

EDITORIAL BOARD

Editor-in-Chief

Dr. Vinita Gaikwad

Editorial Board Members

Ms. Sreeja.S.S

Ms. Puja Agarwal

Student Members

Ms. Anagha Baikar

Ms. Anagha Patwardhan

Ms. Apurva Khedkar

Ms. Bhavika Vartak

Mr. Debrata Pal

Mr. Matsyendra Mishra

Mr. Rohit Maurya

Ms. Shraddha Mhatre

Mr. Sandeep Kamat

Mr. Yogesh Pawar

PUBLISHER

Thakur Institute of Management Studies, Career Development & Research

Thakur Educational Campus,

Shyamnarayan Thakur Marg,

Thakur Village, Kandivali (E), Mumbai – 400 101

Telephone: 022 - 2884 0484/91, 022 - 67308301/02

Fax: 022 – 28852527

Email: timsedr@thakureducation.org, vinita.gaikwad@thakureducation.org

Website: www.timsedrmbai.in

AIM

The research and development has been transforming the computing paradigms and technology in multidimensional directions. Tech Tonics aims to inculcate research culture among post graduate students and make them aware of new innovational happenings in the field of information technology.

TECH TONICS

TIMSCDR Research Journal

Volume 4
Academic Year: 2015-2016

Published By

Thakur Institute of Management Studies, Career Development & Research
Thakur Educational Campus, Shyamnarayan Thakur Marg,
Thakur Village, Kandivali (E), Mumbai – 400 101

Vision

Thakur Institute of Management Studies, Career Development and Research will become a premier Institute renowned internationally for providing education in Software Application to graduates from various disciplines.

Mission

To achieve excellence in providing software education so that students can grasp existing as well as emerging technologies and to inculcate leadership and managerial qualities in them so that they can deliver results in the organization they join.

Quality Policy

We, the staff, faculty and the management of Thakur Institute of Management Studies, Career Development and Research are committed to offer excellence in software education, conducive academic environment and state of-the-art infrastructure to our students

We will work as a team and interact with students in pro-active manner to achieve our Quality Objectives and fulfill all academic, statutory and regulatory requirements to entire satisfaction of our students as well as for continual improvement of QMS.

PROGRAM EDUCATIONAL OBJECTIVES (PEOs)

- PEO-1 :** To enable students to gain knowledge across all domains of Information Technology with in-depth understanding of their applications.
- PEO-2 :** To enable students to analyze problems and to design and develop software solutions using emerging tools and technologies.
- PEO-3 :** To enable students to continue Life-long learning, Research and Entrepreneurial pursuit in their chosen fields.
- PEO-4 :** To develop communication, teamwork, and leadership skills necessary to manage multidisciplinary projects and serve the society as responsible and ethical software professionals.

PROGRAM OUTCOMES (POs)

1. Apply domain specific knowledge of computing and mathematics for designing of software solutions for defined problems and requirements.
2. Understand and analyze a problem and suggest feasible solutions.
3. Design, evaluate, and develop effective solutions for complex computing problems to meet desired needs.
4. Design and conduct experiments and use research-based methods to investigate complex computing problems.
5. Use appropriate techniques and software tools for computing activities.
6. Understand and commit to professional norms, regulations and ethics.
7. Recognize the need for and have the ability to engage in independent learning for continual professional development.
8. Understand and apply project management principles, as a member or leader in multidisciplinary environments.
9. Effectively communicate technical information, both oral and written with range of audience.
10. Analyze societal, environmental, cultural and legal issues within local and global contexts when providing software solutions.
11. Work as a member or leader in diverse teams in multidisciplinary environments.
12. Use Innovation and Entrepreneurship for creation of value and wealth.

RESEARCH PAPERS

1	Remote Patient Monitoring <i>Aakash Varma, Abhishek Singh, Neha Yadav guided by Ms. Sonu Gupta</i>	01
2	Vehicle Monitoring System Using Mobile Technology <i>Angha Anant Baikar , Reuel Domnic Cardozo guided by Ms. Neelam Naik</i>	03
3	Handling Large Terrain Data In GIS <i>Mukesh Koli, Akshay Mathur guided by Ms. Rashmi Vipat</i>	06
4	Neuroevolution Machine Learning Through Natural Selection <i>Clive Machado, Shashikant Mourya guided by Ms. Rashmi Vipat</i>	11
5	Data Mining for Security Applications <i>Neeraj Maurya guided by Ms. Puja Agarwal</i>	14
6	Enhancement Of Educational Experience Through Mobile Technology <i>Deepak Anthati, Neha Chauhan, Deep Singh Dangji guided by Ms. Neelam Naik</i>	17
7	A Survey of Free Space Optical Communication Network Channel over Optical Fiber Cable Communication <i>Prakash Mandhyan, Krishnakant Maurya , Rahul Maurya guided by Mr. Gajendra Dixit</i>	21
8	Application of Information Technology for the Education Of Dumb And Deaf Children <i>Himesh Jha, Satish Gupta, Rashmi Sharma guided by Ms. Neelam Naik</i>	27
9	Data Mining Techniques for Intrusion Detection <i>Jay Kotecha, Kruthika Hariharan guided by Ms. Rashmi Vipat</i>	31
10	Wireless Tracking System <i>Khedkar Apurva , Mhatre Shraddha, Mungekar Aishwarya guided by Ms. Rupali Jadhav</i>	36
11	Artificial Intelligence in Healthcare <i>Parikshit S Bhardwaj, Vivek Chaube, Sandeep Chauhan guided by Ms. Madhulika B.</i>	38
12	Cloud Computing and Database <i>Vikas Pandey guided by Ms. Puja Agarwal</i>	41
13	Text Mining in Healthcare <i>Mamta Patel , Vandana Patel guided by Ms. Puja Agarwal</i>	45
14	Digital Immortality <i>Protima Paul, Nikhil Pednekar guided by Mr. Pankaj Mudholkar</i>	51
15	Use Of Big Data In Social Networking <i>Abhinav Singh, Randhir Singh, Rakhi Singh guided by Mr. Pankaj Mudholkar</i>	54

ARTICLES

1	Role of Computer in Global Warming <i>Kumbhar Harish, Masdekar Mohanish, Mishra Manishkumar guided by Ms. Rupali Jadhav</i>	60
2	Introduction to Machine Learning <i>Trishul Mody, Rohit Maurya , Matsyendra Mishra guided by Ms. Priya Sinha</i>	63
3	A comparative Study: IoT Tools Raspberry Pi and Arduino use in Healthcare <i>Pankaj Singh, Ramratan Singh, Ajeet Kumar Yadav by Ms. Sonam Pareek</i>	68
4	Virtualization In Cloud Computing <i>Prakash singh, Anurag Singh, Ritesh Singh guided by Mr. Pankaj Mudholkar.</i>	71
5	Critical Study on E-Business Trends <i>Rajkumar Pathak, Pankaj Rajbhar guided by Mr. Pankaj Mudholkar</i>	76

Editorial

It gives us colossal pleasure to show the fourth volume of our Research Journal, Tech Tonics – TIMSCDR Research Journal. It is a collection of scholarly research papers and articles written by students of MCA (Masters in Computer Applications) course in the domains of Computer Science, Information Technology and Applications.

The Journal showcases the research endeavors of Post Graduate students. This helps them to understand IT industry problems either analytically or practically. These efforts will inculcate ability to think and elaborate new ideas amongst the students in the dynamic field of Information Technology.

The Journal presents papers and articles in various specializations in Information Technology like Mobile technology, Data Communication, Cloud Computing, Big Data and Social Networking, Digital Immortality, E-business, use of Information Technology in Global Warming, Wireless Technology, Data Mining, use of Text Mining in Health Care, Database Management System, Machine Learning, GIS, Intrusion Detection, Application of IT in the field of Education, Internet of Things(IOT), Artificial Intelligence and Remote Patient Monitoring using cloud.

To ensure the originality of research work, the research papers are thoroughly checked for plagiarism and selected for publication. Students of TIMSCDR get opportunity and exposure to the dynamic field of IT and are able to understand the relevance of research work through this academic exercise of performing research and presenting the same through research papers.

Finally, this Research Journal is a modest effort to encourage the young, enthusiastic and resourceful minds of the students to do research using latest techniques, and innovate and pen down emerging ideas in the field of Information Technology and Applications.

Editor

Dr. Vinita Gaikwad

I/c Director, TIMSCDR

RESEARCH PAPERS

Remote Patient Monitoring

Guided by: Ms. Sonu Gupta

Aakash Varma
TIMSCDR

Abhishek Singh
TIMSCDR

Neha Yadav
TIMSCDR

Abstract- In precise, Remote Patient Monitoring is an initiative by the modern medical science that allows examining patients without the primitive clinical prerequisite that enable easy access to cure and also decreases expense.

Keywords- *Personal health record, Benchmarks, Monitoring, Surveillance, Master Patient Index.*

I. INTRODUCTION

If RPM is implemented in treating disease that are severely chronic in nature it helps in improving the state of patient's life. RPM not only minimises expense but also helps in simplifying the process to treat patients.

A round the clock surveillance helps patients suffering from high-risks disease. RPM causes less readmissions and shorter stay in the hospitals. It also reduces the in person interaction between the doctor and patient and hence reducing the minimal chances of catching an infection.

Patient can be accessed 24x7 and if we include a cloud-based system of recording data than it becomes easier to have a brief view of the patient's medical history at a time.

RPM is also counted as an approach where the patient himself can at times if required perform the required medical check-ups upon them self and send the data to the doctor for reference at real time.

II. EARLY DETECTION

Early detection includes everyday monitoring devices like glucose meters for patients with diabetes or blood pressure monitors for patient receiving cardiac treatment. Mobile connectivity is also being vastly used for self detection however, it comes with its share of pros and cons. They can log into health care communities and gain detailed explanations about the disease they suffer from. Home systems can also help people to join online portals for discussion where anonymous people can share their views, information and experiences with

people. Better and quick health management will keep things affordable. Costly expenses are few to occur.

RPM is relatively cheap and new to the market in present. It not only implements new technology but also is creating new demands because of easy access and the interest capitalist are showing in this new gateway

III. BASIC STEP-BY-STEP PROCEDURE FOLLOWED

Phase 1:

- Activation: a person should nominate for passive data collection which is mandatory here.
- Obtaining data: the device being used for collecting data should store information for later viewing and delivery.
- Wrap data: data to be transmitted is properly packaged.

Phase 2:

- Delivery: an electronic medium is selected for transmission of the packaged data.
- Receipt: An authorised third party than accesses the data for further analysis.
- Indicators: benchmarks are made and available results are compared against them. Any abnormal activity observed is surely highlighted.

Phase 3:

- Observation and Review: the data made available is to be screened, compared and properly tabulated for future reference.
- Alert method: here, an alert is prepared that would be sent via phone or an email. Proper contact details are needed. It is hence necessary that genuine data be collected from the user for this phase.

Phase 4:

- Alert sent: Alert is sent to patient and also immediate people for the patient i.e. assistant, family, the technician.

Phase 5:

- Treat and Adjust: here, the doctors, family intervene in patient activity to provide assistance and treatment.
- Educate: The doctors also teach the family and immediate assistance at the patient's place of how to respond when similar situation arises in the future.

IV. PERSONAL HEALTH RECORD (PHR)

Technically it is an application used by the people undergoing treatment to maintain and manage their health information in a private, secure manner. It helps to store, share and track past information.

It is an approach that helps maintain information about health situations. It has an up-to-date medication list. It has a memorandum of all our previous health issues and treatment underwent. It has the ability to give proper details about the surgeries we underwent when and where. At times health history are an intrinsic to come to a conclusion and helps in identifying the root problem of the patient.

A PHR at times can be accessible from various devices and hence information can be readily available at any time.

V. MASTER PATIENT INDEX (MPI)

It is a database that carries details about every patient registered at any healthcare centre. It may also include information about the physicians as well as patient name, their contact details, birth place, date and also their medical history. It helps to give a clear picture of an individual patient and his details. MPI helps in efficient And accurate collection of data. And MPI also makes sure that every data is presented just once and also completely. They are many times easily accessible from PHR. MPI details are available in the hospitals where the patient has registered himself. And can be accessed by their personal doctors or the specialist they are taken care of under.

An EMPI i.e. Enterprise Master Patient Index is a larger network that connects several MPI's to each other and cover all rehabilitation facilities and

outpatient clinics and also includes all medical hospitals. EMPI has the capacity to merge (aggregate) information stored in separate sections in the same healthcare centre.

VI. PRESENT STATUS

RPM in general implementation faces some big issues in India. Senior citizen might not be well acquainted with use of modern technology. The devices used for collecting data from patient about their health issues are not standardized properly as IT still has a long way to go before we get a basic protocol-following system that manages health electronic records easily.

Another technical aspect is how the monitoring team actually puts to use the health data collected and make an observation that actually nails the issue that they suffer from. Alerts and benchmarks are supposed to be put on good use as any change in their values may put the patient and his life in great risk. Timely warning messages are also supposed to be on time.

VII. CONCLUSION

However, investors who are encouraging these initiatives are pretty sure that this endeavour with time will be a good success as there is enough scope and a big need for an automated system to exist in healthcare system.

RPM in general at this point is quite in its early stages. With time it'll surely be beneficial to the masses as it would not only make the work paper-free but also making appointments, regular reminders very easy for people.

ACKNOWLEDGMENT

It gives us immense pleasure to present this Research Paper. We grab this opportunity to express our heartfelt obligation towards the people without whom completion of this Research Paper have not been possible. We would like to thank our guide Prof. Sonu Gupta for her immense support and continuous encouragement to our team.

REFERENCES

- [1] <http://searchhealthit.techtarget.com/definition/remote-patient-monitoring-RPM>
- [2] http://www.corp.att.com/healthcare/remote_patient_monitoring/
- [3] <http://www.informationweek.com/mobile/remote-patient-monitoring-9-promising-technologies/d-d-id/1110968?>
- [4] https://www.myphr.com/StartaPHR/what_is_a_phr.a_spx

Vehicle Monitoring System using Mobile Technology

Guided by: Neelam Naik

Angha Anant Baikar
TIMSCDR

Reuel Domnic Cardozo
TIMSCDR

Abstract—Important efforts have been made to develop versatile management ways to support the precise options of transport systems. Here the main focus is on the study for the implementation of a true Time Bus Management System (RTBMS), by putting in GPS devices on town buses. The important Time Bus Management System could be a specific system designed to show the period locations of the general public buses or the other transportation company supplier within the town. It'll alter the trailing devices to get GPS information of the bus locations, that it'll then transfer it to centralized management unit and depict it by activating representational process of buses within the approximate geographic positions on the route map. Specific softwares are needed to interface the information received to the map. The study reveals the precise or most closely correct location of the Bus. The instructed models are the mobile Application on those users will track the bus location and internet primarily based backside system for Administrations to manage the whole system.

I. INTRODUCTION

The movement of vehicles is full of totally different unsure conditions as the day progresses, like traffic jam, surprising delays, and randomness in traveller demand incidents and irregular vehicle-dispatching times. Several passengers are typically late to figure, students are late for categories as a result of the arrangement to expect the bus rather than simply merely employing an alternate transportation.

The enhanced system showing the bus location and also the bus point at bus stops may cut back the anxiety of passengers looking forward to the bus. Dispersive point information through different interfaces like sensible phone may create the general public transit system additional easy and therefore increase its fight among varied transportation modes. With the appearance of GPS and cellular network, real time vehicle following for higher transport management has become doable. These technologies will be applied to conveyance systems, particularly buses, that don't seem to be ready to adhere to predefined timetables due to reasons like traffic jams, breakdowns etc. The exaggerated waiting time and also the uncertainty in bus arrival create conveyance system unattractive for passengers. Real Time Bus

Management system uses a spread of technologies to trace the locations of buses in real time and uses this information to come up with predictions of bus arrivals at stops on the route. Once this information is disseminated to passengers by wired or wireless media, they'll pay their time expeditiously and reach the stop simply before the bus arrives, or take alternate if the bus is delayed. They'll even arrange their journeys long before they really undertake them. This can create the general public transport system competitive and passenger- friendly. The employment of personal vehicles is reduced once additional folks use transportation vehicles that successively reduce traffic and pollution.

