hand-eye coordination.

1) Treating Learning Disorders: Special education is the most common treatment for learning disorders. Under the Individuals with Disabilities Education Act (IDEA), all U.S. children with learning disorders are entitled to receive special education services for free in public schools.

After doing an evaluation to pinpoint where your child is having problems, a team of special educators will create an individualized education program (IEP) for your child that outlines what special services he needs to thrive at school. Special educators will then help your child build on his strengths and teach him ways to compensate for his weaknesses.

Many resources are also available outside of the public school system, including

- Private schools that specialize in treating children with learning disabilities
- After-school programs designed for children with learning disabilities
- At-home tutoring and therapy services

Learning disability doesn't have to be a roadblock to success. With the right tools, people with learning disabilities can overcome any challenge.

People with learning disabilities and disorders can learn strategies for coping with their disabilities. Getting help earlier increases the likelihood for success in school and later in life. If learning disabilities remain untreated, a child may begin to feel frustrated with schoolwork, which can lead to low self-esteem, depression, and other problems.

2) Parenting a Child with a Learning Disability: Finding out your child has a learning disability can be overwhelming. Many parents find the process of diagnosing a learning disability incredibly frustrating, and then once the diagnosis comes, they face an uphill battle to get their child the help he or she needs.

The best thing you can do as a parent is simply to love and support your child. These tips can also help you help your child:

- Learn everything you can. Get all the facts about your child's learning disability and how it affects the learning process. Research services and supportive strategies so

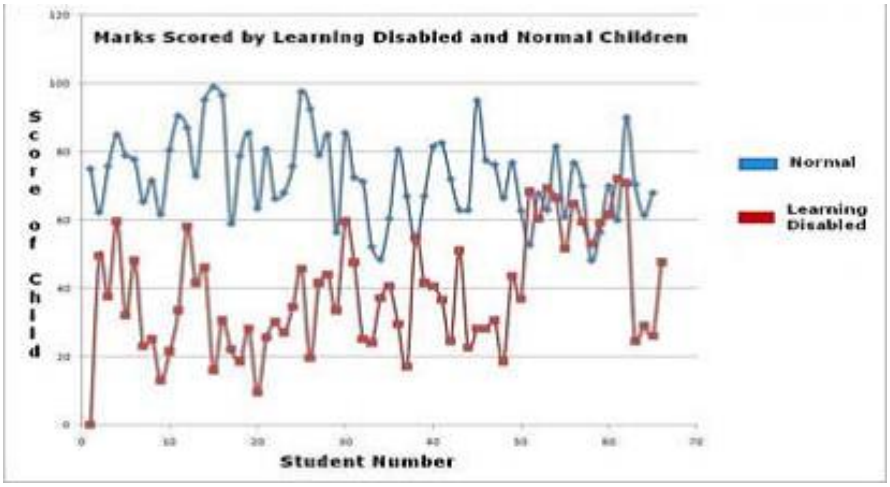


Fig. 1. Marks Scored by Learning Disabled and Normal Children

that you'll be able to take an active role in deciding on the right treatment for your child.

- Be your child's advocate. Work with your child's school to develop an IEP (Individualized Education Plan) - a special plan that sets goals for your child and describes support that may be needed to reach those goals. Understand special education laws and school policies so you can make sure your child is getting the most out of school. Many services may be available, but they may not be offered until you ask for them.

Many children have difficulty with reading, writing, or other learning-related tasks at some point, but this does not mean they have learning disabilities. A child with a learning disability often has several related signs, and these persist over time. The signs of learning disabilities vary from person to person.

These signs alone are not enough to determine that a person has a learning disability. A professional assessment is necessary to diagnose a learning disability.

Each learning disability has its own signs. Also, not every person with a particular disability will have all of the signs of that disability.

A kid might work with a tutor or specialist or even go to a special class. But often, kids with learning disabilities can continue in their regular classrooms and there's no reason they can't do normal stuff, like participate in school activities and sports.

C. Concept Map-Based Adaptive Tutoring System Supporting Learning Diagnosis for Students

An adaptive tutoring system that analyses learning characteristics of students with learning disability, diagnoses learning problems of them and according to that, they provides proper advice.

Learning disability can cause a student to have difficulty in learning and using certain skills. For study about learning disability in students, main thing is accurate diagnosis of him/her learning ability to check knowledge of students before schooling. This detail study will consider all the learning characteristics of each learner such as learning environment, goals etc These methods are integrated in the adaptive instruction system.

In adaptive tutoring system, learning materials are constructed using concept map. It is a graphical representation where nodes represent concepts, and links represent the relationships between concepts. It is used to assess students, achievements by misconception and missing concepts and to reinforce students understanding and learning for key concepts. The structure of concept map offers not only an overall cognition of subject contents but also the relationship to indicate the effect of learning one concept on the learning of other concepts. The system consists of user interface, information-collection, material-generation and diagnosis. The interface is classified into two parts for students and teachers. The tests are taken for students and test result is presented. By identifying those teachers can monitor the path, learning frequency, learning time etc The system was implemented with java on a Windows NT platform.

It is important to provide adaptable learning methods and learning materials considering the learning characteristics of each one of students with learning disability. The system can provide personalized suggestions for each student by analysing student answers and the relationships among the subject concepts and test items. Learning disability cannot be cured, is impact can be lessened through instructional intervention and compensatory strategies.[1]

D. Taxonomy of Computational Diagnosis of Learning Disability

The different computational methods and models used in diagnosing LD can be broadly classified into four groups. **Digital Signal Processing (DSP) methodologies** DSP techniques are to compare spoken words with pre-recorded and properly pronounced phonemes. The mispronounced phonemes were identified which led to the detection of LD.

	mild	moderate	Severe/profound
IQ range	50-69	35-49	<35
% of cases	85%	10%	5%
Ability to self care	Independent	Need some help	Limited
Language	Reasonable	Limited	Basic or none
Reading and writing	Reasonable	Basic	Minimal or none
Ability to work	Semiskilled	Unskilled, supervised	Supervised basic task
Social skill	Normal	Moderate	Few
Physical problems	Rare	Sometimes	Common
Aetiology discovered	Sometimes	Often	Usually
Academic skill	6th grade or higher	2nd to 3rd grade	-

Fig. 2. Features of Intellectual Disability

Digital Image processing (DIP) methodologies This used graphic signals to conclude that eye movements of even an infant could indicate LD.

Soft Computing Methodologies It applied a multi-layer feed forward perceptron to diagnose dyslexia where letter strings were mapped to phoneme strings in multi-syllabic words.

Hybridized computational techniques Hybridized approaches in LD include attempts to apply video and signal processing techniques along with soft computing techniques. They concluded that the reading speed increased with the probability of the patient being healthy. Wu et al combined different feature selection algorithms like brute-force, greedy and GA along with ANN to improve the identification rate of LD.

E. Preventing School Failure: Alternative Education for Children and Youth

Students have disabilities in their learning, especially in the basic education. This leads to the emergence of assistive technology in general education classrooms for disabled students. Its benefits include written expression, reading, mathematics, and spelling. When students have the writing challenges, they are more successful in the general education classroom. A necessary component for this is the collaboration between classroom teachers and assistive technology specialists.

Students with learning disabilities, technology can be an assistive tool replacing an ability that is either missing or impaired. It provides the support needed to accomplish a task. For example, word processing assists students with LD in improving writing

1) Computer Supports for Writing: Computers offer easier writing process for develop ideas, record, edit and to publish and share with others. Different computer supports are useful during different phases in the writing process. Some of them are talking word processors, word prediction, portable note taking devices, prewriting organizers, and multimedia prewriting prompts.

Talking Word Processors - It give the student creativity to the writing process. Letters, words, sentences, paragraphs, or entire documents can be read aloud while the student types. Features can be customized to individual based on students need.

Word prediction - It is a tool which augments spelling and syntax to enable users to make choices, find words, and complete sentences. Word prediction programs display words based on frequency of use, grammatically correct usage of words, and most recently used words. The dictionary will learn the word and predict it the next time it is used.

Portable Note-Taking Devices - It is an efficient means to record ideas and class- room notes to complete assignments, and to demonstrate writing creativity. Portable note takers allow more time for writing and require less concentration on operating the device

Prewriting Organizers - Some students with LD find graphic organizers helpful in mapping ideas during the planning stage. Graphic organizers such as Inspiration provide organizational frameworks to help writers generate topics and content for writing projects.

2) Role of Technology in Inclusive Classrooms: A sense of belonging to group, shared activities with individual outcomes and a balanced educational experience.

Adaptations for students with LD have been widely used to compensate for barriers associated with difficulties in reading, writing, mathematical reasoning, and problem solving. Increased use of assistive technology devices during cooperative learning activities can enhance the participation of students with LD by circumventing specific disability related barriers.