II. EXISTING SYSTEMS

The bus is associated with Automatic Vehicle Location (AVL) system to enhance London's buses exploitation technology put in by Siemens conductor. The system tracks all of London's 8000 buses to produce passengers with audio visual announcements, improved data on bus arrivals, and to trigger priority at traffic junctions. There is conjointly a live map to trace the bus locations [1].

III. LITERATURE SURVEY

Three algorithms were introduced with totally different assumptions on computer file and were shown to beat many algorithms from the literature. Their algorithms, however, didn't contemplate the impact of tie up and dwell time at bus stations. Kidwell [3] bestowed a rule for predicting bus arrival times supported period vehicle location. The rule worked by dividing every route into zones and recording the time that every bus tried and tries each zone. Predictions were supported the foremost recent observation of a bus passing through every zone. However, this rule wasn't appropriate for big cities wherever each period of time and dwell time may well be subject to large variations. Usually speaking, these models measures reliable only when the path within the space of interest is comparatively stable. One among their main limitations is that it needs an in depth set of historical knowledge, which cannot be offered in

application particularly once the path varies considerably over time.

IV. PROBLEM DEFINITION

A considerable quantity of cash is spent on IT-based applications like period, at-stop displays on transport, however actual data concerning the activity affects these customers or potential customers in reality is kind of distributed. This paper [2] focuses specially on user response to transport data via phone, mobile devices, and the web and at-stop displays. Variety of studies is initiated within the past to deal with the bus time of arrival prediction downside.

Vehicle trailing has become therefore advanced and straightforward with varied approaching technologies. However the price effectiveness and its implementation became high. The matter lies in maintaining high grade servers to backup the info for vehicle trailing and therefore the use of web to trace these vehicles. Though real time trailing of vehicles through web offers Associate in nursing overall advantage in vehicle trailing, the system fails to produce fewer prices of investment and maintenance. The system uses an indoor info to appear in the placement of the vehicle with the GPS coordinates received. Therefore it does not use the GPRS and TCP/IP protocol for web for information transfer between servers and therefore the user device. Therefore this reduces the price of the trailing system. Conjointly the placement details are sent quickly through SMS via the GSM electronic equipment.

V. IMPLEMENTATION OF RTBMS

The entire system can accommodate GPS trailing and backend server, User Application (Android Based) and Administrator's management Panel.

A. GPS Trailing and Backend Server

This will accommodate the complete GPS and GPRS, location trailing system, which can obtain this location of the bus, the speed it's moving, the number of times that was at a still position, direction and a couple of different details and store every information on a frenzied server .

1. Maintain the information of all the routes, the buses that ply on a route, the stops on every route etc.
2. To incessantly receive location and speed from the vehicle units of all the buses.

B. User Application (Android Based)

A mobile application that users will use to achieve these services,

1. Get current position of a definite bus.
2. To search for running buses for his journey.
3. To find out the bus point to the stop and calculable time to reach his destination.

These applications are as user friendly as potential for even users World Health Organization does not seem to be technical tech savvy.

C. Administrator's Control Board

1. Directors will amend the system settings in keeping with the new changes or updated.
2. Complete access to all or any modules, information and algorithms may be changed with the administrator's panel. [3]

VI. IMPORTANCE OF RTBMS

By meticulously following every bus, the user will gain complete management over the complete fleet, considerably reducing prices, up safety and boosting productivity.

A. Reducing fuel costs

Fuel usage is clearly inevitable expense for a public utility, whereas such an expense can't be fully eliminated, it will actually be reduced. A recent study by the Aberdeen cluster terminated that GPS vehicle following systems have helped firms scale back their fuel prices by a mean of 13.2%. Virtually everything monitored by GPS vehicle following will work to assist bring down fuel expenditures for a corporation. For instance, dashing may be a Brooding angina think about high fuel usage. A GPS following answer will give speed info and send alerts once a vehicle exceeds the regulation. Excessive inactivity may also contribute to high fuel prices. A GPS fleet following system will facilitate scale back inactivity times by alerting the user whenever a vehicle idles longer than a gift period. Conserving fuel has alternative edges besides saving your company valuable cash. With the continued considerations concerning depleting oil resources and also the environmental impact of vehicles that admit them, a public utility that's actively scaling down on its fuel consumption will even be thought of to be "going inexperienced." That's a key PR angle that ought to charm to the general public at giant, and should even garner raised funds from native governments.

B. Improving client service

It's a reasonably straightforward equation: once rider's area unit assured within the performance of the buses they use, they're happier concerning victimisation them. By victimisation the tools in GPS fleet following, a public utility will actually serve its purchasers higher. Getting reports of a vehicle's route will facilitate the user track arrivals and departures. By providing correct and verifiable proof of services, it's a valuable tool for reviewing routes and frequency of stops, which may facilitate organization contour operations and run a lot of with efficiency. Depending on the community you serve, some aspects of a fleet following system

may also be shared with the general public that depends on your buses. For instance, in Feb 2011 Lehigh University in Bethlehem, Pa., launched a bus-tracking web site that provides students info on bus schedules and locations, and even permits them to attend in an exceedingly safe and environment-controlled building till they see their bus approach their location. This helps students keep out of the rain, snow and alternative harsh environments.[4]

C. Increasing security and safety

GPS fleet following may also facilitate improve security and safety for drivers and passengers alike. Ought to a bus go missing for no matter reason, GPS will facilitate an organization recover it and any on-board instrumentality quickly. Some systems even provide options that send associate alert if a vehicle is employed throughout off-hours. If the bus burst its route at any time, the system can understand. And within the unlikely — however not unregistered — event that a bus has been purloined, there isn't a quicker thanks to track it down. Of course, correct maintenance is vital to make sure that buses area unit compliant with safety pointers and up to now on repair schedules. Some GPS fleet following systems feature alerting functions to let the user understand once vehicles area unit due for routine maintenance or alternative scheduled appointments. This could greatly facilitate scale back repair prices and alternative fees. And again, with the increasing concern. keeping vehicles properly maintained can scale back the chance of harmful emissions and fulfil alternative inexperienced efforts.[4]

D. Monitoring driver behaviour

While not a doubt, any transportation organization is barely as good because the drivers it employs. As mentioned higher than, a GPS fleet following system will relay a spread of knowledge relating to driving behaviour, together with dashing, inactivity time and unauthorized vehicle use. Besides encouraging drivers to behave a lot of responsibly on the task, it may also shield them within the event of associate unwarranted criticism. The reports in an exceedingly GPS fleet following system will prove, for instance, that the motive force was so following native speed laws, created stops at the appointed times and stuck to the established route. This could be a profit not just for mass transit, however is vital for buses that often transport under-age students.

VII. SUGGESTION

One can predict bus arrival times that supported period and vehicle location. This may be worked by dividing every route into zones and recording the time that every bus gone through each zone. Predictions were supported the foremost recent observation of a bus passing through every zone. We are able to conjointly predicting the traffic location supported bus time of arrival. RTBMS is beneficial for safety purpose in class bus. Thus one can be able to track bus.

VIII. CONCLUSION

The RTBMS tracks the location of all the buses and estimates their time of arrival at completely different stops in their individual routes. Estimates area unit updated each time the bus sends associate update. It distributes this data to passengers' mistreatment show terminals at bus stops; net primarily interface and sensible phone application. This analysis serves the wants of passengers, vehicle drivers and directors of the transport system. With the appearance of GPS and also the omnipresent cellular network, real time vehicle trailing for higher transport management has become potential.

REFERENCES

- [1] <http://traintimes.org.uk/map/tube/>
- [2] Ganesh K and Thirvikraman M, Net Logic Semiconductors Pvt. Ltd., Bangalore "Implementation of a real time passenger information system".
- [3] Lin, W.-H. and J. Zeng. "Experimental Study on Real-Time Bus Arrival Time Prediction with GPS Data". In Transportation Research Record: Journal of the Transportation Research Board, No. 1666, TRB, National Research Council, Washington, D.C., 1999, pp1019.
- [4] Kidwell, B. "Predicting Transit Vehicle Arrival Times". GeoGraphics Laboratory, Bridgewater State Colleg, Bridgewater, Mass., 2001.

Handling Large Terrain Data In GIS

Guided by: Ms. Rashmi Vipat

Mukesh Koli
TIMSCDR

Akshay Mathur
TIMSCDR

Abstract: This paper presents an exploration and development project which will give an extension to second geo-databases for handling massive parcel of land knowledge. It first discusses application needs and system style, and then elaborates system design for optimum knowledge organization and change, economical multi-resolution queries, and dynamic DTM generation. It then addresses technical problems associated with knowledge storage, seamless covering, vertical classification, and DTM generalization. Finally, it discusses the restrictions and shortcomings of the current approach, and identifies future analysis and development tasks.

Keywords: - *Intrusion Detection System, Anomaly Detection, Misuse Detection, Data mining, Clustering, Classifications*

I. INTRODUCTION

Many GIS comes, particularly wide and nationwide ones, often have to be compelled to store and manage massive parcel knowledge. Even small scale projects might ought to wear down an outsized quantity of parcel data, thanks to recently accessible knowledge acquisition techniques like LiDAR. Such knowledge may be many tera-bytes in size, or may contain billions of activity points. While most of today's enterprise geo-databases (such as SDE) are capable of handling giant second information, parcel of land information have brought new necessities and challenges. These embrace 1) a way to integrate parcel of land information with second information, 2) what arrangement to use, and 3) a way to support high performance multi-resolution spacial queries and update. Given the actual fact that TIN and GRID area unit the foremost common information formats in digital piece of ground modeling, it's necessary to look at if they are the most effective selections for storing piece of ground information. Because different applications could need information totally different of various special resolutions looking on underlying abstract models (Peng, 2000, 1997), multi-resolution queries have become a lot of and more vital subject in GIS. Some applications could even require a alleged "horizontal" multi-resolution question that specifies totally different levels of vertical resolutions for various parts of a study

space (Kinder et al., 2000). Typical examples include landscape coming up with 3D flight simulation, where the center of interest typically needs higher resolution information, while the rest of the world solely needs information of coarser resolutions.

To address of these problems, and others, a replacement analysis and development project has been enforced at ESRI to produce an extension to current second geo-databases for handling giant terrain information. The remaining of the paper elaborates the planning idea and system design, and addresses connected technical problems. Finally, it provides a top level view for any analysis and development.

II. DESIGN CONSIDERATIONS AND SYSTEM ARCHITECTURE

The design can be boiled down to three aspects:

- 1) what to store;
- 2) where to store it; and
- 3) how to store it..

A. What to Store?

Typically, supply piece of ground knowledge embrace 1) mensuration points (e.g., spot height points like measuring system data), 2) contours, and 3) structure lines (or break lines) that capture the separation of piece of ground and different necessary geophysical science and geographic features. As a result, the set of individual points, contours, and break lines, doesn't represent a decent (continuous) piece of ground representation in a very digital atmosphere (Peng et al., 1996), they are not typically directly used for surface visual image and analysis in GIS. Instead, a typical GIS would build a digital terrain model (DTM), victimising these knowledge, and do analysis based on the DTM. Thanks to this, individuals usually store and manipulate their piece of ground information directly as a DTM, disregarding the supply knowledge.

A DTM might take the shape of a GRID or TIN (Triangulated Irregular Network). Spacial resolution of a GRID DTM is inherently unnatural

to cell size – the smaller the cell size, the higher the resolution, excluding the standard of the initial data. However, once generated, the supply information square measure lost and no improvement is feasible. One will solely down-sample a GRID DTM (i.e., attend a bigger cell size and, thus, lower resolution). Creating a brand new DTM of a smaller cell size out of associate degree existing GRID DTM won't increase its spacial resolution. A TIN DTM, on the opposite hand, doesn't suffer from this constraint due to its adaptative nature, though a little elevation tolerance may be utilized to cut back information amount in constructing a TIN. Many giant information suppliers (USGS, for instance) select GRID for their piece of land information, attributable to its simplicity and comparatively little storage size. TIN is usually employed in places wherever engineering precision is needed. Thanks to its refined structure and heavy overhead in storage (in order to stay topology), TIN is rarely accustomed offer and maintain an outsized quantity of piece of land data.

B. Where to Store?

GIS applications usually need a information atmosphere that supports (among others) 1) geometric and thematic description of abstraction objects, 2) topological relationships at a geometrical primitive level and object level, 3) versioning, 4) multi-user access and redaction, and 5) seamless, scalable, multi-resolution, and high performance abstraction queries. The approach delineated relies on ESRI's geo-database framework outlined in ArcGIS (Zeiler, 1999), because it has the potential to fulfill the necessities listed. As shown in Figure 1, a geo-database (or info for short) contains one or more feature datasets; a feature dataset contains one or a lot of feature categories; a feature class contains one or a lot of options of the same geometric sort (point, line, or area). A feature dataset defines a abstract entity for those feature categories that share the same abstraction reference, cowl identical geographic extent, and are usually thematically associated with one another. Under this framework, it is clear that supply parcel of land knowledge ought to be sorted into totally different feature categories, in line with their geometric sort, source, and thematic description. Typical examples embrace evacuation systems, mass points (e.g., LiDAR data), road networks, ridgelines, and communication system points. Those feature categories that contribute to identical ground space ar then place into identical feature dataset.

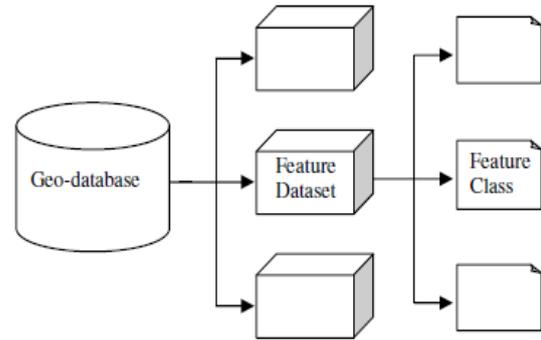


Figure 1: A geo-database structure.

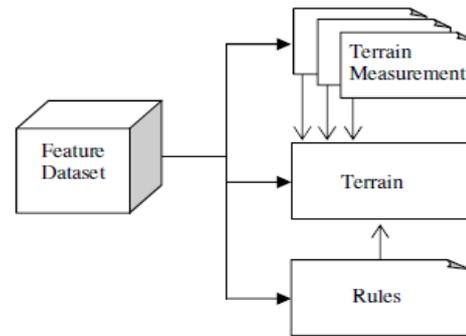


Figure 2: Terrain in geo-database.

C. How to Store?

Feature categories of tract knowledge don't themselves have abundant meaning in tract application. A better level of abstraction is necessary so as to support users to properly model their terrain knowledge and applications. This can be achieved by introducing Terrain into the geo-database framework delineated in section 2.2. A tract is outlined as a special sort living within a feature dataset, consisting of one or additional feature categories among the dataset and a group of rules (Figure 2). Feature categories constituting a tract dataset area unit known as tract measurements. Terrain measurements will be kept as purpose, line, or area feature categories. They will participate in additional than one tract within a similar feature dataset. Tract measurements also are regular feature categories that may be used for different functions, allowing secnd knowledge and tract knowledge to be integrated into a similar database.

The set of rules defines however the measurements area unit applied once constructing a piece of ground illustration (such as a DTM) as a result of spacial question. These rules 1) outline the role of every measurement, whether or not a mensuration ought to be adscititious as a mass elevation purpose, a chance line, a replacement plane figure, or a clipping plane figure, and then forth; 2) specify, for each measurement, wherever the elevation data comes from, whether the Z

coordinate, the worth of a specific attribute, or a result interpolated from a given DTM (which might embrace the one beneath construction); and 3) specify at what resolution, or resolution vary, a mensuration ought to participate. This is necessary so as to support multi-resolution queries and accommodate completely different necessities in piece of ground generalization (to be mentioned in section three.2). The introduction of piece of ground permits a geo-database to store and manage piece of ground knowledge while not being guaranteed to a specific sort of DTM. A DTM may be generated dynamically upon user's request by applying the foundations. the thought of storing mensuration data (rather than express DTMs) in an exceedingly knowledgebase is comparable to the philosophy of Implicit TIN (Kidner et al., 2000).

Terrain provides a substantive and comprehensive entity through that users manage, query, and apply their tract relief data. Such associate degree entity is stated in ArcGIS as terrain dataset. In order to support varied applications, variety of basic requirements are known that tract ought to support. It is also assumed that a DTM continues to be the foremost favorable structure for surface image and analysis in GIS. Therefore, a spatial question on a tract is anticipated to end in a DTM, upon which varied analyses is performed. These necessities include:

- Support an oversized space extent and an oversized quantity of information.
- Support purpose, line, and space knowledge.
- Support update, on each measurements and rules enable sure measurements to be included/excluded in a spatial question.
- Support TIN and GRID DTM output.
- Support spatial question with reference to a given space of interest and vertical resolution. This might need Terrain to dynamically generate DTMs of given resolutions, anyplace at intervals the extent of the tract.
- Support “horizontal” multi-resolution question a special quite question that specifies completely different vertical resolutions for various components of a given space. The query can end in a multi-resolution DTM during which vertical resolution varies across the entire space

III. KEY TECHNICAL ISSUES IN SYSTEM DEVELOPMENT

To implement the supports required, two key technical problems have to be addressed compulsory:

- A. Tilting
- B. DataUpdating.

Although today's sensible geo-databases area unit capable of handling large amounts of knowledge, and second abstraction assortment is essentially a built-in feature, further arrangements area unit still necessary so as to support quick DTM creation, multi-resolution queries, and DTM generalization. These embrace internally arrangement measurement knowledge into tiles in keeping with knowledge extent, density, and hardware/software constraints; and introducing further indexing within the vertical dimension.

A.Tiling

Since information area unit is probably large, it's not possible to handle all data directly. Not solely is memory a drag, however performance can be intolerably poor. Tiling, on the opposite hand, can be a good “divide and conquer” approach for handling information of enormous extent, given the constraints of today's offered technology. A good covering theme may end up in abstraction coherence – information area unit organized and hold on in step with their abstraction proximity, thus increasing the performance in abstraction question and information transfer. It also provides a strong (2D indexing) mechanism for quick searching; permits information to be handled in an exceedingly manageable form; permits memory and central processor intensive tasks to be performed locally while not paralyzing the system; and is crucial for DTM generalization (to be mentioned later). Tiling divides an oversized geographic region into smaller, more manageable, units (Figure 3), which might have totally different forms. This approach uses a daily parallelogram tile for its simplicity and efficiency in computation. Selecting a correct tile size could be a bit more difficult. It depends on information density, CPU speed, available memory, and alternative issues. Basically, the size must not be too huge, so a full resolution DTM of any tile can be generated suitable quantity of system resources.

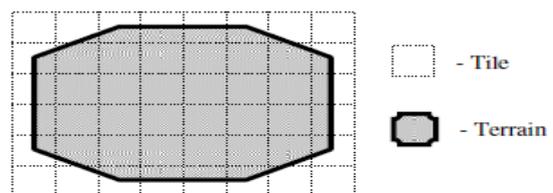


Figure 3: Divide large terrain extent into regular tiles.

Tiles are used as the basis for reorganizing points. Points falling into the same tile can be grouped into, and stored and handled as, one single entity – a so called multi-point. This will reduce storage space, increase spatial coherence and access speed,

and reduce disk I/O and network traffic. As most of the terrain data will be mass points coming from remote sensing and laser scanning, the benefit of this process can be significant. Line and area features may require extra work in order to benefit from this measure, as a single feature can cross many tiles.

B. Updating Data

Requirements for piece of land update come back from two aspects: the measurements, and therefore the rules. Any changes concerning these two will require the interior vertical classification to be updated. Because rules area unit personal to the piece of land dataset, change rules is simple easy. Measurements, on the opposite hand, are shared by alternative applications, and might be changed while not going through piece of land datasets. So as to stay piece of land datasets and measurements in adjust, some mechanisms area unit needed that keep the datasets familiar whenever Associate in Nursing update is performed on the measurements. this can be done through Events and Invalidated-Area. Associate in Nursing Invalidated-Area could be a region wherever changes of measurements have occurred. It permits Associate in Nursing noncurrent piece of land dataset to be updated regionally. When Associate in Nursing update to a mensuration is committed, a happening is broadcast. Those piece of land datasets that area unit affected can update their Invalidated-Areas upon receiving the Event. Users will then decide once to update the affected piece of land dataset.

IV. APPLICATION EXAMPLES – SPATIAL QUERY AND SURFACE ANALYSIS

Spatial question and surface analysis square measure Terrain’s two most important applications. A typical spatial question takes a locality of interest AOI and a (relative) vertical resolution DH, and outputs a TIN or GRID DTM (specified by the user, Figure 4). The output are often associated (transient) object that may be persisted as long as requested by the user. Space of interest AOI might contain multiple regions. During this case, there’ll be an inventory of DHs, each of which corresponds to a section in AOI. A multi-region AOI can result in a multi-resolution (continuous) DTM, whereas a single region AOI can manufacture a single-resolution one. With all the indexing support, the system will quickly apportion those multipoints that contribute to DH. However also are among the question space AOI. Line and space measurements can even be quickly known.

Obviously, it’s a waste to use all the points once zoomed to full extent, as several of them could also be mapped onto a similar pixels of the screen.

During this case, a well-calculated, simplified version of the DTM might serve to produce an honest summary of the tract. This additionally reduces the time utilized in DTM generation and rendering.

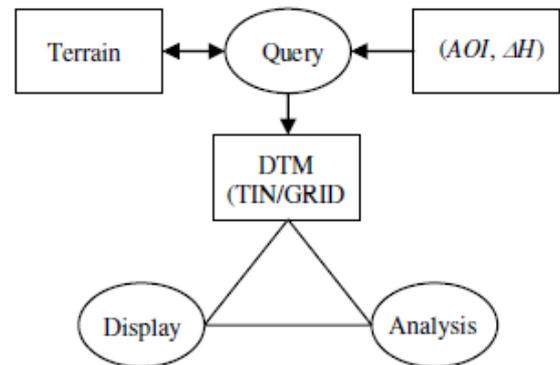


Figure 4: Examples of Terrain application.

V. CONCLUSION

This paper has given associate economical approach for GIS users to handle massive piece of ground knowledge and model surface applications. Because solely measurements and rules area unit hold on an exceedingly information, users will take the benefits of TIN and GRID structures without sacrificing storage or losing data. The tiling scheme makes it attainable to perform large-scale tasks that require acting on a DTM of high resolution. It conjointly helps to achieve spatial coherence, therefore dashing up spatial queries, and reducing disk I/O and network traffic. Vertical compartmentalisation provides another contribution to any speed up spatial queries.

Storing piece of ground knowledge as feature categories in an exceedingly feature dataset permits them to be integrated with second knowledge and be shared by alternative applications, like Topology and Geometric-Network (Zeiler, 1999). The ESRI geo-database framework additionally provides a foundation to implement versioning and multi-user access support. The Invalidated-Area mechanism provides AN economical and elegant thanks to discover changes to parcel of land measurements, and enables native change. The Massachusetts example given during this paper, and different in house testing cases, have incontestable the aptitude and efficiency of the projected approach. The approach may additionally benefit the user WHO will currently contract out for the supply measurement information and handle it in an economical and versatile manner for several applications, as opposition catching out a GRID model that’s restricted. There are, however,

many tasks that also need additional research and development. These embody 1) a much better vertical indexing mechanism for line and space options, 2) on-the-fly 3D generalization of line and space options, and 3) “horizontal” multi-resolution queries involving line and space options. The current style and implementation of the system has introduced some storage overhead so as to support covering and vertical categorisation. Efforts are required so as to attenuate or eliminate the impact of this overhead.

REFERENCES

- [1] Gottshalk, H.J., “Die Generalisierung von Isolinen als Ergebnis der Generalisierung von Flächen”, *Zeitschrift für Vermessungswesen*, Vol. 97, No. 11, pp. 489-494. (1972).
- [2] Heller, M., “Triangulation Algorithms for Adaptive Terrain Modeling”, proceedings of the *Fourth International Symposium on Spatial Data Handling*, Zurich, Vol. 1, pp. 163-174. (1990).
- [3] Hoppe, H., “Smooth View-Dependent Level-of-Detail Control and its Application to Terrain Rendering”, *IEEE Visualization*, October 1998, pp. 35-42. (1998).
- [4] Johns, C.B. and Abraham, I.M., “Line Generalization in a Global Cartographic Database”, *Cartographica* 24: pp. 32-45. (1987)
- [5] Kalvin, A.D., “Superfaces: Polygonal Mesh Simplification with Bounded Error”, *IEEE Computer Graphics and Applications*, May 1996, pp. 64-77. (1996)
- [6] Peng, W., Pilouk, M., Tempfli, K., “Generalizing Relief Representation Using Digital Contours”, proceedings of the *XVIII ISPRS Congress*, July 9-19, 1996, Vienna, Austria, 8p.(1996)
- [7] Reinhard, K., “Multiresolution Representations for Surfaces/Meshes Based on the Vertex Decimation Method”, *Computer & Graphics*, Vol. 22, No. 1, pp. 13-26. (1998)
- [8] Weibel, R., “Models and Experiments for Adaptive Computer-Assisted Terrain Generalization”, *Cartography and Geographic Information Systems*, Vol. 19, No. 3, pp. 133-152. (1992)
- [9] Weibel, R., “Konzepte und Experiment zur Automatisierung der Reliefgeneralisierung”, *Doctoral dissertation*, Geo-Processing Series, Vol. 15, Department of Geography, University of Zurich. (1989)
- [10] Wolf, G.W., “Generalisierung topographischer Karten mittels Oberflächengraphen”, *Doctoral dissertation*, Department of Geography, University of Klagenfurt. (1988)
- [11] Wu, H., “Prinzip und Method der automatischen Generalisierung der Reliefformen”, *Nachrichten aus dem Karten-und Vermessungswesen*, series I, Vol. 85, pp. 163-174. (1981)
- [12] Yoeli, P., “Entwurf einer Methodologie für computergestütztes kartographisches Generalisieren topographischer Reliefs”, *Kartographisches Generalisieren*, *Swiss Society of Cartography* (ed.), (Cartographic Publication Series, Vol. 10, Zurich. (1990)
- [13] Zeiler, M., “Modeling Our World – The ESRI Guide to Geodatabase Design”, *ESRI Press*. (1999)

Neuroevolution

Machine Learning Through Natural Selection

Guided by: Ms. Rashmi Vipat

Clive Machado
TIMSCDR

Shashikant Mourya
TIMSCDR

Abstract: This research paper focuses on the idea of Neuroevolution (NE). In this technique unlike other the machine learns through repeated mistakes and memorizing it. The results through applying NE can be proven on computer games already and the neural network formed can be visualized on the screen itself. Neuroevolution later can be implemented more widely where it can mimic the actual process of the human brain and the neural activity but for now due to computational limits it is commonly applied in computer games or robotics. The neural network formed can be complicated so this approach needs to also be carefully adhered to since humans cannot predict what a particular AI would evolve into.

Keywords: Neural, encoding, population, neuroevolution, feedback

I. INTRODUCTION

Neuroevolution (NE) is an artificial evolution of neural networks using generic algorithms which mean it mimics the process of Natural selection. In this technique unlike other, the machine learns through repeated mistakes and memorizing it similar to a human being although faster. It is most commonly applied in artificial life, evolutionary robotics, and computer games. Benefits of Neuroevolution are that it can be applied more widely than supervised learning algorithms, which require correct input-output.

II. HOW A HUMAN BRAIN WORKS

Basically, a human brain is made up of millions of neurons, also known as a neuron or nerve cell, connected to each other. It is an electrically excitable cell processing and transmitting information through chemical and electrical signals.

In short, when you memorize something your brain does not have a directory to store it like some files or folders in a computer. On the contrary, a human brain makes new connections of these nerve cells every time you memorize something.

III. HOW A NEURO EVOLVED AI WOULD WORK

Similar to the concept of human brains, the machine will work in iterations and find out faults in every iteration and register these faults in its memory so that next time it has evolved to avoid these faults.

IV. BENEFITS OF NEUROEVOLUTION

Neuroevolution is very promising in complex reinforcement learning. Neuroevolution identifies through the behavioral for a network that performs well at a given task. This approach for solving many complex problems represents a different approach to normal statistical techniques that attempt to estimate the objective of a particular action.

Past studies conducted have also proven that Neuroevolution is more efficient and faster method. Reinforcement learning similar to Neuroevolution allows the machine to learn via behavior based on the feedback from the environment around it.

Since Neuroevolution searches for certain behaviors instead of a particular value function, it is effective in problems which change overtime.

V. TWEANN ENCODING

The question for all TWEANNs of how to encode networks using an efficient genetic representation must be addressed. Here several prototypical representational schemes will be discussed. TWEANNs can be segregated between those that use an indirect encoding, and those that use a direct one.

Employed by most TWEANNs, Direct encoding schemes specify every connection and node in the genome that will appear in the phenotype.

On the contrary, indirect encodings specify rules only for phenotype construction. Those rules can be growth rules or layer specifications through cell division. Direct encoding does not allow a more compact representation than indirect encoding, thus every connection and node are not specified in the genome, although they can be derived from it.

VI. INITIAL POPULATIONS AND TOPOLOGICAL INNOVATION

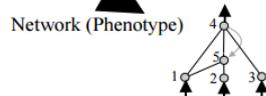
A lot of TWEANN systems, the starting population is a random collection of topologies. Such type of a population makes sure topological diversity from the start. Although, random starting populations turn out to produce varied issues for TWEANNs.

For example, many direct encoding schemes, there is a possibility that a network will have no path from its inputs to its outputs. Such inconceivable networks take time to get out of the population. Although, there is a more silent but more dangerous problem with starting randomly. It is considerable to evolve minimal solutions; in that manner, the number of parameters which have to be searched is reduced.

Initializing with random topologies may not lead to finding minimal solutions, due to the population which starts out with a lot of unnecessary nodes and connections already present. Not many of these nodes or connections withstood a single evaluation, which means there is no justifying their configuration. Any minimization of networks will have to be spent getting rid of apparatus which should have not been there in the first place, also nothing in this process of recombining different topologies launches towards such minimization. Since there are no fitness cost in creating larger networks, those fitness cost will overpower as long as they have high fitness. One solution to force minimal topologies is to include network size into the fitness function, and some of those TWEANNs actually apply this.

In such methods, larger networks have their fitness's penalized.

Genome (Genotype)								
Node	Node 1	Node 2	Node 3	Node 4	Node 5			
Genes	Sensor	Sensor	Sensor	Output	Hidden			
Connect.	In 1	In 2	In 3	In 2	In 5	In 1	In 4	
Genes	Out 4	Out 4	Out 4	Out 5	Out 4	Out 5	Out 5	
	Weight 0.7	Weight 0.5	Weight 0.5	Weight 0.2	Weight 0.4	Weight 0.6	Weight 0.6	
	Enabled	DISABLED	Enabled	Enabled	Enabled	Enabled	Enabled	
	Innov 1	Innov 2	Innov 3	Innov 4	Innov 5	Innov 6	Innov 11	



The above example is a genotype to phenotype mapping example. A genotype is shown that produces the depicted phenotype. 3 input nodes are present, one output node and one hidden node, and seven connections defined, from which one is recurrent. The second gene is not enabled; hence the connection that it describes (between nodes 2 and 4) is actually not expressed in that phenotype. Smaller networks in this way can encourage fitness function, it is not easy to know how much large the penalty must be for any particular network size, because different problems could have quite different topological requirements.