The most significant for introducing technology to the general education classroom are shared responsibility for par-

ticipation and decision making and for securing and sharing resources, and shared accountability for student outcomes. Assistive technology specialist evaluates students technology needs in collaboration with classroom teachers, related services staff, parents, and students. The school staff facilitates the evaluation process by identifying students strengths and the areas in which they are challenged in general education classrooms. Classroom teachers and students will be primarily responsible for the integration of technology into daily classroom routines.

Technology can help students with LD compensate for challenges in learning, especially in the area of writing, providing computer-supported tools. It can also ease frustration, increase motivation, foster a sense of peer acceptance, and improve productivity in the classroom and at home. Collaborative planning teams must develop a vision of technology for individual students and general education classrooms. Team members need to determine the effectiveness of current technology and closely monitor students to ensure that the necessary modifications are made to reflect the changing abilities of the individuals. The potential of assistive technology for students has not been realized; the future is uncertain but holds much promise. For individuals with disabilities, this technology can be one way to break down barriers to learning.[3]

F. Serious Games And Image And Perception Representation Of Words For Learning Disabilty People

This paper deals with games, image and perceptual representation for disabled people. Learning disability is not a health problem.It is a situation that,disabled people cannot perform the actual operations normally.They may be slow in behavior when compared with normal people. There are many ways to help and support disabled people.We can support the disabled people by finding out their interested areas. Mental support for them is more important.First of all we stand with them to convince that disability is just a situation, the situation can be change by their mental power.

1) Serious games support for disabled people: Serious game is an application.By using this application we can create a mind for the disabled people that they can do the same things that normal people do.They will concentrate in games, then by they will complete the game in proper time and if they succeed in games the people will be happier. The people may realize that they can perform actions that normal people do.

2) How imaginal representation helps disabled people: Imaginal representation is also an important thing that support disabled people.Imaginal representation helps them to identify anything in better way.It may take time to understand them to found out that what has been shown there in the picture.If the identified the image properly its a good result of they can understand or observe things clearly,no matter the time delay.The disabled people feel desperate when the normal people respond to the environment easily and fastly.This is not a memory problem.It is dependent to the time needed to recall the things that are already found by their memory.Here is the application of games useful.

3) Role of games and imaginal representation: By looking out these paper we can came into a conclusion that the situation of being disabled can be procure, not completely. But someway by the support of normal people the situation can be change.There are many more games for supporting and devolipng the abilities of disabled people.In imaginal representation the mode of interaction is different.The results indicate that learning disability and normal children do not differ in perceptual processing time,however learning disabled children have smaller image capacity than normals.Correlational results also suggest that the imaginal mode is independent of other modes for normals than learning disabled.[12]

III. Conclusion

In this paper we discussed about learning disabilities A learning disability is not anyones fault. Many people with learning disabilities say that their problems seemed worse before they understood their disability. By learning about his differences, a child will grow to understand them. Through this understanding and support at home and in school, he can avoid shame. The effects of the disability can be reduced through intervention. And he can develop strong skills in other areas.

This paper also discussed about Dyslexia which is a complex problem that has no simple solution. The most widely accepted view is that dyslexia is a language based disorder.It is important that any therapy for learning disabilities be scientifically established to be valid before it can be recommended for treatment the evidence does not support the concept that vision therapy or tinted lenses or filters are effective, directly or indirectly, in the treatment of learning disabilities. Thus, the claim that vision therapy improves visual efficiency cannot be substantiated. Diagnostic and treatment approaches that lack scientific evidence of efficacy are not endorsed or recommended.With early recognition and individualized, interdisciplinary management strategies, children with learning disabilities can enjoy successful academic experiences.

In future, it is intended to apply other soft computing techniques for early diagnosis of Learning diability.

References

[1] Sook-Young Choi *A concept map-baesd adaptive tutoring system supporting learning diagnosis for students with learning disability*,2004
[2] Bender W.N, *Learning disabilities : Characteristics, identification and teaching strategies*,1992
[3] Jane quenneville *Preventing school failure : Alternative education for children and youth*,2015
[4] Alves S, Marques A, Queiros C, Orvalho V *LIFEisGAME prototype: a serious game about emotions for children with autism spectrum disorders*,2013
[5] Karal H, Kokoc M, Ayyildiz U *Educational computer games for developing psychomotor ability in children with mild mental impairment*,2010
[6] Lecavalier L, Snow A, Norris M *Autism spectrum disorders and intellectual disability*,International Handbook of Autism and Pervasive Developmental Disorders Autism and Child Psychopathology Series, Springer, Heidelberg,2011
[7] Connolly T.M *A systematic literature review of empirical evidence on computer games and serious games*,2012
[8] Wouters P, van der Spek D, van Oostendorp H *Current practices in serious game research: a review from a learning outcomes perspective*,2010
[9] Blank M, Bridges W.H *Deficiencies in verbal labeling in retarded readers*,1996

[10] Sotiriou S, Anastopoulou S, Rosenfeld S, Aharoni O, Hofstein A, Bogner F, Sturm H, Hoeksema K *Visualizing the invisible: the CONNECT approach for teaching science*,2006

[11] Bialo ER, Sivin-Kachala J *The effectiveness of technology in schools: a summary of recent research*,1996

[12] Ana R. Cano, Alvaro J. Garcia-Tejedor, Baltasar Fernandez-Manjon *A Literature Review of Serious Games for Intellectual Disabilities*,2006

[13] Kathleen M, McCoy, Robert J. Weber *Image and Perceptual Representation of Words in Learning Disabled and Normal Children*,2014

IoT Security Issues : A Survey

Shabab E, Robin Paul, Prasad K V
Department of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501

Aparna S Balan
Assistant Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur - 680501

***Abstract*—Internet of things is emerging as the third wave in the development of the internet .It expands the Internet services.The Internet of Things at large will foster billions of devices, people and services to interconnect and exchange information and useful data.In this paper, the IoT and existing security threats, and open challenges in the domain of IoT are discussed. The current state of issues on IoT security is understood by researching on newspapers. It encompasses many aspects of life from connected homes and cities to connected cars and roads, roads to devices that track an individual and use the data for collected for push services. This paper provides an overview of IOT with emphasis on enabling technologies and IOT Applications security issues from all international newspapers ,Websites and Books specifically to the issues of security paramount in IOT devices in all over the world .**

***Index Terms*—IoT Security issues, IoT Security Issue Reports, IoT applications.**

I. Introduction

The Internet of Things is an emerging topic of technical, social, and economic significance. IoT (Internet of Things) is an advanced automation and analytics system which exploits networking, sensing, big data, and artificial intelligence technology to deliver complete systems for a product or service. These systems allow greater transparency, control, and performance when applied to any industry or system. IoT systems have applications across industries through their unique flexibility and ability to be suitable in any environment. They enhance data collection, automation, operations, and much more through smart devices and powerful enabling technology.

A. IoT Advantages

1) **Improved Customer Engagement :**
   Current analytics suffer from blind-spots and significant flaws in accuracy; and as noted, engagement remains passive. IoT completely transforms this to achieve richer and more effective engagement with audiences.
2) **Technology Optimization :**
   The same technologies and data which improve the customer experience also improve device use, and aid in more potent improvements to technology.
3) **Reduced Waste :**
   IoT makes areas of improvement clear. Current analytics give us superficial insight, but IoT provides real-world information leading to more effective management of resources.
4) **Enhanced Data Collection :** Modern data collection suffers from its limitations and its design for passive use.

B. IoT Disadvantages

1) **Security :**
   IoT creates an ecosystem of constantly connected devices communicating over networks.
2) **Privacy :**
   The sophistication of IoT provides substantial personal data in extreme detail without the user's active participation.
3) **Complexity :**
   Some find IoT systems complicated in terms of design, deployment, and maintenance given their use of multiple technologies and a large set of new enabling technologies.
4) **Flexibility :**
   Many are concerned about the flexibility of an IoT system to integrate easily with another.
5) **Compliance :**
   IoT, like any other technology in the realm of business, must comply with regulations

II. IoT Applications And Issues

From building automation and smart factories to wearables, the IoT touches every facet of our lives. TI makes developing applications easier with hardware, software and support to connect anything to the internet. We have identified six key markets for the IoT with potential for exponential growth.

A. Smart Home

Smart Home has become the revolutionary ladder of success in the residential spaces and it is predicted Smart homes will become as common as smartphones. The cost of owning a house is the biggest expense in a homeowners life. Smart Home products are promised to save time, energy and money[1].