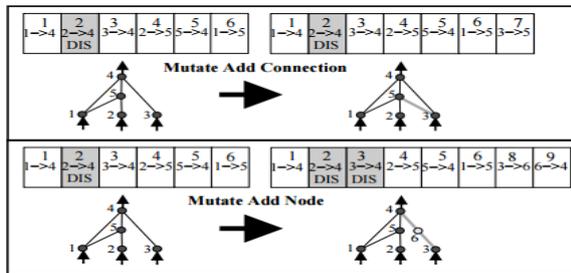
Modifying the fitness function could cause evolution to perform unfamiliarly than what the designer of the unmodified original fitness function intended to. A different way could be that for the method of Neuroevolution could tend towards minimalist. If there are no beginnings in hidden nodes of the population and grows structure while it benefits the solution, there is no need for fitness modification to minimize networks.

Hence, initializing with a growing structure and minimal population and from there is a design principle mentioned in NEAT. By beginning minimally, NEAT makes sure that the system searches on the solution in the lowest-dimensional weight space and its possibility over the course of all generations. Hence, the target is not to minimize only the final product and also all intermediate networks along the route as well.

This concept is the key to access and gain an advantage from the evolution of topology: it brings in the possibility for us to allow and minimize the search space, which can result in dramatic performance boosts.

One possibility for the TWEANNs and as to why they do not start out minimally is that without topological diversities presence in the starting population, topological innovations might not survive. The issue of protecting innovation is not adhered by these methods; hence networks with major structural additions are more likely not to reproduce. Thus, the evolutionary process through which new biological species arise, the population does not disable starting minimally in NEAT.

VII. WHAT IS NEUROEVOLUTION OF AUGMENTING TOPOLOGIES (NEAT)



Both the types add a connection and add a node, are demonstrated above with the connection genes of a network displayed above their phenotypes.

The number at the top in each genome of that gene is the innovation number. The original historical ancestor of each gene identified via the historical markers also known as the innovation number. New genes are assigned new increasingly higher numbers.

While adding a connection, a single new connection gene is added at the end of the genome and provided with the next available innovation number. Also while adding a new node, the connection gene is being split and also disabled, also two new connection genes are added at the end of the genome. The new node is between the two new connections. A new node gene (not depicted) representing this new node is added to the genome as well.

VIII. WHY IS IT STILL NOT IMPLEMENTED ON A LARGE SCALE

The reason being obvious, since this being a neuroevolution where the machine learns on its own humans still have to understand the complex behaviors it might produce.

Also the capacity to have more complex neural network similar to a human brain is still not possible due to the factor that our current technology cannot compute the large quantities of neural networks or patterns it generates.

IX. CONCLUSION

NE is a very powerful technology for particular sequential decision tasks

- Neural nets and evolutionary computation is a good match
- Adapts itself to many extensions
- Powerful in applications

Easy adaptation to applications

- Artificial life, robotics
- Control, biology
- Gaming, entertainment, optimization, training

Applicable in many future work opportunities

- Indirect encodings
- Inter action and evolution
- Knowledge, Learning, novelty

REFERENCES

- [1] <http://nn.cs.utexas.edu/downloads/papers/stanley.ec02.pdf>
- [2] <https://www.youtube.com/watch?v=qv6UVOQ0F44>
- [3] <http://pastebin.com/ZZmSNaHX#>
- [4] <https://en.wikipedia.org/wiki/Neuroev...>
- [5] <https://en.wikipedia.org/wiki/Evoluti...>
- [6] <https://en.wikipedia.org/wiki/Artific...>
- [7] <http://tasvideos.org/BizHawk.html>

Data Mining for Security Applications

Guided by: Ms.Puja Agarwal

Neeraj Maurya

TIMSCDR

Abstract: Database mining can be described as the strategy of burrowing for certain, once in the past unidentified and conceivably critical information from appallingly colossal databases by compelling learning revelation techniques. In this paper we focuses on key online security and security issues and concerns, the a portion of self-direction and the customer on security and security protections, data certification laws, regulatory examples, and the outlook for assurance and security authorization. Data mining strategies that we have successfully associated for advanced security. These applications fuse yet are not limit to noxious code area by mining parallel executable, framework interference recognizable proof by mining framework development, eccentricity acknowledgment, and data stream mining.

Keyword: Security, Threats, terrorism, theft

I. INTRODUCTION

The expansive measure of information, accumulated from different channels, contains much individual data. Whenever individual and delicate information are distributed and/or broke down, one essential inquiry to consider is whether the examination damages the security of people whose information is alluded to. The significance of data that can be utilized to build income cuts costs or both. Information mining programming is one of various expository apparatuses for investigating information. It permits clients to break down information protection is becoming continually.

Information mining has numerous applications in security incorporating into national security (e.g., observation) and also in digital security (e.g., infection location). The dangers to national security incorporate assaulting structures and annihilating basic frameworks, for example, power lattices and telecom frameworks. Information mining methods are being utilized to distinguish suspicious people and assemble, and to find which people and gatherings are fit for doing terrorist exercises. Digital security is worried with shielding PC and system frameworks from debasement because of malevolent programming including Trojan stallions and infections. Information mining is additionally

being connected to give arrangements, for example, interruption location and evaluating. In this paper we will concentrate basically on information digging for digital security applications. There has been a considerable measure of work on applying information digging for both national security and digital security.

A significant part of the center of our past paper was on applying information digging for national security. Information mining, the revelation of new and intriguing examples in substantial datasets, is a blasting field. One angle is the utilization of information mining to enhance security, e.g., for interruption recognition. A second angle is the potential security perils postured when an enemy has information mining capacities .Protection issues have pulled in the consideration of the media, legislators, government organizations, organizations, and protection advocates.

II. DATA MINING AS A SECURITY CONCERN

There has been a ton of work on applying information digging for both national security and digital security. A great part of the center of our past paper was on applying information digging for national security. In this a player in the paper we will talk about information digging for digital security.

A. Digital terrorism, Insider Threats

Digital terrorism is one of the real terrorist dangers postured to our country today. As we have specified prior, this risk is exacerbated by the tremendous amounts of data now accessible electronically and on the web. Assaults on our PCs, systems, databases furthermore, the Internet infrastructure could crush to organizations. It is assessed that digital terrorism could cause billions of dollars to organizations. A work of art illustration is that of a managing an account data framework. On the off chance that terrorists assault such a framework and exhaust records of reserves, then the bank could free millions and maybe billions of dollars. By injuring the PC framework a great many hours of profitability could be lost, which is eventually proportionate to coordinate money related misfortune. Indeed, even a basic force blackout at

work through some mischance could bring about a few hours of efficiency misfortune and as a result a noteworthy monetary misfortune. In this manner it is important that our data frameworks be secure. We talk about different sorts of digital terrorist assaults. One is the engendering of pernicious portable code that can harm or break touchy documents or other information; another is interruptions upon PC systems. Assaults, however data related insider dangers are regularly ignored or thought little of. Individuals inside an association who have contemplated the business' practices what's more, systems have a colossal point of interest when creating plans to injure the association's data resources. These individuals could be standard representatives or even those working at PC focuses. The issue is entirely genuine as somebody might be taking on the appearance of another person and bringing about a wide range of harm. In the following few segments we will look at how information mining can be utilized to recognize and maybe avert such assaults.

B. Charge card Fraud and Identity Theft

We are listening to a ton nowadays about Visa Misrepresentation and wholesale fraud. On account of Visa extortion, an aggressor gets a man's Visa and uses it to make unapproved buys. When the proprietor of the card gets to be mindful of the misrepresentation, it might be as well late to turn around the harm or capture the offender. A comparable issue happens with phone calling cards. In reality this sort of assault has transpired by and by. Maybe while I was making telephone calls utilizing my calling card at air terminals somebody saw the dial tones What's more, replicated them to make free calls.

The issue was managed quickly. A more genuine robbery is data fraud. Here one expect the character of someone else by obtaining key individual data, for example, government disability number, and uses that data to do Exchanges under the other individual's name. Indeed, even single such exchange, for example, offering a house and keeping the salary in a deceitful ledger, can have obliterating results for the casualty. By the time the proprietor discovers it will be excessively late. It is likely that the proprietor may have lost a large number of dollars because of the data fraud.

We have to investigate the utilization of information digging both for Visa misrepresentation discovery and additionally for wholesale fraud. There have been a few endeavors on identifying MasterCard extortion.

C. Assaults on Critical Infrastructures

Assaults on basic bases could injure a country and its economy. Foundation assaults incorporate assaulting the telecom lines, the electric, power,

gas, supplies and water sup-employs, sustenance supplies and other essential substances that are basic for the operation of a country. Assaults on basic foundations could happen amid an assault whether they are non information related, data related or bioterrorism assaults. For instance, one could assault the programming that runs the information transfers industry and close down all the telecom lines. Thus, programming that runs the force and gas supplies could be assaulted. Assaults could likewise happen through bombs and explosives. That is, the telecom lines could be physically assaulted. Assaulting transportation lines for example, interstates and railroad tracks are additionally assaults on frameworks. Frameworks could likewise be assaulted by normal debacle, for example, tropical storms and earth shudders. Our fundamental enthusiasm here is the assaults on foundations through pernicious assaults, both data related and non information related. We will probably inspect information mining and related information administration advances to recognize and counteract such foundation assaults.

D. Information Mining for Cyber Security

Information mining is being connected to issues, for example, interruption discovery and reviewing. For instance, inconsistency discovery procedures could be utilized to recognize surprising examples and practices. Join investigation might be utilized to follow self-engendering malevolent code to its creators. Arrangement might be utilized to gather different digital assaults and after that utilization the profiles to identify an assault when it happens. Forecast might be utilized to decide potential future assaults depending in a path on data learnt about terrorists through email and telephone discussions. Additionally, for a few dangers non real time information digging may suffice while for certain other dangers, for example, for system interruptions we may require constant information mining. Numerous specialists are researching the utilization of information digging for interruption location. While we require some type of constant information mining, that is, the outcomes must be produced in real time, we likewise need to assemble models progressively. For sample, MasterCard extortion discovery is a type of real time handling. In any case, here models are typically worked early. Building models progressively remains a test. Information mining can likewise be utilized for breaking down web logs and also dissecting the review trails. In view of the aftereffects of the information mining apparatus, one can then figure out if any unapproved interruptions have happened and/or whether any unapproved questions have been postured.

III. CONCLUSION

This paper has examined information digging for security applications. We initially began with an exchange of information digging for digital security applications and afterward gave a brief outline of the instruments we are creating. Information digging for national security and also for digital security is an exceptionally dynamic examination area. Various information mining systems including join investigation and affiliation principle mining are being investigated to distinguish irregular examples. As a result of information mining, clients can now make a wide range of connections. This additionally raises protection concerns.

One of the zones we are investigating for future examination is dynamic barrier. Here we are researching approaches to screen the foes. For such observing to be powerful, the screen must keep away from discovery by the static and element examinations utilized by standard hostile to malware bundles.

REFERENCES

- [1] Thuraisingham, B., "Web Data Mining Technologies and Their Applications in Business Intelligence and Counterterrorism", *CRC Press*, FL, 2003.
- [2] Chan, P, et al, "Distributed Data Mining in Credit Card Fraud Detection", *IEEE Intelligent Systems*, 14 (6), 1999. Lazarevic, A., et al., "Data Mining for Computer Security Applications", Tutorial *Proc. IEEE Data Mining Conference*, 2003.

Enhancement of Educational Experience through Mobile Technology

Guided by: Neelam Naik

Deepak Anthati
TIMSCDR

Neha Chauhan
TIMSCDR

Deep Singh Dangi
TIMSCDR

Abstract – The mobile technology have shown that they need an outsized quantity of potential for academic use. This technology is developed for master skill, learn world languages, crucial thinking skills, and improve skills in arithmetic, engineering, technology, and science among students. A number of these samples of mobile learning utilized the quality options of mobile devices, whereas others used new innovative options. Mobile devices facilitate students to find out information within or outside the schoolroom. Mobile technologies are comparatively cheap and might reach underserved students with restricted incomes. Mobile devices teach students social skills that are important and necessary for fulfillment within the twenty first century. Mobile devices are tiny and having a natural suitable use among the educational setting. Mobile technology provides students with tailored academic expertise as a result of the devices will be customized and utilized in many various ways that. This work highlights a number of these options of mobile technology.

I. INTRODUCTION

Nowadays Mobile devices are introducing exciting new prospects once it involves digital learning. For-profit companies are jumping on board; however the academic trade is basically leading the manner for innovation during this area. Mobile devices can be smart phones, tablets, ultra-portable laptop.

These new, efficient technologies are greatly established with today's students. A recent survey saw forty seventh of scholars say that they need used a mobile app for learning functions. Students aren't the sole ones jumping on board. Another study found that hour of college thought that students were actuated by lessons that incorporated iPad (compared to those that did not). Roughly sixty seven of the school felt that the iPad conjointly improved the standard of the students' work.

II. USES OF MOBILE TECHNOLOGY IN TEACHING

1. Preparing Students for the Future: The fundamental and important job as a lecturer is to organize your students for the longer term. Well so as to try and do therefore you would like to include mobile technology within the room. Operating with

mobile devices won't solely be a section of their everyday lives as adults, however it'll even be very important a part of several career methods. Knowing a way to suitably use mobile devices is a very important facet during this progressively connected world. So as to properly prepare your students for the longer term as they transform into the labour force, incorporating mobile technology within the room is essential. [1]

2. Up-to-date learning: The recent days of searching for data in encyclopaedias area unit long gone. Having mobile devices within the room permits students instant access to the most recent news, data, statistics, etc. nearly each question they need is at their fingertips, keeping them connected with what's occurring around them and making certain that they're continuously well aware with the foremost up-to-date data.

3. Alternative to textbooks: Many textbooks don't seem to be the foremost relevant sources of knowledge. Today's generation has become full-grown conversant in instant, updated data. Textbooks will give students with the most recent data sort of a mobile devices can. Also, having digital textbooks on their mobile devices keeps students additional organized and provides them easy accessibility to their materials. Nobody likes lugging around massive textbooks. Several digital textbooks area unit perpetually updated and sometimes additional vivid, helpful, creative, and plenty cheaper than those recent serious books [2].

4. Learning goes outside of the room: By permitting mobile devices in class you'll be able to expand learning outside of the classroom. Students won't solely have access to data throughout pc research lab time (which is additionally changing into extinct). They will hunt data from anyplace on field. Collaboration can increase as students will use these devices as analysis tools throughout.

Plus students love technology continues learning outside of college hours. Due to digital text books and learning apps on mobile devices, it allows them to urge in further learning and learning in throughout the time. They will carry their books and notes with them the least bit times and have

instant access to materials. If students are extremely excited and engaged in learning within the room, they're really to continue learning outside of the room and that they will do so with mobile technology.

5. Easy access: It is additionally convenient to the learners that Mobile learning provides easy accessibility anywhere and anytime. Learners have a bonus of paying their free time throughout travelling, in between conferences or throughout weekends to concentrate on subject they require to find out.[3]

6. Collaborative Learning: mLearning encourages cooperative learning permitting the learners at totally different locations to urge to bear with their peers or others groups to debate and learn.

7. Learner engagement: Mobile learning is unique platform to be engaged in coaching. The coaching at the work place consists of principally verbal and desktop communication; however adapting mobile learning will bring many opportunities to have interaction the learner on a social level outside of the work.

8. Self-Pace Learning: Each has his or her own approach of understanding the content or methods to find out. With mobile learning, learners are currently able to learn in their own vogue at their own pace.

9. Address all learning designs: Mobile learning will slot in to totally different learning styles because it permits them to try and do the subsequent Learning through Videos, Reading, Research on the net and Listening to Audio.

III. LIMITATIONS

1. Connectivity: There are conjointly some property problems whereas downloading and uploading of data because of poor mobile network signals.

2. Screen size: As a result the screen size is just too small; it'll strain the eyes of the learners if they use it for long quantity of it slow. In addition entirely less knowledge or just the gist of content could also be provided as results of size constraints.

3. Device: Learners ought to possess the mobile device that is supporting the courseware and this device is additionally expensive. As technology keeps dynamically changing at a quick rate, these devices need to be upgraded oftentimes. In addition to the worth of these devices, there are monthly data charges from mobile network providers; so downloading large content not entirely takes time but in addition costs heaps.

4. Distraction: Whenever accessing the courses through mobile devices, if the learner gets a choice or SMS or social media updates, then they are bound to get distracted. Therefore, mobile learning can extraordinarily facilitate employees to avail the benefits of anytime and anywhere learning in their organizations. It opens various doors to new technology and may continue as a result of the years go.

IV. SOME DOWNSIDES

As with something, there are unit natural downsides to mobile learning. First, it's troublesome for college students to avoid distractions, opting to look at a funny YouTube video rather than the lesson lecture. Also, it's tougher to verify the credibleness of a top quality instructional supply. The net could be a little bit of a scrap, and also the actual validity of abundant of the content are often questioned. If students leverage this content for his or her studies, they'll be relying upon inaccurate data.

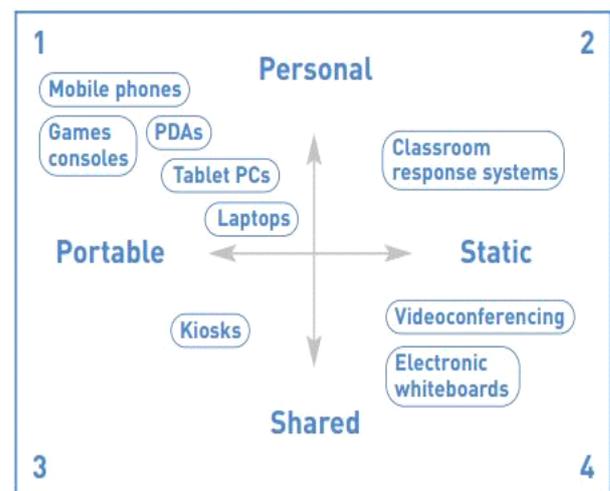


Fig1: Existing types of mobile apps to support teaching

There are many alternative styles of technology which will be classed as 'mobile'. Mobile, to most, means that 'portable' and 'movable'. It additionally appears to implicate a 'personal' as hostile 'shared' context of use, and also the terms 'mobile' and 'personal' are typically used interchangeably – however a tool can be one while not essentially being the opposite.

Quadrant one show devices which will be classified as each moveable and private. These styles of devices area unit what folks most typically consider in relevancy mobile technologies: mobile phones, PDAs, pill PCs and laptops. It additionally includes hand-held game consoles, with Rosas et al (2003) and Lee et al (2004) coverage on early

evaluations of their instructional use. Since these devices usually support one user, they're usually perceived as being terribly personal. The networked nature of such devices affords communication and data sharing, which means that whereas the devices themselves are unit personal, the knowledge at intervals them are often shared simply. These devices are moveable as a result of they're taken from place to position and thus they'll be on the market in many alternative locations. These are unit personal moveable technologies. Other technologies, less moveable than mobile phones and PDAs, will still supply personal interactions with learning experiences. Schoolroom response systems, shown in quadrant a pair of, encompass individual student devices that won't respond to anonymously to multiple alternative queries administered by an educator on a central server. This technology is static within the sense that it will solely be utilized in one location, however remains personal as a result of its small size and allocation to (typically) only one single user. These are personal static technologies. Being physically captive from one place to a different isn't the sole means during which mobile technologies are often 'portable'. In quadrant three, there are unit samples of technologies which will offer learning experiences to users on the move; however the devices themselves don't seem to be physically movable. Street kiosks, interactive depository displays and other forms of installations supply pervasive access to data and learning experiences, however it's the learner United Nations agency is moveable, not the delivery technology. Such devices are unit usually seen as being less personal, and are probably to be shared between multiple users. Their larger size means that they're additionally higher suited to multiple-user interactions. For additional shareable interactions, the devices themselves should become larger and thus less moveable. Examples video-conferencing facilities and embody interactive schoolroom whiteboards and, as shown in quadrant four. These technologies are enclosed to point out the entire area of potentialities engendered by our classification; however they'd usually not be classed as mobile technologies. We have a tendency to believe that 'mobile technologies' comprise all devices from quadrants 1-3 and people from quadrant four that don't seem to be at the acute finish of the 'static' dimension.

A) Mobile Learning: Mobile learning is taken into account to be the flexibility to use mobile devices to support teaching and learning. It's the 'mobile' facet of mobile learning that produces it stand except alternative kinds of learning, specifically planning learning experiences. E.g. - within the education management business, podcasts of lectures are often created on the market

for downloading. Learners are unit expected to interact with these learning resources if far from the standard learning areas. The most challenge in mobile learning is property and battery life.

B) Native Apps: These are developed specifically for one platform, and might take full advantage of all the device options. Native apps are unit put in through an application store (such as Google Play or Apple's App Store). Once you begin the app, it starts in real time. Once you use the app, you get quick performance, consistent platform look and feel. Once your app wants to update, it tells you therefore.

C) Mobile web optimized or HTML5 Apps: They have to go past a browser usually written in HTML5. An HTML5 mobile app is largely an online page, or series of websites, that are unit designed to figure on a little screen. As such, HTML5 apps are unit device agnostic and might be opened with any fashionable mobile browser. And since your content is on the online, it's searchable, which may be quite profit looking on the app (shopping, for instance. HTML5 apps are unit easier to develop, easier to support, and might reach the widest vary of devices.

D) Hybrid Apps: Hybrid apps are unit half native apps, half net apps. (Because of that, many of us incorrectly decision them "web apps"). Like native apps, they sleep in an app store and might make the most of the various device options on the market. Like net apps, they have confidence HTML being rendered in an exceedingly browser, with the availability that the browser is embedded at intervals the app.

V. SUGGESTION FOR DEVELOPMENT OF RECENT APPLICATION

The purpose of this study is to spot the advantages that mobile technology which is in education. From a review of the literature and interviews with people within the field of education, we have a tendency to developed associate activity based model to associate the uses of mobile technology with their edges. The model first identifies the final uses of mobile technology in education and also the specific activities that they cover. Then, the advantages from their uses are known with overarching themes rising as follows: bigger efficiency and effectiveness in learning, multiplied individual support and opportunities for private development, greater strategies of cooperating and human activity and larger exposure to information technology.

Mobile Learning (M-Learning) is associate evolving space of e-Learning because of the regularly increasing demand by the learners of associate ever-increasing variety of mobile (hand-held) devices to move with advanced services offered by learning platforms. These services

modify mobile learners to use the benefits of net, mobile computing and mobile communications. The interaction between mobile learners and repair suppliers is accomplished through completely different mobile and wireless networks. M-Learning is associate knowledge domain study covering mobile communications, applied science, education and different studies. Therefore, a short introduction of the present mobile devices and technologies, further because the technical problems touching the standard of m-Learning applications and services is critical for the reader to know the conception of this paper.

M-Learning services will be adopted through completely different systems supporting quality and plenty of solutions for wireless access, like wireless telecommunications networks, specifically cellular networks (GSM, GPRS, UMTS) and different wireless networks, as as an example IP Wireless native space Networks (WLANs) mistreatment the IEE. The system should embody all the mandatory options to accomplish the specified tasks. Accuracy, quality, compliance, ability and privacy area problems that have to be investigated in planning an M-Learning system to confirm that the system can perform as predicted to. These factors ought to be rated consistent with their importance and enforced consequently. The teacher ought to be able to manage the implementation of those options, particularly on privacy [3]. The sessions will be used as a tool to facilitate the teacher in student assessment.

VI. CONCLUSION

The mobile technology is exclusive for academic reform. Education needs a lot of analysis and support from each of the general public and personal sectors to succeed. They have to develop interventions for mobile learning to make it public understanding of the technology so to enhance education for youngsters of all ages. A method to accomplish this is often by making samples of mobile technology in education and presenting them to the general public. It is needed to actively encourage the employment of mobile technologies within the classroom to the general public and to academic policymakers. Educators should steel oneself against the employment of mobile technologies within the schoolroom by coaching colleagues to use and incorporate mobile devices into learning and obtain support from the country's leadership for the tutorial use of mobile technologies.

REFERENCES

- [1] <http://www.qrcodescanning.com/mobile-technology-classroom-learning.html>
- [2] <https://net.educause.edu/ir/library/pdf/ers0502/cs/ecs0502.pdf>
- [3] <http://teach.com/teach100/blogs/378-Educational-Technology-and-Mobile-Learning>

A Survey of Free Space Optical Communication Network Channel over Optical Fiber Cable Communication

Guided by: Mr. Gajendra Dixit

Prakash Mandhyan
TIMSCDR

Krishnakant Maurya
TIMSCDR

Rahul Maurya
TIMSCDR

Abstract: In the recent past, free-space optical correspondence (FSO) wound up being a key substitution to radio repeat correspondence. In optical fiber join correspondence, there is an obstacle that optical fiber gets hurt in a long partition foundation. Radio repeat correspondence moreover oversees confinements of transmission limit and data rate. In this paper, a brief depiction about free space optical (FSO) correspondence and its accessibility with central cell framework is cleared up. This paper in like manner gives the information about the piece of free space optics in the central cell framework. It contains a RF framework, which has relationship within the Radio Access Framework (RAN) and the Central Framework (CN) for uniting the adaptable customer endorser of Individuals by and large Traded Telephone Frameworks (PSTN) and Web. Plus, free space optical correspondence makes prepared for a high exchange speed requiring applications with license free operation, high transmission security, full duplex transmission and tradition straightforwardness will give a more broad point of view to free space optical correspondence.

Keywords: SO, RAN, CN, PSTN, PAT.

I. INTRODUCTION

Optical correspondence framework can likewise be named as optical telecom framework. It began in 1970 and works in 800nm-1600nm wavelength range. In this framework, light is utilized for conveying the data. The data transmission of the optical correspondence framework is 1000 times higher which helps in simple dispatching of the articles. In this framework, numerous clients impart in the meantime because of high data transmission effectively. In the present situation, optical correspondence underpins 300 THz data transfer capacity [1]. In optical correspondence framework, the assignment is performed by utilizing three parameters: Transmitter, Channel and Recipient.

The capacity of the transmitter is to encode the message electrical sign into an optical sign. Semiconductor gadgets are generally utilized as a

part of optical transmitters such as Drove (Light-Emitting Diodes) and LASER diodes. The distinction between the two is that, Driven is utilized for short range correspondence and produces unintelligible light while LASER is utilized for long range correspondence and produces cognizant light. Semiconductor optical transmitters ought to be exceptionally proficient, minimized and solid. It can be specifically regulated at high frequencies [1]. A semiconductor Drove discharges light through unconstrained emanation. LEDs are extremely helpful for minimal effort applications on account of its basic outline. A semiconductor laser discharges light through the unconstrained outflow [2-7].

The channel takes the sign to its destination. It is an adaptable and straightforward fiber made up of excellent silica or plastic. It works likewise as a waveguide for transmitting light starting with one end of the fiber then onto the next. It is for the most part utilized for long separation transmission and with higher information rates than wire links. The beneficiary produces the message signal from the received data signal. The photograph locator is the prime segment of the optical collector which changes over light into power by utilizing the photoelectric impact. It is a semiconductor based photodiode. PN photodiodes, PIN photodiodes and torrential slide photodiodes are the sorts of photodiodes. The coupling mix of electrical converters, trans-impedance intensifier and a restricting enhancer delivers an advanced sign in the electrical space from the approaching optical sign, which might be lessened while going through the channel [8, 9-11].

Optical fiber correspondence has a few favorable circumstances due to which it has been utilized for quite a while. It has to a great degree high data transfer capacity and alongside this, numerous clients convey effortlessly in the meantime. It has a

low assembling taken a toll with Bit Blunder Proportion (BER) of 10^{-6} [12]. Be that as it may, it has a few drawbacks likewise that it must be done through point to point correspondence and proliferation of light is unidirectional. Be that as it may, the detriments of optical fiber correspondence are uprooted by Free Space Optical Correspondence. It is characterized as the innovation which utilizes light to propagate as a part of vacuum or air to transmit information for PC organizing. This innovation is useful on account of high cost because of the illogicalness of physical associations. It is anything but difficult to convey and has high piece rates and low piece blunder rates. It is sans permit and utilized for long range operation. Be that as it may, it has some evil impacts like Climatic retention, Bar scattering and Deduction from foundation light sources.

The remaining paper is sorted out as takes after. In section II, a point by point portrayal about free space optical correspondence is given and section III includes the utilization of free space optical correspondence in the focal cell system. The paper is finished up in section IV.

II. FREE SPACE OPTICAL COMMUNICATION

The communication through the cable infrastructure has become a part of our past. Since optical fiber communication is better in some respects like large distance communication up to 40 km without even requiring a repeater device, flexible nature, which made it ready to be used anywhere, duplex mode of communication, EMI (electromagnetic interference) resistive nature and large bandwidth supporting capacity. But apart from all these, modern environment demands for more efficient bandwidth utilization and better data sheet, which lead to another method of correspondence called as free space optical correspondence. Following optical fiber cabled framework has neglected to achieve a less cell end point. In any case, this constraint is evacuated by free space optical correspondence. It utilizes free space for the correspondence reason, so otherwise called fiberglass optics or optical remote [13].

A free space optical system uses simple on-off keying as modulation format which allows a free space optical system to be designed as bandwidth and protocol transparent physical layer connection. In order to implement a free space optical network, we need to have following two crucial components as a part of it, radio access network (RAN) and central network (CN). RAN provides the connectivity between the mobile users and CN by using RF frequency signal as a medium. RAN uses a base station and base controller, while CN provides the PSTN for mobile to mobile telephony

The microwave RF signal is also used as a source of connectivity.

Free Space Optical framework has a few favorable circumstances over the optical fiber correspondence framework:

- Ease of deployment: FSO framework has supplanted the base based OFC framework without letting its execution down. Besides, it is additionally financially savvy technique and free from bothers of burrowing and burring the links.
- License free operation: FSO framework requires no RF permit so it can be effectively sent. It is effortlessly upgradable and its open interface bolster hardware from an assortment of sellers, which helps ventures and benefit suppliers ensure their interest in installed telecom framework.
- High speed and bit rate: FSO correspondence is staggeringly quick and permits an information rate of 2.5 GB/s. This is an abundant measure of BW for exchange of documents between two closures. Besides, these rates are relied upon to increment up to an astounding are around 10 GB/s. Subsequent to in FSO, pillar restriction is so vital and successful that two bars work freely and such differences between two data conveying bars gave for all intents and purposes boundless level of recurrence reuse in any ordinary situations.
- Security: FSO is an extremely secure remote arrangement of correspondence. Because of tight pillar or bar restriction nature required for transmitting a sign, a sign by means of a laser to RX it will be, it is for all intents and purposes difficult to break that data in the system. Subsequent to the laser pillar can't be recognized by a range of RF meter so it is one more accreditation for FSO'S secure correspondence. Besides, one must be in the viewable pathway in the middle of RX and TX to have the capacity to over drop the data. Be that as it may, in the event that it is conceivable, then it would to accepting site have lost associations there are no security overhauls that are required for FSO. Figure 1 depicts a general free space optical system.

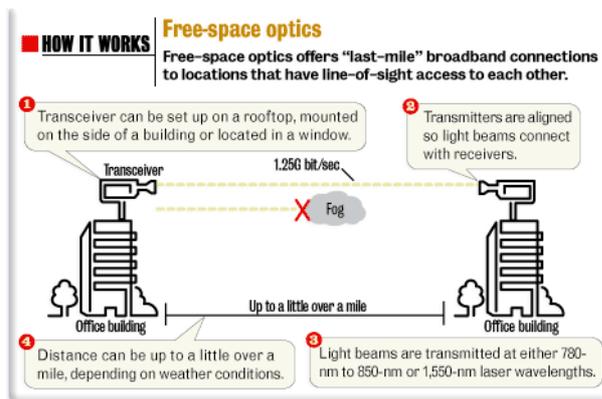


Fig 1: A General Free Space Optical Network

Each innovation accompanies some feeble focuses additionally which will make alive the examination in the theme. FSO has some frail focuses and they are

- Distance: The separation up to which a FSO framework conveys is not all that vast and is restricted inside of 2 km range. In spite of the fact that it is a force full framework with extraordinary throughput, yet separate restriction is a major advantage for its utilization.
- Line of sight communication: A tight laser pillar or a more restricted bar is basic in the event of FSO, yet LOS in FSO is influenced by another element as:
 - o Fog: Thick mists for the most part influence the pillar control nature. It either totally traps the shaft or diminishes its power after various refractions, scrambling and ingestion. A technique to diminish this is to utilize it for short separations.
 - o Atmospheric effect: The different air impacts that can truly change the way of the laser bar or to lessen its energy are retention of the water atom, scrambling by suspending particles in nature physical deterrents or developments in the method for LOS furthermore the sun oriented impedance. Disseminating happens when the wavelength of light crashes into a molecule. The sort of dispersing which happens is by choosing the wavelength of light.
 - If wavelength < scatterer size (Rayleigh scattering)
 - If wavelength > scatterer size (Mie scattering)

Subsequent to in FSO lasers are utilized as transmitters and as a part of FSO since the correspondence happens in free space, along these lines an uncommon consideration must be taken that laser pillar does not influence our eyes or the particular wavelength is critical in light of the fact that just certain wavelength between 0.4 and 1.4 μm are gone by eye on the retina. In this way, strict worldwide principles have been set for wellbeing and execution.