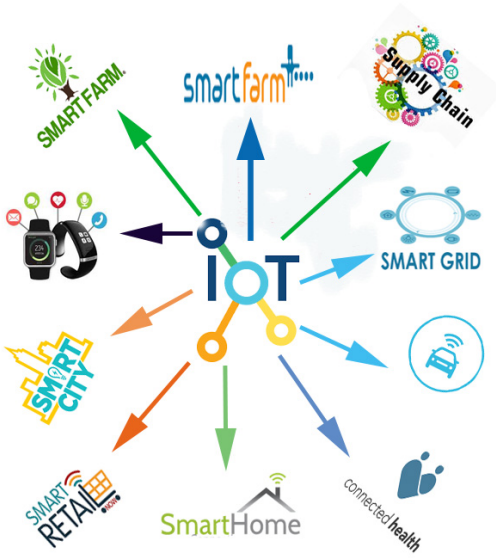


Fig. 1. IoT

Reported Issues:

1) **'Smart' home devices used as weapons in website attack :**
Hackers used internet-connected home devices, such as CCTV cameras and printers, to attack popular websites on Friday . security analysts say. Twitter, Spotify, and Reddit were among the sites taken offline on Friday. Each uses a company called Dyn, which was the target of the attack, to direct users to its website. Dyn is a DNS service - an internet "phone book" which directs users to the internet address where the website is stored. It came under attack - a distributed denial of service (DDoS) - which relies on thousands of machines sending coordinated messages to overwhelm the service [5].
2) **Takeaways from October's IoT DDoS attack :**
Last month, there was a massive DDoS attack that was made possible by hacking into unsecured IoT devices, mainly home surveillance cameras. The truth of the matter is that the attack was the result of a vulnerability on cheap cameras and other IoT devices. They are designed to be accessed over a local network and they come with unsecured, hard-coded default passwords. Unfortunately, many people own these types of devices, which led this cyberattack having such a wide reach. The problem here is that most consumers dont know how to secure networks and unknowingly expose themselves to such vulnerabilities [6].
3) **Dumb mistakes open smart lighting to hackers :**
An IoT security consultant shows an overflow crowd at LuxLive just how easy it is to waltz into a Wi-Fi network, steal passwords, and all that stuff.
LONDON A security consultant who makes a living from hacking the Internet of Things (IoT) warned the lighting industry here that if it's not careful, its smart LED lighting systems could be extremely vulnerable to attack [7].

B. Wearables

Wearable devices are installed with sensors and softwares which collect data and information about the users. This data is later pre-processed to extract essential insights about user.These devices broadly cover fitness, health and entertainment requirements. The pre-requisite from internet of things technology for wearable applications is to be highly energy efficient or ultra-low power and small sized [1].

Reported Issues:

1) **Security firm BitDefender demonstrated that the Bluetooth communication between Android devices and smartphones could be deciphered using brute-force attacks :**
Recently, security firm BitDefender demonstrated that the Bluetooth communication between Android devices and smartphones could be deciphered using brute-force attacks. hackers opt for persistent trial and error, trying username and password combinations until they crack the code and are able to access contents stored on devices [8].
2) **Survey: Wearable devices most likely to pose IoT security threat :**
Across the U.S., Europe, the Middle East and Africa, 53 percent of IT professionals said that smart watches, fitness trackers and other wearables were somewhere between likely and extremely likely to create a threat. Moreover, almost 90 percent of said the flood of IT devices entering the market creates security and privacy issues in the workplace, with 84 percent of survey-takers naming the growing number of entry points into the network as a major concern [9].
3) **Wearables leaking password details :**
Wearable is an IOT application. Researchers from Binghamton University and the Stevens Institute of Technology have revealed that wearable devices have the ability to leak passwords. Wearable Devices Reveal Your Personal PIN, the researchers collated data from embedded sensors in wearable technologies, such as smart watches and fitness trackers, along with a computer algorithm to ascertain PINs and passwords. An attacker can also place a wireless sniffer close to a key-based security system to eavesdrop sensor data from wearable devices sent via Bluetooth to the victim's associated smartphones [10].

C. Connected Cars

The automotive digital technology has focused on optimizing vehicles internal functions. But now, this attention is growing towards enhancing the in-car experience. A connected car is a vehicle which is able to optimize its own operation, maintenance as well as comfort of passengers using onboard sensors and internet connectivity. Most large auto makers as

well as some brave startups are working on connected car solutions. Major brands like Tesla, BMW, Apple, and Google are working on bringing the next revolution in automobiles [1].

Reported Issues:

1) **Internet Of Things Meets Cars : Security Threats Ahead :**
   One major topic is how to collect data while protecting the vehicle owner's privacy. the "hacker" may simply be your five-year-old kid using a backseat touchscreen to inadvertently purchase something via an online retailer's (all-too-convenient) one-click-buying feature. Today, this factory-default reset button doesn't exist, a potential privacy problem as consumers start to buy and sell used connected cars. We're going to have to have an easy way as consumers to do that in our automobiles, because we don't want my driving record, driving history, and how much I'm going over the speed limit transferred to a third party [22].
2) **Connected cars: security and privacy risks on wheels :**

   In July 2015, Wired magazine broke the story that hackers had taken control and killed the accelerator of a Jeep in motion on the freeway. 1.4 million vehicles to fix the bug that enabled the attack, and chip maker Intel formed an industry task force, the Automotive Security Review Board to focus on securing "cyber-physical systems in vehicles. But in the race to secure vehicles, the industry is coming from behind[23].

D. Industrial Internet

The industrial internet is also one of the special Internet of Things applications. While many market researches such as Gartner or Cisco see the industrial internet as the IoT concept with the highest overall potential, its popularity currently doesnt reach the masses like smart home or wearables do. The industrial internet however has a lot going for it. The industrial internet gets the biggest push of people on Twitter (1,700 tweets per month) compared to other non-consumer-oriented IoT concepts [1].

Reported Issues:

1) **Passing the burden on to third party cloud and security providers may not help with increased security or cost savings :**
   Industrial customers turn to third party VPN providers and cloud access companies, but in doing that they arent solving the problem, just passing it along onto others for convenience. The DragonFly attack was just that; a sophisticated and very well funded campaign targeting cloud VPN services

E. Smart city

Smart city spans a wide variety of use cases, from traffic management to water distribution, to waste management, urban security and environmental monitoring. Its popularity is fueled by the fact that many Smart City solutions promise to alleviate real pains of people living in cities these days. IoT solutions in the area of Smart City solve traffic congestion problems, reduce noise and pollution and help make cities safer [1].

Reported Issues:

1) **Risk And Repeat: DNS DDoS attacks raise concerns over IoT devices :**
   The major DDoS attacks on DNS provider Dyn Inc. Last week have raised new concerns about how insecure IoT devices can be used for malicious purposes. The domain name system (DNS) DDoS attacks caused problems for Dyn and intermittently disrupted major websites such as Netflix, Amazon, Twitter, Reddit and others. Tens of millions of discrete IP addresses associated with the Mirai botnet were part of the attack [11].
2) **Most Internet of Things devices have privacy issues:** The Information said by Privacy Commissioner for Nova Scotia Catherine Tully . Privacy Commissioner office focused on 14 health products that were found in drugstores. So, generally, blood glucose monitors, heart monitors, blood pressure monitors, some fitness bands, among other products, The study determined that 59 per cent of devices reviewed didnt properly explain to buyers how their personal information was collected, used and disclosed. The products reviewed in Nova Scotia, the rate is 90 per cent [12].
3) **IoT broadens attack surface of Smart Cities :**
   It may sound like the plot of a Philip K. Dick novel, but headlines in recent months have decried several attacks on public and private websites. To be sure, the Internet of Things promises more reliable and easy access to myriad industrial and municipal systems. However, as smart cities start investing in smart meters and other devices that could fall prey to attacks engineered by botnets taking advantage of unsecured IoT devices and other IP-connected electronics and systems [13]
4) **Risk And Repeat: US accuses Russia of state-sponsored cyber attacks :**
   The issue of cyber attribution has become a topic of debate following the U.S. government's formal statement accusing the Russian government of state-sponsored cyber-attacks. Intelligence released a joint statement last week accusing the Russian government of orchestrating cyber-attacks against state election systems and the Democratic National Committee, among others. .The attacks have in some cases resulted in data breaches exposing confidential emails, which the U.S. intelligence community believes are politically motivated [14].
5) **Over 33M records leaked from US corporate database :**

The database contains email addresses and other contact information for thousands of corporate and government employees, reports ZDNet. It contains information - such as names, job titles and functions, work email addresses and phone numbers - on employees of thousands of companies and government agencies, including the US Department of Defense, the US Postal Service and US Army, ZDNet reported Wednesday. It's still unclear how the data was exposed. As general practice, Dun and Bradstreet uses an agile security process and evaluates and evolves security controls to protect the integrity of our data [15].

6) **Mirai malware to create a huge network of botnets which hacked the servers :**
Mirai malware to create a huge network of botnets which hacked the servers of DYN, a web service provider resulting in disruption in services of popular sites like Twitter, Netflix and Amazon alike. The report even pointed out that 70% of such devices are vulnerable. Most of the devices are plug-n-play, sell-and-forget. Botnet harvesters identify firmware vulnerability and proceed to locate and exploit all the devices deploying the same firmware [16].

F. Smart farming

Smart farming is an often overlooked business-case for the internet of Things because it does not really fit into the well-known categories such as health, mobility, or industrial. However, due to the remoteness of farming operations and the large number of livestock that could be monitored the Internet of Things could revolutionize the way farmers work. But this idea has not yet reached large-scale attention. Nevertheless, one of the Internet of Things applications that should not be underestimated. Smart farming will become the important application field in the predominantly agricultural-product exporting countries [1].

Reporting Issues:

1) **FBI Warns of Smart Farm Risk :** Farmers who are looking to make better use of technology need to start paying attention to security, or suffer the same fate as industries such as healthcare. A dangerous precedent as farmers invest in connected and data intensive farming equipment and related services . Possible risks include hacktivists who destroy data to protest the use of genetically-modified organisms (GMOs) or pesticides. Farm-level data may also be vulnerable to ransomware and data destruction [27].

G. Smart Retail

The potential of IoT in the retail sector is enormous. IoT provides an opportunity to retailers to connect with the customers to enhance the in-store experience. Smartphones will be the way for retailers to remain connected with their consumers even out of store. Interacting through Smartphones and using Beacon technology can help retailers serve their consumers better. They can also track consumers path through a store and improve store layout and place premium products in high traffic areas [1].

H. Smart Grid

Power grids of the future will not only be smart enough but also highly reliable. Smart grid concept is becoming very popular all over world. The basic idea behind the smart grids is to collect data in an automated fashion and analyze the behaviour or electricity consumers and suppliers for improving efficiency as well as economics of electricity use. Smart Grids will also be able to detect sources of power outages more quickly and at individual household levels like nearby solar panel, making possible distributed energy system [1].

Reported Issues:

1) **Report: Energy Sector Security Improvements Imperative :**
A new report from McAfee details the thoughts of industry leaders on the state of energy security. A cybercriminal could debilitate a major city by a single targeted attack on the energy grid and compromise anything from the lights and appliances in homes, to heart monitors in hospitals, to air defense systems, the press release states. According to the report, the most prevalent cyberthreat reported by the global energy sector is extortion. Criminals gain access to a utility's system, demonstrate that they are capable of doing damage, and demand a ransom[17].
2) **Algeria Attack Uncovers Fresh Energy-Sector Security Concerns :**
The deadly attack on an Algerian natural gas complex will do little to discourage the drive for lucrative energy exploration in northern Africa. the attack is forcing companies to increase security after largely ignoring risks of operation in the remote desert region, The Associated Press reports. Last Wednesdays assault on Algerias Ain Amenas gas complex by a multinational band illustrates the danger posed by Al-Qaida and its offshoots. Algeria has taken a strong tack against the terrorists [18].
3) **Experts Point to Terrorism for Metcalf Power Substation Attack :**
An April attack on Silicon Valleys phone lines and power grid was terrorism . The FBI has released multiple statements that is has found no indications to support this claim.It involved snipping AT And T fiber-optic lines to knock out phone and 911 service, and firing shots into a PG and E substation . This is the most sophisticated and extensive attack thats ever occurred on the grid to my knowledge [19].
4) **PG And E Offers $250K Reward for Substation Attack Leads :**
According to The Associated Press, the attack left AT and T fiber-optic lines disconnected. the attackers fired shots into the substation, knocking out 17 transformers and causing more than $15 million in damage.The reward would be funded by shareholders and comes

nearly a year after AT and T offered its own $ 250,000 reward for information leading to an arrest [20].

5) **Thousands Could Die in Power Grid Attack : "Damaged Transformers Could Take Years to Replace :** A terrorist attack on the U.S. power grid could result in the deaths of thousands of people, as well as incurring costs of hundreds of billions of dollars, according to the National Academy of Sciences. While such an attack would not immediately kill people. If the event occurred during extreme weather, heat or exposure to cold could lead to hundreds or even thousands of deaths, . A 2011 report from the Electric Power Research Institute says that about $3.7 billion in investment is needed to protect the grid from cyber attacks. Energy companies, including utilities, would have to increase their investment in computer security more than seven-fold to reach an ideal level of protection [21].

I. Connected Health (Digital health/Telehealth/Telemedicine)

Connected healthcare yet remains the sleeping giant of the Internet of Things applications. The concept of connected healthcare system and smart medical devices bears enormous potential not just for companies, but also for the well-being of people in general. Research shows IoT in healthcare will be massive in coming years. IoT in healthcare is aimed at empowering people to live healthier life by wearing connected devices. The collected data will help in personalized analysis of an individuals health and provide tailor made strategies to combat illness. The video below explains how IoT can revolutionize treatment and medical help [1].

Reported Issues:

1) **Healthcare Cybersecurity Attacks Rise 320 % from 2015 to 2016 :** Healthcare organizations are a prime target for cyber-attacks. Healthcare providers have to prioritize a proactive approach to security - balancing people, process and technology to improve the protection of their informational assets and patient information. Most of the 2016 healthcare data breaches stemmed from hackers, with an overall increase in healthcare cybersecurity attacks of that kind rising 320 percent, according to recent research from Redspin. Healthcare providers have become the primary targets of malicious hackers, and their attacks are becoming increasingly sophisticated and disruptive to operations [24].

J. Smart supply chain

Supply chains have been getting smarter for some years already. Solutions for tracking goods while they are on the road, or getting suppliers to exchange inventory information have been on the market for years. So while it is perfectly logic that the topic will get a new push with the Internet of Things, it seems that so far its popularity remains limited [1].

Reporting Issues:

1) **Land Rush: Race is On To Hack Vulnerable IoT Devices :** Cyber criminal groups are racing to gain control over a population of insecure "Internet of Things devices . NyaDrop looks for embedded systems that use the MIPS architecture, a common feature of broadband routers and other small, embedded systems. Login information captured by the researchers at Malware Must Die suggest that the attacks are looking for common hardware like DAHUA brand IP-enabled network security cameras, though the attackers are likely targeting a range of Internet connected devices [25].

2) **Shoddy Supply Chain Lurks Behind Mirai Botnet :** China-based supplier of management software is the common thread that ties together the myriad digital video recorders, IP-based cameras and other devices that make up the Mirai botnet. The software used by over five-hundred thousand devices on public IPs around the world, meaning they cannot be changed and make the devices susceptible to trivial compromise. Those attacks were the largest denial of service attacks, measured by the volume of bogus Internet traffic used to cripple their targets [26].

III. Application of TRIZ Framework for Resolving Security Issues in IOT

Apply the TRIZ methodology to resolve the underlying contradictions in order to the findthe solutions to the problems with IOT devices:

The current security mechanisms are not capable of dealing with IOT security issues. In this paper, The authors use TRIZ methodology to find solution to the security threats making IOT more robust

A. Methodology

Somebody somewhere has already solved the problem. The challenge now is to seek that solution and modify it according to the specific problem at hand .

1) Specific Problem
2) General Problem
3) General Solutions
4) Specific Solutions

B. Applying TRIZ to IOT security problems

1) Specific Problem : Find the specific problem that concerns the area, to which an effective solution has to be found.

There are some critical security issues, effects of which can be detrimental if not addressed.

- Security
- Enterprises Security
- End-User Privacy
- Data
- Storage Management
- Server
- Device tampering

2) General Problem : Reduce the problem into its elemental components and apply the contradiction matrix. Those generic problems are then put through the TRIZ contradiction matrix.

Some of the worsening feature for the security issues with IoT include :

- Loss of information
- object generated harmful factors
- device complexity
- ease of manufacture
- power

3) General Solution : The contradiction matrix gives the specific inventive principles out of the 40 inventive principles that give us the possible solution to eliminate the contradictions.

Some of the principles that we get from the matrix are :

- Principle 1 Segmentation
- Principle 2 Taking Out (Extraction)
- Principle 10 Preliminary action
- Principle 22 Blessing in disguise
- Principle 27 Cheap short living objects
- Principle 35 Parameter changes

4) Specific Solutions : We then formulate and draw analogies from those generic principles into our specific problem and find specific solutions. Several security measures have been devised in order to prevent data theft and hacking these methods adopted for IoT along the lines of inventive principles are :-

- Principle 1 Segmentation
- Principle 2 Taking Out (Extraction)
- Principle 10 Preliminary action
- Principle 22 Blessing in disguise
- Principle 27 Cheap short living objects
- Principle 35 Parameter changes

C. Applications of the Proposed Security Solutions

The new encryption methods as block chain cryptography techniques can be to provide the much-needed encryption to the low processing IoT devices.

One of the major problem which is likely to be faced by this project would be generation of huge amount of data that is useless. In such a scenario, The superfluous data can be removed and valuable storage space can be saved.

Device to device communication over a network is one of the imperative features of IoT. Data from one device to another needs to be encrypted with a very immune technology.

D. Limitations of the Work

- The application of TRIZ matrix and deducing the general problems and solutions out of the matrix is based on the authors discretion.
- The extent of the search for articles on IOT and its security and concerns was limited to a few publishers and thus the whole panorama of security issues may not have been covered.
- This research work will be the first step in applying TRIZ to IOT problems [2].

IV. Security Requirements Analysis for the IoT

IoT(Internet of Things) is described as collaborative ecosystem of context-aware, intelligent and automated device connected network for specific purpose. Gartner, Cisco and IDC evaluate IoT as a promising technology of future. A lot of corporations in the world are developing IoT-related devices, services and technologies to dominate the market in advance. However, they do not consider security as a functional requirement. This paper proposes security requirements based on three characteristics of IoT and six key elements in IoT.