III. APPLICATION OF FREE SPACE OPTICS IN CENTRAL CELLULAR NETWORK

Free space optical items are installed with cell systems as observable pathway successor of natural connections in the focal cell system. Free space optic items having advantages of no cabling and no range permitting. They offer fast information rate, with the limit of optical fiber. The fundamental weaknesses of free space optic correspondences utilized as a part of practice are connection unwavering quality because of climate, and separation [15]. These parameters limit free space optic connections to a couple of kilometers.

The laser has an imperative property of profoundly directional pillars. Free space optical framework recipient is outlined in such an approach to get a little uniqueness and focus less optical vitality on a beneficiary. Every last optical handset is pointed in an animating way to another in the correspondence process. In view of the natural impacts such as barometrical weight, air and warm stacking of hardware and building structure, free space optical framework need to execute Guiding, Securing, and Following (PAT) subsystems [16]. The more troublesome undertaking is to grow PAT however in RF correspondence transmitter and beneficiary have fundamental hardware. Flying out from point to point is non-paltry. The difference of the transmitted shaft and the beneficiary field-of-perspective must be more noteworthy than the pillar jitter so as to get careful information.

Traditional Free space optic frameworks, first fiber sending end handset change over an electrical sign into an optical sign. At that point in same side electrical sign is opened up by a laser driver sufficiently giving momentum to drive the laser diode. A regulated light rush of laser diode is coordinated through the channel to the comparing beneficiary which centers the light shaft onto a photograph locator [17, 18-21]. Finally the optical sign changed over once more into an electrical sign. Figure 2 outlines the free space optical correspondence framework

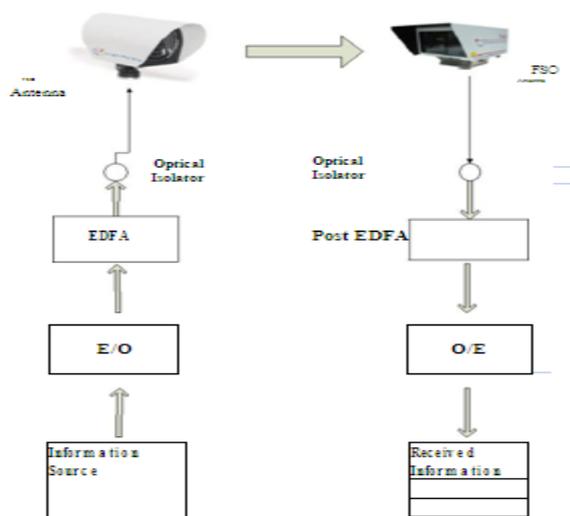


Fig 2: Free Space Optical Communication System

In remote optical correspondence frameworks, light discharged specifically from a fiber end to free space through an optical radio wire. At the collector end, the transmitted optical pillar is engaged, utilizing the recipient optics, specifically to a fiber and after that sent down the fiber for recognition. Thus the transformation of a sign from electrical to optical and optical to electrical is evacuated, which offer ascent to data transfer capacity and convention straightforward correspondence connect and is additionally much less demanding to incorporate with cabled foundation.

We have contemplated for quite a while about radio recurrence and free space optics on account of their coordinating nature. Radio and free space join correspondence; both have the ability to utilize both at once. The consolidate utilization of both will have preference on a solitary correspondence framework. Both are influenced by environment emergencies. Free space optic connections are seriously lessened in foggy conditions, and microwave RF frequencies are fundamentally constricted by downpour, because of dispersing that happens in the event of stormy season. Optical remote permits high information rate and high throughput which is profitable in postponement impolite applications [22].

The move from wired to the remote system client by keeping up the availability independent of client's versatility and land limitations is satisfied by the radio access system. With expanding request or expect levels of the portable client, the radio access arrange alone is not possessed the capacity to give expanded limit and expanded scope with constrained radio assets. In prior correspondence

frameworks, numerous base handset stations were associated with a Base Station Controller, which assumes the liability of Radio Asset Administration. Be that as it may, in a LTE propelled correspondence framework, anodes are in charge of Radio Asset Administration choices. Suitability of free space optical connections in the Radio access systems is gotten to in this correspondence framework.

Recently created cell frameworks will bolster the client with shorter separation between the client and the working settled systems. This can be accomplished by expanding the quantity of reception apparatuses to 4 times while moving from 3G to 4G advancements. Composed Multipoint transmission and gathering, transfer arranges and disseminated receiving wires are the financially savvy choices that will lessen to manage approaches like cell part and sectoring, which will expand the sign to commotion proportion to and from portable clients. Free space optical correspondence will display a satisfying alternative of transmission of RF signs over optical fiber to connection remote system associations. This is termed as Radio over Fiber or RF Photonics [23]. In this technique, optical transporters are utilized to place simple Radio Recurrence signals and afterward transmitted through optical fiber links having high limit. In optical fiber transmitter tweak system, the optical bearer is regulated with the radio sign. Optical fiber offers a little measure of constriction and is free from multipath blurring. It is likewise shielded from electromagnetic obstruction. The confinement in establishment of optical fiber links has demonstrated the handiness of free space optic correspondence frameworks that are equipped for supplanting Radio over Fiber joins.

For low-transfer speed based administrations like voice administrations, propelled remote frameworks are unbendable and cost-incapable as a result of consistently changing movement properties and system prerequisites. Way differences are accomplished from short various jump joins between system components in cross section topologies which are valuable with the expectation of complimentary space optical correspondences. Expansive recurrence reuse makes it easy to execute and scale in a cross section topology in free space optic connections when contrasted with RF.

The little administration territories with low transmit power Pico-base stations are known as Pico cells [24]. They will connect for less portable endorsers which are inside a building and give them administrations. Since telecom organizations have spent a lot of their money related asset for

giving optical fiber network in the middle of nations and urban areas yet this availability is still distracted for some high information throughput necessity structures. In this manner a huge issue of "last mile" is still present. Since it is impractical dependably to set out an optical fiber, as it is unreasonable and additionally tedious. So Free space optic correspondence frameworks are an answer of the "last mile" issue, chiefly in overpopulated urban ranges.

IV. MERITS

- ✓ Flexible system arrangement over routine broadband administrations.
- ✓ Straight forward sending no licenses required
- ✓ Low beginning speculation
- ✓ Ease of establishment
- ✓ Re-deployability
- ✓ High piece rates and low blunder rates

V. DEMERITS

- ✓ Haze
- ✓ Physical checks
- ✓ Glimmer
- ✓ Sunlight based impedance
- ✓ Diffusing
- ✓ Assimilation
- ✓ Building influence/Seismic movement

VI. ADVANTAGES OF FIBER OPTICS

1. Greatly high transfer speed – No other link based information transmission medium offers the data transfer capacity that fibre does.
2. Simple to accommodate expanding data transfer capacity – Using a large portion of the late eras of fiber optic cabling, new gear can be added to the idle fiber link that can give tremendously extended limit over the initially laid fiber. DWDM, or Dense Wavelength Division Multiplexing, loans fiber optic cabling the capacity to turn different wavelengths of light going down the fiber on and off freely. These two attributes of fiber link empower dynamic system transmission capacity provisioning to accommodate information movement spikes and respites.
3. Imperviousness to electromagnetic obstruction – Fiber has a low rate of bit mistake (10 EXP-13), as a consequence of fiber being so impervious to electromagnetic impedance. Fiber-optic transmission is for all intents and purposes commotion free.
4. Early recognition of link harm and secure transmissions – Fiber gives a to a great degree secure transmission medium, as there is no real way to distinguish the information being transmitted by "listening in" to the electromagnetic vitality "spilling" through the link, as is conceivable with customary, electron-based transmissions. By continually checking an optical

system and via painstakingly measuring the time it brings light to reflect down the fiber, grafts in the link can be effortlessly identified.

VII. DISADVANTAGES OF FIBER OPTICS

1. Establishment costs, while dropping, are still high – Despite the way that fiber establishment expenses are dropping by as much as 60% a year, introducing fiber optic cabling is still generally exorbitant. As establishment costs diminish, fiber is extending past its unique domain and significant application in the bearer spine and is moving into the nearby circle, and through advances, for example, FTTx (Fiber to the Home, Premises, and so forth,) and PONs (Passive Optical systems), empowering endorser and end client broadband access.
2. Extraordinary test gear is frequently required - The test hardware regularly and generally utilized for ordinary electron-based systems administration is of no utilization in a fiber optic system. Hardware, for example, an OTDR (Optical Time Domain Reflectometer) is required, and costly, particular optical test hardware, for example, optical tests are required at most fiber endpoints and association nexuses keeping in mind the end goal to legitimately give testing of optical fiber.
3. Vulnerability to physical harm – Fiber is a little and smaller link, and it is exceedingly helpless to getting to be cut or harmed amid establishment or development exercises. Since railways regularly give privileges of-approach to fiber optic establishment, railroad auto crashes represent a critical link harm danger, and these occasions can upset support of expansive gatherings of individuals, as fiber optic links can give colossal information transmission abilities. In light of this, when fiber optic cabling is picked as the transmission medium, it is important to address rebuilding, reinforcement and survivability.
4. Natural life harm to fiber optic links – Many winged creatures, for instance, discover the Kevlar strengthening material of fiber link coats especially engaging as settling material, so they peck at the fiber link coats to use bits of that material. Beavers and different rodents use presented fiber link to hone their teeth and creepy crawlies, for example, ants fancy the plastic protecting in their eating regimen, so they can frequently be discovered snacking at the fiber optic cabling. Sharks have likewise been known not fiber optic cabling by biting on it when laid submerged, particularly at the rehashing focuses. There is a plant called the Christmas tree plant that regards fiber optic link as a tree root and wraps itself around the link so firmly that the light driving forces going down the fiber are interfered with.

VIII. CONCLUSION

The outline of development RF system is dependable for complex cross sections and useful for a free space optics prospect. By and large networks are considered for short connection removals between radio wires. This short connection gives dependability, speed, and differing nature of sign in transitory and awful climate conditions. Free space optical connection is more nonexclusive yet to some degree dangerous, however we can accomplish some perpetual applications. The portability diminishes as the separation increments, yet it can be uprooted by utilizing circular radio wires. There is work going ahead in a positive heading to expand the framework unwavering quality and coherence, yet the size and unpredictability of the framework likewise needs an answer sooner rather than later. There is likewise a need of steadiness in network of portable clients on account of indoor and semi dissemination transmission. The fundamental issues that are blocking the development are temperature stacking impacts, awful climate conditions; tempests, building, and vibrations. As optical strands have high data transfer capacity, which is the primary element why still optical links are utilized for long separation transmission. Free space optical correspondence is for suitable for short separation up to a greatest of 4 km. Because of unpredictability of access focuses, Opportunities will develop for short connection transmission.

REFERENCES

- [1] S. Ghosh, K. Basu and S. K. Das, "An architecture for next generation radio access networks," *IEEE Network*, vol. 19, 2005, pp.35-42
- [2] J. Hou and D. C. O'Brien, "Vertical handover-decision-making algorithm using fuzzy logic for the integrated Radio-and-OW system," *Wireless Communications, IEEE Transactions on*, vol. 5, 2006, pp. 176-185.
- [3] S. Ghosh, "Emergent technology based Radio Access Network (RAN) design framework for next generation broadband wireless systems," M.S. thesis, Dept. Comp. Sci. and Eng., Univ. Texas at Arlington, 2004.
- [4] T. Kamalakis, I. Neokosmidis, A. Tsipouras, S. Pantazis and I. Andrikopoulos, "Hybrid free space optical/millimeter wave outdoor links for broadband wireless access networks," in *Personal, Indoor and Mobile Radio Communications, IEEE 18th International Symposium on*, 2007, pp. 1-5.
- [5] M. Safari and M. Uysal, "Relay-assisted free-space optical communication," in *Signals, Systems and Computers, Forty-First Asilomar Conference on*, 2007, pp. 1891-1895.
- [6] J. Akella, C. Liu, D. Partyka, M. Yuksel, S. Kalyanaraman and P. Dutta, "Building blocks for mobile free-space-optical networks," in *Wireless and Optical Communications Networks, Second IFIP International Conference on*, 2005, pp. 164-168.
- [7] O. Bouchet, M. El Tabach, M. Wolf, D. C. O'Brien, G. E. Faulkner, J. W. Walewski, S. Randel, M. Franke, S. Nerreter, K. D. Langer, J. Grubor and T. Kamalakis, "Hybrid wireless optics (HWO): Building the next-generation home network," in *Communication Systems, Networks and Digital Signal Processing, 6th International Symposium on*, 2008, pp. 283-287
- [8] D. O'Brien, "Cooperation in optical wireless communications," in *Cognitive Wireless Networks: Concepts, Methodologies and Visions Inspiring the Age of Enlightenment of Wireless Communications*, F.H. P. Fitzek and M. D. Katz, Eds. Springer, 2007, pp. 623-634
- [9] J. P. Javaudin and M. Bellec, "Technology convergence for future home networks," in *Wireless Days, 2008, 1st IFIP*, 2009, pp. 1-5
- [10] R. Ramirez-Iniguez and R. J. Green, "Indoor optical wireless communications," in *Optical Wireless Communications, IEE Colloquium on*, 1999, pp. 14/1-14/7.
- [11] R. J. Green, H. Joshi, M. D. Higgins and M. S. Leeson, "Recent developments in indoor optical wireless systems," *IET Communications*, vol. 2, 2008, pp. 3-10.
- [12] Gupta, Akhil. "Improving Channel Estimation in OFDM System Using Time Domain Channel Estimation for Time Correlated Rayleigh Fading Channel Model." *International Journal of Engineering and Science Invention*, vol. 2, issue 8, pp. 45-51, August. 2013
- [13] K. Kazaura, K. Wakamori, M. Matsumoto, T. Higashino, K. Tsukamoto and S. Komaki, "RoFSO: A universal platform for convergence of fiber and free-space optical communication networks," *Communications Magazine, IEEE*, vol. 48, 2010, pp. 130-137.
- [14] R. Ramirez-Iniguez, S. M. Idrus and Z. Sun, *Optical Wireless Communications: IR for Wireless Connectivity*. CRC Press, 2008
- [15] J. C. Juarez, A. Dwivedi, A. R. Mammons, S. D. Jones, V. Weerackody and R. A. Nichols, "Free-space optical communications for next-generation military networks," *Communications Magazine, IEEE*, vol. 44, 2006, pp. 46-51.
- [16] D. M. Jeganathan and D. P. Ionov. (2001, Dec 27). Multi-Gigabits per-second optical wireless communications [Online]. Available: www.freespaceoptic.com.
- [17] S. S. Muhammad, B. Flecker, E. Leitgeb and M. Gebhart, "Characterization of fog attenuation in terrestrial free space optical links," *Optical Engineering*, vol. 46, 2007, pp. 066001-1-066001-10
- [18] C. C. Davis, I. I. Smolyaninov and S. D. Milner, "Flexible optical wireless links and networks," *Communications Magazine, IEEE*, vol. 41, 2003, pp. 51-57.
- [19] E. Ciaramella, Y. Arimoto, G. Contestabile, M. Presi, A. D'Errico, V. Guarino and M. Matsumoto, "1.28 Tb/s (32 x 40 Gb/s) WDM transmission over a double-pass free space optical link," in *Optical Fiber Communication, Conference on*, 2009, pp. 1-3
- [20] Kim and E. Korevaar, "Availability of free space optics (FSO) and hybrid FSO/RF systems," *Optical Wireless Communications IV*, vol. 4530, 2001, pp. 84-95.
- [21] K. Kazaura, P. Dat, A. Shah, T. Suzuki, K. Wakamori, M. Matsumoto, T. Higashino, K. Tsukamoto and S. Komaki, "Studies on a next generation access technology using radio over free-space optic links," in *Next Generation Mobile Applications, Services and Technologies*, 2008, pp. 317-324.
- [22] J. S. Seybold and I. NetLibrary, *Introduction to RF Propagation*. Wiley Online Library, 2005P. T. Dat, A. M. Shah, K. Kazaura, K. Wakamori, T. Suzuki, K. Takahashi, M. Matsumoto, Y. Aburakawa, T. Nakamura, T. Higashino, K. Tsukamoto and S. Komaki, "A study on transmission of RF signals over a turbulent free space optical link," in *Microwave Photonics*, 2008, 173-176.
- [23] M. Bass, "Atmospheric optics," in *Handbook of Optics, Third Edition*, vol. 5, M. Bass, Ed. McGraw-Hill, 2010, 3.3-4.1

Application of Information Technology for the Education of Dumb and Deaf Children

Guided by: Neelam Naik

Himesh Jha
TIMSCDR

Satish Gupta
TIMSCDR

Rashmi Sharma
TIMSCDR

Abstract - Mobile phones have the flexibility to form an honest academic content of pictures, shapes, and graphics and illustrations applicable signs to the Deaf and Dumb students and also the production of academic content appropriate for individual variations in education between them and meets their wants mental and their interests that are totally different from standard students in Education. The technology is employed to realize the interaction between deaf and dumb youngsters with others.