A. Analysis of characteristics in IoT

1) **Heterogeneity :**
In IoT, heterogeneity means diversity of hardware performances, protocols, platforms, policies etc. The biggest problem of heterogeneity is the absence of common security service. Heterogeneity weakens interoperability and also it make security-related policies and updates more complex. In order to solve these problems, we can use some technologies such as MDR, middleware. But this is not a fundamental solution. For providing common security services, unified IoT security standards has to be established. Recently, standard organizations develop some standards for the security in IoT.
2) **Resource Constraint :**
Most IoT devices are lacking performance and battery capacity. Therefore, legacy security services, such as TLS,AES cannot be applied to IoT devices directly. Therefore, these services or algorithm should be designed to be lightweight and straightforward to increase efficiency of CPU, memory and battery.
3) **Dynamic environment:**
Due to mobility and bad connections, IoT has a dynamic network topology. In very demanding cases, numerous devices send a large number of requests. Hence flexibility and scalability is required in IoT communication protocols.

B. Security issues and requirements for IoT environments

Six key elements of IoT are:

1) **IoT network:**
IoT network is basically not different from conventional networks. Therefore most existing problems could happen in IoT network.
   - **Privacy :** Encryption and authentication is necessary to be made use of bitwise operation rather than mathematical algorithm like ECC in order to make lightweight security services. Privacy does not always have to be protected.
   - **Security in multitasking :** When using multicasting, multicast group should be created with

authenticated users and secret key that is shared with group members is required to keep security.

- **Security in bootstrapping :** Bootstrapping is a process that sends data to participate secure IoT network. So bootstrapping is required to designed in a flexible, scalable and lightweight manner.

2) **Cloud :**
IoT devices use clouds because they cannot save the data in their low memory capacity. If cloud out of order for some reason, IoT devices cannot save data. Thus critical data may be missing. In this case, availability is highly necessary so that device should have backup cloud.
There are a lot of data sent from many devices in cloud. To protect the data from unauthorized users, cloud should use proper access control, encryption, data anonymity etc.
3) **User :**
User is the most vulnerable element in IoT security. Even if information system is implemented securely, if a user is careless to manage, any security system will be useless.
4) **Attacker :**
Due to IoT devices are connected to network, it can be attacked any time. Most of IoT devices cannot apply strong security services because of its constrained resources. In IoT environment, security threats can be categorized into non-physical threats and physical threats. Non-physical threats can be described as threats which uses network and physical threats can be described as all threat except for non-physical threat. Most of the non-physical threats are attacks on confidentiality, integrity and availability.
5) **Service :**
To take the advantage of a service, the user needs to trust the server, and the server need to provide privacy to the user. If the user decides the server is trustworthy, the user will use the services provided by the server or group of devices with smart phone, smart watch or some kind of network devices. After that, the device have to progress bootstrapping and access control. Thereby, devices obtain trust from server. Finally, attacker can compromise the server for malicious intentions.
In IoT environment, middleware can be used for providing security to devices and data. Access control is also a security service which include authentication and authorization for ensuring security and privacy.
6) **Platform :**
oneM2M, OIC and other standards organizations have been established platform standards. The platform needs to minimize its vulnerability and should be verified itself using the trusted platform module(TPM) to avoid attacks [3].

V. Mitigating IoT Security Threats with a Trusted Network Element

Securing the growing amount of loT devices is a challenge for both the end-users bringing loT devices into their homes, as well as the corporates and industries exposing these devices into the Internet as part of their service or operations. The exposure of these devices, often poorly con figured and secured, offers malicious actors an easy access to the private information of their users, or potential to utilize the devices in further activities, e.g., attacks on other devices via Distributed Denial of Service. Network Edge Device offloads the security countermeasures of the individual devices into the trusted network elements. The major benefit of this approach is that the system can protect the loT devices with user-defined policies, which can be applied to all devices regardless of the constraints of computing resources in the loT tags. Additional benefit is the possibility to manage the countermeasures of multiple loT devices/gateways at once, via a shared interface, thus largely avoiding the per-device maintenance operations.

A. IoT Security Challenges

IoT Security Challenges is a four-layer loT security architecture. In essence, this architecture aggregates the layered interfaces of Functional Decomposition view of ARM (Architectural Reference Model for loT) to attack surfaces. ARM is the main model of this architecture developed in FP7 research project loT-A. The four-layer architecture is applicable for NED benefit analysis in loT protection.

The architectural layers are defined as follows :

- **Perception layer:** Comprises of the tags, the physical sensor and actuator devices, with RF connection.
- **Network layer:** Forms the communication network, commonly wireless sensor network, connecting the tags to the information processing (back -office) system.
- **Middleware layer:** Consists of the information processing systems, the databases providing storage capabilities; service oriented with the goal to provide similar services to all the connected nodes.
- **Application layer** Comprises of the various business-logic applications, creating the added-value loT applications to implement a smart space, smart logistic, or even smart grid.

Selected IoT attack types in different architecture layers are listed in table 1.

B. Secured Platform and NED

The idea of SECURED is to off-load the security controls (e.g., malware protection, IDS/IPS, network monitoring, etc.) from the individual devices (laptop, tablet, mobile, loT device) to the closest trusted network element, a network edge device (NED), thereby providing the same security level for each user's device in the SECURED platform, without heavy processing power restrictions on the individual devices. The SECURED platform provides a userfriendly way to define the security policies in a high-level language for the actual

TABLE I
IOT SECURITY ATTACK TYPES IN DEPARTMENT ARCHITECTURE LAYERS

Perceptional Layer	**Network Layer**	**Middleware Layer**	**Application Layer**
Unauthorized access	Sybil attack	Unauthorized access	Code injection
Tag cloning	Sinkhole attack	Denial of service	Denial of service
Eavesdropping	Sleep deprivation	Insider attack	Special-Phishing
Spoofing	Denial of service		Sniffing
RF jamming	Code injection		
	Man-in-the-Middle		

end-users without the need of having to be an expert in, e.g., frrewall configurations.

C. Current and Proposed Countermeasures

The proposed NED-hosted countermeasures help in mitigating security threats in the typical situation, where terminal nodes have only limited communication and computation resources. Concentrating the countermeasures on the network edge, where more resources are available, makes it possible to select targeted and strong countermeasures, depending on the requirements and situations in the loT application.

1) Analysis of countermeasure off-loading possibilities : Countermeasures are technical or non-technical, and some set of technical countermeasures can be implemented outside of loT device. When the countermeasures are located at the network edge, through which the loT devices are communicating to the Internet, they can be easily updated in case a vulnerability is discovered in the loT devices and/or the countermeasures can be used to mitigate the impacts of a successful exploit in one of the loT devices. The defmed policies for individual users (loT gateways, for instance) are fetched from the policy repository, so the configuration of the policies for a general type user loT gateway can be updated from one place to multiple devices. A typical loT system can be protected by one or multiple NEDs depending on the network deployment[4].

VI. CONCLUSION

In conclusion, the Internet of Things is closer to being implemented than the average person would think. Most of the necessary technological advances needed for it have already been made, and some manufacturers and agencies have already begun implementing a small-scale version of it. The main reasons why it has not truly been implemented is the impact it will have on the legal, ethical, security and social fields. Workers could potentially abuse it, hackers could potentially access it, corporations may not want to share their data, and individual people may not like the complete absence of privacy. For these reasons, the Internet of Things may very well be pushed back longer than it truly needs to be.