I. INTRODUCTION

Around 15 percent people are disabled due to any reason in India. People do suffer with disabilities as a result of weak or no capability to utilize a sensory organ like eye ear or management over their body elements, mental weakness etc. The matter may be by birth, as a result of aging method, accident etc. incapacity has its result on quality of human life no matter its origin and reason. Totally different style of disabilities needs support associated with handling that drawback. Issues of individuals with Multiple disorders becomes a lot of sophisticated as answer for such a sort of incapacity could not work for them, thus further efforts are required to supply devices and technological support to create their life easier. [2]

About 9 million folks within the world are thought-about to be a deaf or dumb. Communications between deaf-dumb and a standard person have perpetually been a difficult task. Typically dumb folks use signing for communication however they notice problem in human action with others World Health Organization don't perceive signing. This work aims to lower this barrier in communication by suggesting a transportable device that act as a wise assistant. The primary goal during this work is to supply a typical fashion for deaf and dumb peoples as traditional ones.[3]

II. MOBILE TECHNOLOGY TO HELP THE DEAF AND DUMB

A lot of individuals get sensible phones not just for pleasure, however to resolve necessary business-problems. They're gadgets of just about daily use for individuals of various ages and occupations. The rationale is clear; sensible phones and tablets

area unit portable; offer continuous access to the net and different helpful services. However, they have also become indispensable assistants for individuals with restricted talents (for example for deaf and dumb persons). Supplied with special computer code, these gadgets will solve issues, however vital for such individuals. It covers a large vary of problems from shopping for food at the grocery store to being educated at the university. In fact, in combination with some medical invention such mobile applications are capable to form the lifetime of deaf individuals. A lot of eventful and richer activities are giving them plenty of opportunities in social sphere. A sensible phone will become a private signer for a dumb person. The mobile application created by Alexander Griff catches the oral communication, records and rework it into a wave file then into video showing explicit signs. The appliance will operate with 3500 signs and it's the most important info in Russia taking into consideration that there is a unit regarding 6000 signs within the signing. The simplest Russian signer took half in making this info. Additionally, this mobile application will build correct sentences from the point of view of Russian syntax. Two third of the world's population use the foremost widespread languages. In each country, in each city there are deaf individuals, and many agencies would like such associate degree assistant providing no-hit communication. This circumstance makes new niches within the mobile computer code development. The innovation are often utilized in totally different spheres of social life from going searching and obtaining appropriate treatment in hospital to being educated in class and obtaining chance to vote throughout parliamentary elections, thus the impotence of such a computer code answer is simple.

III. LITERATURE REVIEW

N.Hema, Ms. P. Thamarai, Dr. T.V.U. KiranKumar (2013) [1] indicates however technology is employed to realize the interaction between deaf and dumb folks with others. This technology used for information acquisition and transmission. Opt for low hardware instrumentation and transport protocol, add the node sleep mechanism, so the system has low energy consumption, giant

communication vary, high stability characteristics. It is used because of its portability for Deaf and Dumb. Short distance as well as long distance communications is feasible. The screen gesture technique eliminates the utilization of advanced hand gestures. Thus removes the necessity for hand movement sensing systems that are quite giant, complex, valuable and slower. It supports for uneducated folks with Image translation feature. Learning mode helps uneducated folks to find out English words through it. Additional advanced user will use word translation. [1]

Dalia Nashat, Abeer Shoker, Fowzyah Al-Swatand Reem Al-Ebailan, (2014) [5], in their paper, have discussed an economical application for uneducated Deaf-Dumb application. This application aims to assist deaf and dumb by providing them with a lovely communication and learning tool. This work introduces a Mobile application that alters communication between uneducated Deaf-Dumb and traditional folks in our society. It additionally develops aid tool for deaf and dumb in several fields like restaurants, Hospitals and transportation. Moreover, this application introduces a straightforward translator from signing to English or Semitic and contrariwise [5]. Kimberly A. Weaver, Kimberly A. (2012) [6] have suggested that to form a mobile informal learning tool for hearing oldsters of deaf youngsters making an attempt to find out signing. They have a tendency to delineate the appliance and printed a study which can able to investigate victimization of application that will impact parents' ability to find out vocabulary. The ultimate contribution of this paper may be a reflection on the however device platform selections might impact the biological process cycle and also the analysis of mobile applications. Within the case of sensible Sign authors have attempted initial experimental validity and can then expand development to supply wider access to the target population. [6]

IV. MOBILE APPLICATION FRAMEWORK

Responsive Web Design (RWD) is associate best viewing expertise; simple reading and navigation with a minimum of panning, resizing, and scrolling across a good range of devices from mobile phones to desktop. Media queries permit the page to use totally different CSS rules supported the device. The victimization grid system of CSS media question of the web site are going to be converted into a Mobile internet App and one will access the web site by victimization any mobile iPhone, Android, Windows and conjointly iPod, desktop PC. Users tend to use complete resolution for making hybrid mobile apps with all you would like in one place with variety of mobile platform App and associated Cloud service to use as an app at backend. It comes with platform-independent App

that makes it easier to access phone hardware. A MVC framework supports to change speedy development of mobile apps. Modules created are simple to employ across totally different apps, therefore considerably reduces the event time and also the lines of code.

V. DATA SCALABILITY

Mobile applications have some special needs like location-based search. They are also usually not as weight as their desktop counterparts, as memory and computing cycles square measure restricted on a mobile device. What's additional they ought to have a reduced network demand and create fewer spherical visits between the device and also the server, load leveling the net server tier, load leveling the info tier. It conjointly introduces quick location primarily based search and lookups and least network usage. Optimize pictures of Deaf and Dumb sign education and/or take away or update them if doable. It also changes layout for the foremost well-liked devices, and pays time beyond regulation testing. This inventive resolution could work business needs and that we can bring immense speedups to the info tier, and so your overall Deaf and Dumb mobile education.

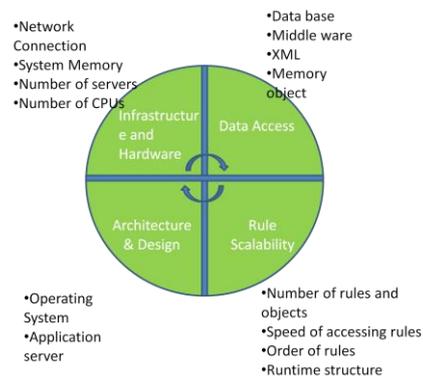


Fig 1: Mobile Application Scalability

VI. MOBILE DATA INTEGRATION

Mobile integration solutions from deliver a no-coding approach to application integration, facultative developers to deliver on business necessities is possible in an exceedingly fraction of the time. It'd take alternative mobile development tools. It's designed to facilitate integration between applications and systems that aren't generally mobile friendly. One can continue with the quantity, selection and speed of information needed to power mobile learning applications.

VII. MOBILE LEARNING APPLICATION

"Text to language Converter" is a Mobile learning application to Deaf and Dumb. It is used for conversion of text to language for effective communication with deaf and dumb individuals. It's conjointly a vicinity freelance language creator tool wherever new words and various pictures can be interplanetary into the library by internet cam and mobile camera. The applier takes text as associate input and generates the output dynamically by looking out many various pictures for each word and displays output by adjoining them with explicit time gap System searches and the occurrences of input words. Stage one describes the flows of operations between the users and also the mobile application. This stage contains 2 steps, namely, the registration and language generation. In the registration step, user may be associate institute or a personal user (home user). In such cases, the client should register with the mobile application information.

The registration method can be institute registration or home registration. In institute registration: The user goes to the institute to produce his personal data, registration method to user may be done by their institute. After that, the user name and Arcanum of the user creates in information and provides it by initial credentials. For later procedures between the user and mobile application for analysis these credentials can be used. In Home registration, the user and mobile application data are passed through personal channel to the registration server. And also the authentication of client personalities is created by him. As before, the information entry can created for user data and initial credentials are going to be causation back to the user or his parent.

The user make modifications of the mobile info initially once access the appliance interface. Synchronization happens between the servers and also the mobile application to make sure the information is consistent. So as to complete the synchronization, a publication is required. A publication is that the meta-data package of knowledge concerning that data is replicated. The info server will synchronize with the mobile info properly. The publication will solely be accessed by the various varieties of users once they are documented. A home Users will access the appliance interface so as to retrieve the meanings of straightforward words or sentences composed of a most of 3 words via linguistic communication. The appliance acknowledges the sort of user (Home User, Institute User, and supplier User) by registration information that's entered into the Mobile Application interface type (username and password). If there are straightforward words or

sentences introduced by the house user, the system sends Associate alert message to the computer user (Provider User) to insert the new which means of the word desired by the house user. During this manner the new linguistic communication video and animation generated within the info by supplier User. The Institute User will access the appliance via information recorded within the system by the Institute. There numerous activities are performed by the user, like inquiring concerning the meanings of words, sentence created from over two words, games, tutorials. Their cluster activities like chatting and discussions forums between the teams are recorded.

Mobile applications enable the user to manage the data flow and communication through the mobile device. These devices are customized or personalized. Second, mobile property improves collaboration via real time or instant interactivity, notwithstanding time and placement, resulting in higher deciding. Finally, mobile property enhances client orientation as users have higher access to their service suppliers and do a far better job in equalization their Work life through a productive use of your time. These edges will prove equally helpful for rising the educational. Therefore, mobile phones measure thought-about the simplest analysis tools for deaf and dumb students. The analysis method was broken into two phases. The primary section of the analysis is related to Home Student and second section to Institute Student. In initial section there is easy application like games to act between the house student and application, merely the user realize the way to confirm. In second section there is technique to judge the student's causes that learning should be accessed through performance like what students will do with their learning. Assessing students' performance will involve assessments that have high- or low-stakes, anonymous or public, formal or informal, individual or collective communication. Here we offer suggestions and techniques for assessing student learning and performance as ways in which they clarify their expectations and improve performance criteria of students. It takes into account the way to focus students' thinking in inventive, difficult, and motivating ways. It is totally different from the far side of the traditional assignment kind and should produce examination that meet the individual distinction in learning between students and also provides the ability of scholars to assimilation.

VIII. CONCLUSION

This model provides Mobile Application strategy for Deaf and Dumb sign language and to increase the efficiency of the operations in the Mobile Learning. The results of some studies show that it is advantageous to develop interactive and dynamic

content that to build Mobile Application. We have proposed new strategy to develop Mobile Learning Application by using data scalability, data integration and mobile frame work, and how to develop deaf and dumb e-content on any type of Mobile Device. This suggestion describes the required procedures to provide conversion of text to sign language for effective communication with deaf and dumb people. It is also a region independent language creator tool where new words and respective images could be added into the library by using web cam and mobile camera. The application takes text as an input and generates the output dynamically by searching many different images for every word and displays output by playing them with particular. It can be integrated with others mobile learning services like integrity services.

REFERENCES

- [1] <http://Www.Intexsoft.Com/Blog/Item/154-Mobile-Applications-For-Deaf-And-Dumb-People.Html>
- [2] <Https://Www.Mepits.Com/Project/288/Advanced-Projects/SMART-ASSISTANT-FOR-DEAF-AND-DUMB->
- [3] Majzoob Kamal Aldein Omer, Mohmed Sirelkhem Adrees, Osama E. Sheta, "Alternative Central Mobile Application Strategy toDeaf and Dumb Education in Third WorldCountries", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-5 Issue-3, July 2015
- [4] N.Hema , Ms. P. Thamarai , Dr. T.V.U. KiranKumar (2013), Handheld Deaf and Dumb Communication Device based on Gesture to
- [5] Voice and Speech to Image/Word Translation with SMS Sending and Language Teaching Ability .
- [6] Dalia Nashat, Abeer Shoker, Fowzyah Al-Swatand Reem Al-Ebailan , (2014) , An Android Application To Aid uneducated Deaf-Dumb People
- [7] Kimberly A. Weaver, Kimberly A. (2012), Mobile Sign Language Learning Outside the Classroom.

Data Mining Techniques For Intrusion Detection

Guided by: Ms. Rashmi Vipat

Jay Kotecha
TIMSCDR

Kruthika Hariharan
TIMSCDR

Abstract: With the dramatically development of web, security of network traffic is turning into a significant issue of computer network system. Attacks on the network square measure increasing day-by-day. The foremost publicised attack on network traffic is taken into account as Intrusion. Intrusion detection system has been used for ascertaining intrusion and to preserve the security goals of data from attacks. Dataprocessing techniques square measure want to monitor and analyze great deal of network knowledge; classify these network knowledge into normal and traditional knowledge. Since knowledge comes from varied sources, network traffic is giant. Data processing techniques like classification and agglomeration square measure applied to make Intrusion Detection System. A good Intrusion Detection System needs high detection rate, low warning rate additionally as high accuracy. This paper presents the review on IDS and totally different data processing techniques applied on Intrusion Detection System(IDS) for the effective detection of pattern for each malicious and traditional activities in network, that helps to develop secure information system.

Keywords: - *Intrusion Detection System, Anomaly Detection, Misuse Detection, Data mining, Clustering, Classifications*

I. INTRODUCTION

With the speedy increase of internet, there is enhancement among the manner of people but at the worth of threats, that unit of measurement created by either folks or any organization. They are accustomed break the protection of network. Security suggests that degree of protection given to the network or system. The foremost goals of security unit of measurement confidentiality, Integrity and convenience of data. Attacks on network can be referred as Intrusion. Intrusion suggests that any set of malicious activities that decide to compromise the security goals of the info. In youth, entirely typical approaches were used for network like cryptography, firewalls, virtual personal network etc but they weren't enough to secure network completely. It is difficult to bank completely on static defense techniques. This may increase the requirement for dynamic technique, which could be monitors system and confirm illegal activities. Thus to spice up the

network security dynamic approach is introduced and named as Intrusion Detection System. Intrusion detection System collects online data from the network then monitors and analyzes these data and partitions it into normal and malicious activities, supply the result to system administrator. IDS is that the space, wherever data processing is employed extensively, this is owing to restricted quantifiability, ability and validity. In IDS knowledge is collected from varied sources like network log data, host knowledge etc. Since the network traffic is massive, the analysis of information is just too exhausting. This bring about to the requirement of mistreatment IDS in conjunction with totally different data processing techniques for intrusion detection . Lee & amp; Salvatore J. Stolfo, Columbia University were first to use data processing techniques within the IDS [3]. Data processing techniques like classification and clump simply extract the knowledge from massive dataset. The remaining a part of the paper is structured during this approach. Section I introduction, Section II review the connected work on IDS mistreatment data processing techniques, Section III explanation of IDS. In Section IV, data processing and its techniques that are utilized in IDS ar delineated and at last Section V brings U.S.A. to the conclusion.