REFERENCES

[1] "Internet of things applications,".[Online].Available:https://iot-analytics.com/10-internet-of-things-applications/.
[2] Pulkit Sharma, Rishabh Rahul Khanna, Vishal Bhatnagar,"Application of TRIZ Framework for Resolving Security Issues in IOT",*International Conference on Computing, Communication and Automation* ,2016
[3] Se-Ra Oh, Young-Gab Kim,"Security Requirements Analysis for the IoT" ,2017
[4] Jarkko Kuusijarvi, Reijo Savola, Pekka Savolainen, Antti Evesti,"Mitigating loT Security Threats with a Trusted Network Element ",*The 11th International Conference for Internet Technology and Secured Transactions* ,2016.
[5] "Smart home devices used as weapons in website attack".[Online].Available:http:// www.bbc.com/ news/technology-37738823.
[6] "Takeaways from October's IoT DDoS attack".[Online].Available: http://internetofthingsagenda. techtarget.com/ blog/ IoT-Agenda/ Takeaways- from- Octobers-IoT-DDoS-attack.
[7] "Dumb mistakes open smart lighting".[Online].Available: http://www.ledsmagazine.com/articles/2016/11/ dumb-mistakes-open-smart-lighting-tohackers.html.
[8] "Five Potential Security Concerns Related to Wearables".[Online]. Available:http://mobile.itbusinessedge.com /slideshows/ five-potential-security-concerns-related-to-wearables.html.
[9] "Survey wearable devices most likely to pose iot security threat".[Online]. Available:https://www.scmagazine.com /survey-wearable-devices-most-likely-to-pose-iot-security-threat/article/528566/.
[10] "Research revealsthreat wearables leaking password details".[Online]. Available:http://www.wearable technology-news.com/news/2016/jul/13/research-revealsthreat-wearables-leaking-password-details.
[11] "Risk Repeat DNS DDoS attacks raise concerns over IoT devices" [Online]. Available:http://searchsecurity. techtarget.com/podcast/Risk-Repeat-DNS-DDoS-attacks- raise-concerns-over-IoT-devices .
[12] "Most internet of things devices have privacy issues study" [Online]. Available:http://globalnews.ca/news /2957825/most-internet-of-things-devices-have-privacy-issues-study/
[13] "IoT threat smart city".[Online].Available: https://gcn.com /articles /2016/12/23/iot-threat-smart-city.aspx
[14] "Risk Repeat US accuses Russia of-state sponsored cyber-attacks" [Online]. Available:http://searchsecurity. techtarget.com/podcast/Risk-Repeat-US- accuses-Russia-of-state -sponsored-cyberattacks
[15] "More than 33 million records leaked from us corporate database" [Online]. Available:https://www.cnet.com/ news/more-than-33-million-records-leaked-from-us-corporate-database/
[16] "How techgiants startups tackle iot security challenge".[Online]. Available:http://tech.economictimes.indiatimes.com/news/ internet/how-tech-giants- startups-tackle-iot -security-challenge/55691691.
[17] "Energy sector security improvements imperative".[Online].Available: http://www.securitymagazine.com/articles/83314-report--energy- sector-security-improvements-imperative.
[18] "Algeria attack uncovers fresh menergy sector security concerns".[Online].Available:http://www.securitymagazine.com/ articles/83949-algeria-attack-uncovers-fresh-energy-sector-security-concerns.
[19] "Experts point to terrorism for metcalf power substation attack".[Online].Available:http://www.securitymagazine.com/ articles/85208-experts-point-to-terrorism-for-metcalf-power-substation -attack
[20] "Pge offers 250k reward for substation attack leads".[Online]. Available:http://www.securitymagazine.com/articles/ 85419-pge-offers-250k-reward-for-substation-attack-leads.
[21] "Thousands could die in power grid attack".[Online]. Available:http://www.securitymagazine.com/ articles/ 83754-thousands-could-die-in-power-grid-attack
[22] "Internet of things meets cars security threats ahead".[Online].Available:http://www.informationweek.com/

big-data/big-data-analytics/internet-of-things-meets-cars-security-threats-ahead/d/d-id/1127737.

[23] "Connected cars security and privacy risks on wheels".[Online]. Available:https://iapp.org/news/a/connected-cars-security-and-privacy-risks-on-wheels/

[24] "Healthcare cybersecurity attacks"[Online].Available:http://healthitsecurity.com/news/healthcare-cybersecurity-attacks-rise-320-from-2015-to-2016&hl=en-IN

[25] "Land rush race is on to hack vulnerable iot device".[Online]. Available:https://securityledger.com/2016/10/land-rush-race-is-on-to-hack-vulnerable-iot-devices/&hl=en-IN

[26] "securityledger"[Online].Available:https://securityledger.com/2016/10/shoddy-supply-chain-lurks-behind-mirai-botnet/&hl=en-IN

[27] "FBI warns of smart farm ris".[Online].Available:https://securityledger.com/2016/04/fbi-warns-of-smart-farm-risk/

Confidentiality and Access Control in Some Popular Cloud Service Providers

Sigma Kochumon, Haritha T, Silpa C A
Department of Computer Applications
Vidya Academy of Science & Technology
Thrissur-680501

Sajay K R
Associate Professor of Computer Applications
Vidya Academy of Science & Technology
Thrissur-680501

***Abstract*—In recent technologies cloud computing is the one of the most popular and advance technology. However when providing the data and business application to cloud third party there is a chance to some security and privacy issues. In this paper we discuss about the cloud confidentiality,security challenges in cloud computing and solution to these issues by access control and study about various popular cloud service providers.**

***Index Terms*—cloud computing,security,privacy,trust, confidentiality,integrity,accountability,availability,Access control,Cloud Service Providers,Dropbox,GoogleDrive,Microsoft OnDrive,Amazon Cloud Drive,iCloud**

I. Introduction

cloud computing providers users and organization to store their data in cost effective manner and empowering their business by delivering software and services over internet to large user base.Now world wide spend billions of money for cloud computing.However because cloud is open platform there is a chance to different attacks.Security of stored data access management and trust are among the primary security aspect in cloud computing .here we discuss about such security challenges and some method to protect data. Many cloud services providers offer a large amount of space for consumer.Here we focus on google drive,one drive ,drop box,amazon cloud services and icloud.All cloud service providers had their own merits and demerits .A survey of cloud service providers are perform in this paper.

II. Confidentiality in Cloud Computing

Cloud computing is one of the emerging technologies in computer science. Cloud computing is a mechanism in which all servers, networks, applications and other elements related to data centers are made available to IT and end users via the internet. Confidentiality is one of the greatest concerns with regards to cloud computing. When dealing with cloud environments, confidentiality implies that customer's data and different computational task are kept to be confidential from the cloud providers and other customers. Trusted cloud providers have the ability to create a unified data protection policy across all clouds. Data confidentiality is one of the challenges in the ongoing research in cloud computing.

A. Background for Confidentiality

Data storage is the main part of the cloud computing, which renders data confidentiality as one of the critical issues in the cloud. In this section describes the concept of database as a services and benefits, architecture of database outsourcing model challenges associated with it.

1) Database as a Service: Database as a service is a service that is managed by the cloud operator.it provides users with some form of access to a database without the need of for setting the physical hardware, installing software. It consists of database manager component which controls all underlying database instances via an application programming interface. Each database services as a unique name and connection request can include database service name. services are built into the oracle database,providing a single system image for workload and prioritization of workload. Database as a service is a cloud strategy used to facilitate the accessibility of business critical data in a well timed, protected and affordable manner. It depends on the principle that specified useful data can be supplied to users on demand, irrespective of any organizational or geographical separation between consumers and providers.it eliminates the redundancy.

2) Architecture of Outsourced Database Model: Generally three entities are involved in database outsourcing that are data owner, clients and service provider. The data owner is responsible for uploading the data at the service provider's site. The data owner has the authority to permit or deny the clients for accessing the database. The clients are also called as queriers who access the database according to the privilege level acquired by them. The service provider performs all the data maintenance tasks. For efficient data communication, the transmission link between the service provider and the data owner as well the link between service provider and the clients should be of high bandwidth.

There are three kinds of architectural models of outsourced databases. That are uniform client model, Multiple client outsourced database model , multiple data owner database outsourcing . In uniform client model the data owner and

the client are the same. The data owner performs all the operations on database (Insert, update, delete). This is the simplest database outsourcing model. The Multiple client outsourced database model comprise of single data owner and multiple clients (queriers). In multiple data owner database outsourcing model each data owner uploads the data at the service provider's site. In this individual access control policies for each group of data owner and clients are needed to be implemented. The Attribute based access control, accounting and authorization are some mechanisms which can be incorporated in multiple data owner model .

B. Difference Between Data Confidentiality and Security

When we talk about confidentiality of information, we are talking about protecting the information from disclosure to unauthorized parties.Information has value, especially in today's world. Every one has information they wish to keep a secret. Protecting such information is a very major part of information security.A very key component of protecting information confidentiality would be encryption. Encryption ensures that only the right people can read the information. Encryption is very widespread in today's environment and can be found in almost every major protocol in use. Other ways to ensure information confidentiality include enforcing file permissions and access control list to restrict access to sensitive information.

Confidentiality agreements are often applied to situations where someone trusted with personal data must safeguard this data from being released. Alternately, some may define confidentiality as issues about the data that gets collected, where privacy issues have to do, again, with the core principle of an individual not being recorded or monitored.

Data security is commonly referred to as the confidentiality, availability, and integrity of data. In other words, it is all of the practices and processes that are in place to ensure data isn't being used or accessed by unauthorized individuals or parties. it ensures that the data is accurate and reliable and is available when those with authorized access need it. A data security plan includes facets such as collecting only the required information, keeping it safe, and destroying any information that is no longer needed. These steps will help any business meet the legal obligations of possessing sensitive data.

Data Security is a different term that's applied to enterprise or government systems. It may include the idea of customer privacy, but the two are not synonymous. Likewise, security may provide for confidentiality, but that is not its overall goal. The overall goal of most security systems is to protect an enterprise or agency, which may or may not house a lot of vulnerable customer or client data. Sometimes, the objectives for privacy and security are the same. In other cases, security may not automatically provide for privacy concerns.

One example is where a business or government agency may be able to keep its data safe from outside attackers, but where employees may be able to view consumer information. Another scenario might involve situations where a company doesn't face any liability by releasing customer data, and so chooses to do so. Here, the company's security is not jeopardized, but the consumer's privacy is violated. New contracts between businesses and federal agencies are also good examples of how IT issues cut through the different layers between privacy, confidentiality and security.

C. Key Requirements of Confidentiality Aspects in Outsourcing

1) Confidentiality: authorized users and systems are given consent to access the data. Data confidentiality refers to keep data from unauthorized access when it stored and also when the data is in transit state. User privacy and access privacy are the primary requirements. In user privacy the user identity when he fetches or manipulates the data. Access privacy is assured when the access pattern of database and intended database records for a particular user are kept secret.

2) Integrity: it assures that the data stored in database are not modified/ manipulated except by trusted persons. Completeness and correctness are two important dimension of integrity. Completeness means that query results obtained by fetching all records from the database. Correctness means that the query result obtained from server are generated by original server.

3) Availability: it ensures that data are available to the trusted users and system when they access database in an authorized manner. It is recommended for service providers to provide always on availability of database to their valued and authorized users.

4) Authenticity: it ensures that query transaction and communication are genuine. To provide authenticity the digital signature are used.

5) Freshness: it guarantees that the query results are obtained by executing the queries over the most updated database.

6) Risk Management : it includes the set of activities to identify and track the data security and also the control measures are set to avoid the further risk to security.

III. Access Control in Cloud Computing

Cloud computing is the most popular technology of internet that store large amount organizational and individual data.Privacy and security of data are important concerns in cloud computing.Access control is the fundamental aspect of information security that is directly tied to the primary characteristics such as confidentiality,integrity and availability .It is originated in 1960's.Access control is a procedure that restrict other users from accessing to the data stored in cloud by granting access permission to different users.It ensure the use of data resource legal.

From the access control perspective the cloud service provider must provide basic functionalities that

- control the access to service feature of the cloud based specific policies and level of service purchased by the customer

- control access to consumer's data from other consumers in multitenant environment
- control access to regular user functions and privileged administrative function and maintain access control policy
- up to date user profile information.

A. Classification of Access Control Model in Cloud Storage

Access control model can be traditionally classified into three categories that is DAC(Discretionary Access Control),MAC(Mandatory Access Control) and RBAC(Role Based Access Control).In DAC model owner of the of the object decide the access permission of other users.DAC model is only used in legacy application .It has huge management overhead in modern and multi-application environment.The MAC model is more adaptable for distributed system compared to DAC model.In MAC model model administrator of the system decide the access permission and the subject and object is identified with a security level of classification .In RBAC model access permission is assigned to users based on role of the users.Traditional access control models are not available in dynamic,distributed,multiple secure domain environment.So the advantages of these model are taken and develop new models.Here we discuss about some of the access models.

1) Role Based Access Control Model : Role based access control(RBAC) model was introduced by American National Standardization Technical Committee in 90s.RBAC model are more scalable than traditional discretionary and mandatory access control model.RBAC model help to control the access of data in cloud storage based on the concept role.In RBAC model the access permission are assigned to role and these roles are assigned to different users based on their business function in organization.Role is act as a bridge between users and access permission.The access control procedure include two steps one is association of access permission and role and other is association of role and user.

TRBAC(Task RBAC) model is an important access control model in cloud computing environment where the traditional discretionary,mandatory and role based access control model cannot be employed.TRBAC model validate access permission for user dynamically based on assigned role and the task user perform with assigned role.ARBAC(Attribute RBAC) model is another role based access control model proposed for cloud computing.In ARBAC model certain attributes and value are assigned to data object.User with specific role has to submit a value for these attribute and access to the data object is provide if the user submitted value is valid.

2) Trust Based Access Control Model : Trust management provide a new thought of solving the security problems in cloud computing environment.In this model access control is based on the trust rank assigned to users.In TBAC (Trust Based Access Control) model trust relationship between user and cloud computing platform is established based on user's behaviour and trust degree calculated by trust model. Trust management used the interpersonal relationship trust model of sociology for reference, it assess the complex trust relationship in computer network by using scientific method, and considered all kinds of factors that could impact the trust adequately in the procedure of assessing the relationships. By analyzing the user's action identify the trust relationship between cloud computing platform and users,based on this a trust rank is assigned to users.Using trust rank access permissions are assigned to users.

TBDAC (Trust Based Dynamic Access Control) is a trust based access control model proposed for cloud computing.It combined trust management and RBAC.It provide a light weight certificate to users by using this certificate user could not only certify its identity validity but also get the access right via trust rank and role information in certificate.

3) Access Control Model of Cloud Storage Based on Attribute Based encryption: The access control based on ABE (Attribute Based Encryption) combine the data encryption with access control.ABE is suitable to protect the privacy and secrecy of data in cloud computing environment.ABE is more useful when the source data not know about identity and public key of recipient only know about certain attribute of recipient. In access control based on ABE the plaintext is encrypted and stored in the cloud storage and control the access to data by limiting the decryption ability.For limiting the access to to data and decryption ability we need to implement access control to cipher text. Access control based on ABE used traditional identity based encryption(IBE) for reference,define user's identity as group of attitudes.The user can only decrypt the ciphertext if the attitude group of identity met the access control structure.

CP-ABE (ciphertext-policy attribute-based encryption)access control is a proposal for ABE based access control model.In this the user was associated to a group of attributes and the data was associated to a group of attribute condition.The user could decrypt the data if satisfied the attribute condition.

4) Other Access Control Models: BLP (Bell-La Padula) model was introduced by David Bell and Leonard La Padula. It was an important access control model and was used to protect confidential data by the military.In this model different safety grades are assigned to access subject and object . The subject could read the object whose safety grade was lower than the subject and the subject could edit the object whose safety grade was higher than the subject. Using this way, the security of the data was ensured.

C. D. Shen and M. X. Yan.had introduced a scheme that was to apply the BLP model to the cloud computing environment. The cloud service provider store the user data in cloud storage and security and stored information of data is stored in server. When the user logged in, the login server would check the user's validity and define the user's safety grade. The server would get the security information and operation matrix information of the data, and compare with the user's security information. If the user's operation was permitted, the data's stored information would be given. Otherwise the denial of service information would be returned. IFC(information flow control) and DIFC-AC (Decentralized Information Control

based Access Control) are another proposed access control models in cloud computing.

IV. Survey of Popular Cloud Service Providers

Cloud service providers (CSP) are companies that offers network services, infrastructure, or business applications in the cloud. The cloud services are hosted in a data center that can be accessed by companies or individuals using network connectivity. The large benefit of using a cloud service provider comes in efficiency and economies of scale. Rather than individuals and companies building their own infrastructure to support internal services and applications, the services can be purchased from the CSP,which provide the services to many customers from a shared infrastructure. The popular cloud service providers are

- Dropbox
- Google Drive
- Microsoft OneDrive
- Amazon cloud drive
- iCloud

A. Dropbox

Dropbox is one of the most popular online file storage/ sharing services currently available. Each user must register an account on dropbox to obtain a certain amount of storage space and the first few gigabytes are free.

Here is what makes Dropbox so popular.Advantages of dropbox are

- Cost : Dropbox is completely free, with no hidden charges leaping out at users at any time.
- Capacity : Users are given 2 GB of storage capacity. This may not seem like much, but it is certainly plenty to start off with. Adding storage is also possible - and without incurring charges. Users have a whole array of free options to increase their storage capacity, including, among others.
- Accessibility : Extremely easy to use, Dropbox can either be used online or by installing the Dropbox application .With the app installed, it is possible to access files from wherever a user is, any time on any device, even if offline at the time (the Dropbox folders on all devices are updated automatically every time a file is saved). Mobile users are able to download only required files in order to save space.
- Sharing : Files can be shared with individual users or public by designating folders to specific users or enabling them for public sharing.
- Back-up : Dropbox is a convenient solution to backing up files in a safe and easy to access fashion. By storing files in the cloud, data sticks, CDs, DVDs and other easily lost or damaged storage devices - and the cost involved - become a thing of the past.
- Synchronization : We could synchronize our files effortlessly after we uploaded the files to our account. What's more, deleting files accidentally will not result in losing them, because Dropbox keeps easy to retrieve copies of all deleted files for 30 days.

Security weaknesses of Dropbox are

- Collaboration : One con found during our Dropbox review was the inability to collaborate on files. The only files that could be synchronized were the files we had uploaded. The files on our computer could not be synchronized with shared users unless we uploaded them first.
- File Location : At times we found it hard to find the location of certain files. No specific tabs were available and we had to initially select different tabs to find some older file versions or deleted copies.

B. Google Drive

Google Drive came into existence after Google decided to re-brand the popular Google Docs service. Since then, Google Drive has found its own fans and haters in equal amounts. With amazing collaboration features, this cloud syncing and storage service is definitely a game changer in the market.

Some advantages of Google drive are:

- Anytime accessibility : Files and documents stored in the data centers of Google can be accessed by a user while sitting in any part of the world, by simply logging into their Google drive account. Corporations have gained enterprise mobility dues to this reason.
- Device independency : There is no specific platform defined for using the file uploaded on Googles drive. Any device, which is connected to network, can be used to access the file stored on drive.
- Integration : Documents stored on Google drive can easily be opened in any other Google app. This provides a better integration of this service with other cloud applications, so as to provide hassle-free experience to the user.
- Ease of Searching : Documents can easily be searched for, using Google Drive service. Text can also be searched, even from images using OCR (Optical Character Recognition) technology. Its an intelligent way of searching for the text even in the image files.
- Limitless file size sharing : Sharing has been made easy among a specified group of people as special editing permissions can be given to them. Moreover, the storage limit has also been increased from a few MBs to a few GBs. Users can share any size of the file without any limit to any specific file size.

With a number of advantages, certain security and privacy issues have also been highlighted in Google drive service. Some of them are listed below:

- Google holds the authority to index your data, stored on Google drive. So in case, some keyword matches with any title of you photograph, some text in you document etc. could float in the search result.
- Internet based service means full dependency on internet to access the data. As, data is stored remotely so some connection is required to access the data. In case of a

scenario where the Internet isnt functioning, access to your data in such situation becomes impossible.

C. Microsoft OneDrive

Microsofts OneDrive is one of the leading cloud storage service platforms on the market and an obvious choice for both businesses and users who employ Microsoft Office packages such as Word and Excel on a regular basis. SkyDrive was renamed and revamped into OneDrive by Microsoft due legal disputes with Sky Broadcasting, and since then, the company has been pushing to get more and more customers to use their cloud services.

The benefits of using OneDrive are

- Free storage : OneDrive offers users 15GB of free storage space as well as the chance to earn extra free storage space. Microsoft has introduced a referral incentive where users gain extra storage for every friend that signs up to an account through them. Additional storage is also offered if users link OneDrive to their mobile phones camera so that it automatically backs up their photos online.
- Easy organisation : You can store any kind of file on OneDrive be it photos, video, and documents, and then access them from any of your Windows PCs or mobile devices. Files are organised by type, so its easy to find what you need.
- Close collaboration with Office As a Microsoft platform, OneDrive works closely with Microsoft Office apps, such as Word or PowerPoint when you launch one of these applications youll see a list of recent documents, including those saved to OneDrive. If you have an Office 365 subscription and open a document saved in OneDrive, you can collaborate with it in real-time with other people.

Potential security issues:

- User error : Microsoft Windows remains the number one targeted platform for hackers and while OneDrive has remained fairly free of any serious breaches, users must follow standard security procedures. To ensure security users must use strong passwords and ensure they have chosen the right file sharing permissions.
- Encryption: Users of the standard OneDrive service will find that data is encrypted in transit using SSL but it will remain unencrypted at rest. However OneDrive for Business incorporates per-file encryption which encrypts files individually each with a unique key; so if it is compromised only one individual file would be accessed rather than the whole store.
- Syncing: There are also suggestion that some files can get altered when they are synced or uploaded to OneDrive but it is unclear how common this problem is.

D. Amazon Cloud Drive

In 2006, Amazon Web Services (AWS) started to offer IT services to the market in the form of web services, which is nowadays known as cloud computing. With this cloud, we need not plan for servers and other IT infrastructure which takes up much of time in advance. Instead, these services can instantly spin up hundreds or thousands of servers in minutes and deliver results faster.

The benefits of Amazon cloud drive are

- Easy to use: AWS is designed to allow application providers, ISVs, and vendors to quickly and securely host your applications whether an existing application or a new SaaS-based application. users can use the AWS Management Console or well-documented web services APIs to access AWSs application hosting platform.
- Flexible: AWS enables you to select the operating system, programming language, web application platform, database, and other services you need. With AWS, you receive a virtual environment that lets you load the software and services your application requires. This eases the migration process for existing applications while preserving options for building new solutions.
- Cost-Effective: Only pay for the compute power, storage, and other resources you use, with no long-term contracts or up-front commitments which makes AWS cost efficient.
- Reliable: With AWS, you take advantage of a scalable, reliable, and secure global computing infrastructure, the virtual backbone of Amazon.coms multi-billion dollar online business that has been honed for over a decade.

The drawbacks of Amazone cloud drive are

- Sharing: Amazone cloud drive provides limited sharing capabilities.
- Synchronization: No synchronization features.
- Desktop app is very limited,you can only restore entire drive,not individual folders via the desktop.
- This service is also available for computers only, no mobile devices.

E. iCloud

iCloud is a storage and cloud computing service from Apple.It allows users to store data sucha as music files on remote computer server for download to multiple devices.This sevice is providedby the Apple to make data backup and restore.In this service you will be given an ID and password from Apple with the iOS device.Using that ID you can easily login into the account of iCloud and upload or download their iOS data. That data is saved to remote server making it accessible worldwide. .Apples iCloud is more of an extra perk available to Apple users rather than an individual service. Most iClouds features only work if youre using an Apple device, and if youre not, its really unnecessary. iCloud fits perfectly into the Apple ecosystem, but lacks integration into other systems and apps.

Advantages of using Icloud are

- Storage: The unlimited storage capability allows one to store things like your personal device, ringtones, text messages, applications, music and photos from your camera roll feature. After storing in the iCloud, pictures can be seen by anyone with access and it can be used from any computer with internet access.

TABLE I
COMPARISON OF CLOUD SERVICE PROVIDERS.

	Drop Box	Google Drive	One Drive	Amazone cloud Drive	iCloud
Synchronization	Y	Y	Y	N	Y
Web Access	Y	Y	Y	N	Y
Backup	Y	N	Y	N	Y
Sharing	Y	Y	Y	Y	Y
Android Support	Y	Y	Y	N	N
Local Encryption	N	N	N	N	Y
Multiple Account	N	N	N	N	Y

- Cost: It is free of cost and the user is not required to pay for it. iCloud provides 5GB of free online storage.If you need more space than that,you can subscribe to an upgrade with a low monthly service charge.
- Safety: The biggest advantage of the iCloud service is how you can use it to back up and restore data on all your Apple products. If you update any media in one device, other device that sync with the same iCloud account will also be updated with the same media. This service come in handy especially when an individual purchases a new device since all of the date is saved on this virtual cloud, it is incredibly easy to sync the saved data onto the new device.
- Sync feature: The convenience of having all of your data synced in one place,no matter which device you happen to be using.you dont have to use a USB chord to sync information and furthermore this is a free service.

Disadvantages of using iCloud are

- Cost: The service is free but data transfer charges do apply.Transferring data from or to iCloud requires active Internet connection which is not free.If you upload any purchased Cd songs and tracks via iTunes then they would not be send to iCloud and hence to any other device. So all you have to do is pay $25 yearly matching up your tracks and it incurred additional cost.
- security and privacy : One of the biggest concerns about iCloud is security and privacy. Users might not be comfortable to upload their data to a third party service and feel that their data might be accessed by any unauthorized users.
- iCloud service is compatible only with IOS devices and has widespread accessibility issues.

V. CONCLUSION

In this paper, we have explained the concept of database as a services and its benefits. The architectural models of the outsourced databases are also elaborated. The thorough analysis of general security requirements for the outsourced databases is done in this paper. In survey of access control firstly sum up the main security threats to data confidentiality then classifies and introduce the existing access control models. Cloud storage comes in all shapes and forms.Direct comparison between providers is often difficult because they focus on different aspects of the service.Often people will base their decision on the amount of free storage available.However this is the only one element.Each of the cloud service providers have their own merits and demerits.

REFERENCES

[1] Zhifeng Xiao and Yang Xiao,"Security and Privacy in Cloud Computing",IEEE communication survays and tutorials 2013
[2] Xinlu li and Xiaoxia Zhao,*"Survay on access control model in cloud computin environment"*,*International Conference on cloud computin and big data*,2013
[3] Z Tan, Z Tang, R Li, A Sallam,"Research on trust-based access control model in cloud computing",IEEE Technology and Artificial 2011
[4] Meghanathan N " Review of access control models for cloud computing". In: ICCSEA, SPPR, CSIA, WimoA Computer Science & Information Technology (CS & IT). 2013. p. 7785.
[5] Z Abdul Raouf Khan,"Access Control in Cloud Computing Environment",ARPN Journel of Engineering and Applied Science 2012
[6] Ms.Mayuri R Gawande and Mr.Arvid S. Kazse,*"Analysis of data confidentiality techniques in cloud computing"*,International journal of computer science and mobile computing,2014
[7] Jin Hn,joseph k Liu,Jia xu,and Jianing Zhou,*Security Concerns in Popular Cloud Storage Services*,2013
[8] Zahir Tari , Xua yi , Uthpala S.Premarathne,Peter Bertok and Ibrahim Khail,*"Security and Privacy in Cloud Computing: Vision, Trends, and Challenges"*,IEEE computer society,2015
[9] Cheng Kang Chu,Wen-Taozhn ,*"Security Concerns in Popular Cloud Storage Services"*,2013