II. INTRUSION DETECTION SYSTEM

The thought of IDS was planned by Denning(1987), to identify, notice and trace the intrusion. Associate in Nursing IDS could be a combination of software system and hardware that area unit used for detecting intrusion. It gathers and analyzes the network traffic and notice the malicious patterns and at last awake to the proper authority. The most operate of IDS includes:

1. Observance and analyzing the data gathered from each user and system activities.
2. Understanding configurations of system and evaluating the file integrity and system integrity.

3. For static records, it finds out the normal pattern.
4. To acknowledge abnormal pattern, it use static records and awake to supervisor.

A. Classification of IDS.

According to techniques used for intrusion detection primarily based on whether or not attack's patterns area unit identified or unknown.

IDS classified into two class:

- (1) Misuse detection
- (2) Anomaly detection

Misuse detection: It's Signature primarily based IDS wherever detection of intrusion is predicated on the behaviors of identified attacks like antivirus software system. Antivirus software system compares the information with identified code of virus. In Misuse detection, pattern of identified malicious activity is hold on in the knowledge set and determine suspicious data by comparison new instances with the hold on pattern of attacks.

Anomaly detection: It is completely different from Misuse detection. Here baseline of traditional knowledge in network knowledge in network eg load on network traffic, protocol and packetsize etc is outlined by supervisor and according to this baseline, Anomaly detector monitors new instances. The new instances area unit compared with the baseline, if there is any deviation from baseline, knowledge is notified as intrusion. For this reason, it's conjointly referred to as behavior primarily based Intrusion Detection System.

Comparison Between Misuse Detection And Anomaly Detection.

Signature – Based(Misuse Detection)	Behaviour–Based(Anomaly Detection)
Advantages	Advantages
<ul style="list-style-type: none"> -Higher Detection rate, Accuracy for known behaviors. -Simplest and effective method. -Low False alarm rate. 	<ul style="list-style-type: none"> -can examine unknown and more complicated intrusions. - Rate of Missing report is low. -Detect new and unforeseen vulnerabilities.
Disadvantages	Disadvantages
<ul style="list-style-type: none"> - It can detect only known attacks. - Needs a regular update of the rules which are used. - Often no differentiation between an attack attempt and a successful attack. - Rate of Missing report is high. 	<ul style="list-style-type: none"> - Needs to be trained and tuned model carefully, otherwise it tends to false–positives -low detection rate and high false alarm rate. - It can't identify new attacks because intrusion detection depends upon latest model.

B. Working of Intrusion Detection System

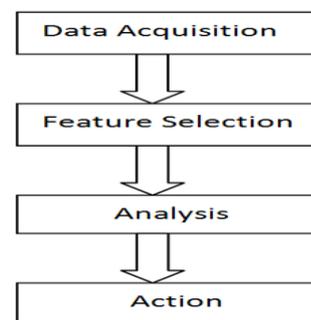
Author presents 4-steps for operating of IDS.

1) **Information Acquisition:** Information is collected from numerous sources by mistreatment explicit package.

2) **Feature Selection:** Brobdingnagian quantity of knowledge is collected from network traffic. Thus dataset for IDS becomes large. For performing on giant dataset generate feature vectors, that contains solely necessary information.

3) **Analysis:** During this step, collected information is analyzed to determine whether or not information is suspicious or not. Here, various data processing techniques are used for Intrusion detection.

4) **Action:** IDS alarms the administrator regarding attack which has been detected.



Operating Of Intrusion Detection System

C. Performance activity of IDS

There ar some primary factors that are used throughout performance activity of Intrusion Detection System.

True positive (TP): The whole variety of traditional information that are detected as a standard information throughout intrusion detection process.

True negative (TN): In Intrusion detection, number of detected abnormal information that are literally abnormal information in dataset.

False positive (FP): Or warning, total variety of detected traditional information however they're actual attack.

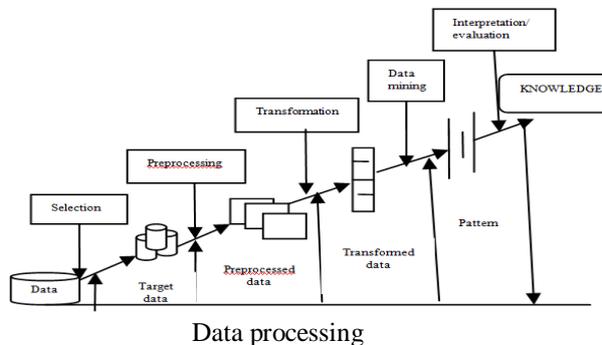
False negative (FN): Variety of detected abnormal instances however in real they're traditional information.

Performance of IDS is measured in terms of detection rate, accuracy and warning rate.

Detection Rate (DR) = $(TP/TP+FN) \times 100$ percent
 warning Rate (FAR) = $FP/\text{Number of Attacks}$
 Accuracy = $(TP+TN/TP + \text{Volunteer State} + FP + FN) \times 100\%$

III. DATA MINING BASED INTRUSION DETECTION SYSTEM

Data mining is that the activity of extracting relevant information from an oversized quantity of knowledge.



Network traffic is huge and data comes from different sources, that the dataset for IDS becomes massive. Hence the analysis of information is incredibly piece just in case of enormous dataset. data processing techniques square measure applied on IDS because it will extract the hidden data and dealswith massive dataset. Presently data processing techniques plays a vital role in IDS. By victimisation data processing techniques, IDS helps to discover abnormal and traditional patterns. This section describes totally different dataprocessing techniques such as agglomeration and classification, that square measure employed in IDS to get data regarding vulnerability by monitoring network knowledge.

A. Classification

Classification is that the task of taking every and each instances of dataset into account and assignment it to a particular category traditional and abnormal means that no table structure is employed for brand new instances. It will be effective for both misuse detection and anomaly detection, but more frequently used for misuse detection. Classification categorized the datasets into preset sets. It is less efficient in intrusion detection as compared to agglomeration. Different classification techniques like call tree, naive Bayes classifier, K-nearest neighbour classifier, Support vector machine etc square measure employed in IDS.

1. Call Tree- Decision tree could be a algorithmic and tree like structure for expressing classification rules. It uses divide and conquer method for cacophonous in line with attribute values.

Classification of the information yield from root node to leaf node, wherever every node represents the attribute and its value and amp; every leaf node represent category label of information. Tree based classifier have highest performance just in case of enormous dataset. Totally different call tree algorithms square measure delineated below.

ID3 formula- It is far-famed call tree formula developed by Quinlan. ID3 formula essentially attribute based mostly formula that constructs call tree in line with coaching dataset. The attribute that has highest data gain is employed as a root of the tree.

J48 formula -It is supported ID3 formula and developed by Ross Quinlan. In WEKA, C4.5 call tree formula is thought as J48 formula. It construct call tree victimization information gain, attribute that have highest datagain is chosen to form call. the most disadvantage of this formula is that it takes a lot of C.P.U. time and memory in execution. Another totally different tree based mostly classifier:

AD Tree -Alternating call tree is employed for classification. AD. Tree have prediction node as each leaf node and root node.

NB Tree- NB Tree formula uses each call tree and naive Bayes classifier. Root node uses call tree classifier and leaf nodes uses naive Bayes classifier.

Random Forest- Random Forest is initial introduced by Lepetit et.al. and it is ensemble classification technique that consists of two or more call trees. In Random Forest, each tree is prepared by indiscriminately choose the information from dataset. By using Random Forest improve the accuracy and prediction power as a result of it's less sensitive to outlier information. It can easily deals with high dimensional information.

2) K-Nearest Neighbor

It is one amongst the best classification technique. It calculates the space between totally different information points on the input vectors and assigns the untaged datum to it snearest neighbor category. K is a crucial parameter. If $k=1$, then the thing is appointed to the category of its nearest neighbor. once price of K is massive, then it takes massivetime for prediction and influence the accuracy by reduces the result of noise.

3) Naive mathematician classifier

Naive mathematician classifier is probabilistic classifier. It predicts the class in step with

membership chance. To derive conditional chance, it analyzes the relation between independent and variable quantity.

Bayes Theorem:

$$P(H/X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Where, X is that the information record and H is hypothesis that represents information X and belongs to category C. P(H) is that the previous probability, P(H/X) is that the posterior chance of H conditioned on X and P(X/H) is that the posterior chance of X conditioned on H.

Construction of Naive mathematician is simple with none complicated reiterative parameter. It is going to be applied to massive number of information points however time complexness will increase.

4) Support Vector Machine

Support Vector Machine is supervised learning methodology used for prediction and classification. It separate information points into two categories +1 and -1 victimisation hyperplane as a result of it is binary classification classifier. +1 represents traditional data and -1 for suspicious information.

Hyperplane may be expressed as: $W \cdot X + b = 0$

Where W= ar weight vector for n attributes A=, X= area attribute values and b may be a scalar. The most goal of SVM is to find a linear best hyper plane so the margin of separation between the two categories is maximized. The SVM uses some of the information to coach the system.

B) Cluster

Since the network knowledge is simply too vast, labelling of every and every instances or knowledge points in classification is pricey and time intense. Cluster is that the technique of labelling knowledge and assign into teams of comparable objects without victimisation better-known structure of knowledge points. Members of same cluster square measure similar and instances of various clusters are completely different from one another. Cluster technique may be classified into four groups: hierarchal rule, Partitioning rule, Grid based mostly rule and Density based rule. Some cluster algorithms square measure explained here.

1) K-Means cluster rule

K-Means cluster rule is simplest and wide used clustering technique projected by James Macqueen.

In this algorithm, variety of clusters K is such by user means that classifies instances into predefined variety of cluster. The first step of K-Means cluster is to decide on k instances as a center of clusters. Next assign every instances of dataset to nearest cluster. For example assignment, live the distance between center of mass and every instances victimization Euclidean distance and in keeping with minimum distance assign every and each knowledge points into cluster. K-Means algorithm takes less execution time, once it applied on small dataset. Once the information purpose will increase to most hen it takes most execution time. It is quick repetitious algorithm, however it's sensitive to outlier and noise.

2) K-Medoids cluster rule

K-Medoids is cluster by partitioning rule as like as K-means rule. The foremost centrally settled instance in an exceedingly cluster is taken into account as center of mass in situ of taking norm of the objects in K-Means cluster. This centrally set object is named reference and medoid. It minimizes the space between center of mass and data points means that minimize the square error. K-Medoids algorithm performs higher than K-Means algorithm once the quantity of knowledge points will increase to maximum. It is strong in presence of noise and outlier because medoid is a smaller amount influenced by outliers, but processing is costlier.

IV. CONCLUSION

On the idea of detection rate, accuracy, execution time and warning rate, the paper has analyzed totally different classification and agglomeration data processing techniques for intrusion detection. per given necessary parameter, execution time of Support vector machine is less and produces high accuracy with smaller dataset, while construction of Naive mathematician classifier is simple. Also decision tree has high detection rate just in case of huge dataset. In agglomeration techniques, execution time of K-Means clustering rule tiny amount just in case of small dataset, but once range of knowledge purpose will increase, K-Medoids performs higher.

REFERENCES

- [1] Deepthy K Denatious & Anita John, "Survey on Data Mining Techniques to Enhance Intrusion Detection", International Conference on Computer Communication and Informatics (ICCCI -2012), Jan. 10 - 12, 2012, Coimbatore, INDIA
- [2] Rung-Ching Chen, Kai-Fan Cheng and Chia-Fen Hsieh, "Using Rough Set And Support Vector Machine For Network Intrusion Detection", International Journal of Network Security & Its Applications (IJNSA), Vol 1, No 1, April 2009

- [3] Deepak Upadhyaya and Shubha Jain, "Hybrid Approach for Network Intrusion Detection System Using K-Medoid Clustering and Naïve Bayes Classification", IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 3, No 1, pp 231-236, May 2013
- [4] Xiang, M.Y. Chong and H. L. Zhu, "Design of Multiple-level Tree classifiers for intrusion detection system", IEEE conference on Cybernetics and Intelligent system, 2004
- [5] Peddabachigiri S., A. Abraham., C. Grosan and J. Thomas, "Modeling of Intrusion Detection System Using Hybrid intelligent systems", Journals of network computer application, 2007
- [6] Mrutyunjaya Panda and ManasRanjanPatra, "A Comparative Study Of Data Mining Algorithms For Network Intrusion Detection", First International Conference on Emerging Trends in Engineering and Technology, pp 504-507, IEEE, 2008.

Wireless Tracking System

Guided by: Ms. Rupali Jadhav

Khedkar Apurva

Mhatre Shraddha

Mungekar Aishwarya

TIMSCDR

TIMSCDR

TIMSCDR

Abstract: When children are moving in a wide area, such as a park or mall many times kids may be lost. Due to large size the security officers getting hard to locate the lost kids. This paper describes the design of children Tracking System. It is designed to track a moving child in a wide area, such as a park or mall, using Radio frequency identification terminology in wireless technology. Proposed system consist of both hardware and software. The software architecture consists of a communication driver that handles all communication functions done at the master station, Application Programming Interface (API) that handles and analyzes the data, GUI and a database that saves all readings and client information. RFID active tag, RFID tags reader, web server and database server forms the hardware architecture. The web server and database server are located in the master station. The tag readers are distributed around the open area, e.g. mall. The tags are programmed with kid's profiles and are worn by the kids. Using wireless LANs the tag reader and the web server communicate with each other.

Keywords: Radio Frequency Identification, RFID object tracking, Wireless system

I. INTRODUCTION

Radio Frequency Identification (RFID) is currently being used for auto-identification of objects, assets, pets, and people. RFID technology can be used in a variety of applications such as factory automations and integrations material tracking information systems, libraries management systems, hospital management systems, Airline baggage Identification and Pharmaceutical manufacturing. The basic architecture of an RFID system consists of a tag that includes an antenna and a chip, a reader equipped with antenna and a transceiver, and a workstation to host then Middleware and database.

The aim of this paper is to design and implement an RFID-based consistent and efficient solution to track lost kids in a large open area. The system allows users (e.g. security officers) to monitor the position of "tagged kids" from a sufficiently large distance by implementing an application that reveals the position of a kid at anytime and anywhere in the coverage area.

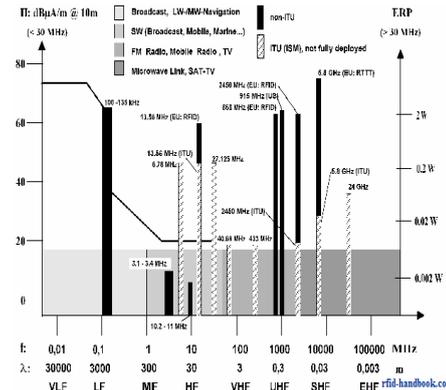


Fig 1: The frequency-ranges used for RFID-systems shown with the corresponding field strength and power levels.

II. REQUIREMENTS

A. Software

Communication: Analyze the connection between the Station and the reader (Master and Slave)

- Prepare the station(Master) packets(command)
- Send the packets (command)
- Receive the reader (Slave) packets (responds)
- Analyze the reader (Slave) packets (responds)
- Update the database.

The module function at the reader (Slave) side performs the following:

- Prepare the slave packet (commands)
- Send Packet (commands)
- Receive tags Packet (responds)
- Analyze tags Packets (responds)
- Receive station (Master) Packet (Commands)
- Analyze reader (Master) Packet command and act.

Application Programming Interface (API):

- Enable and disable tags.
- Get the status of the tags.

User Visualization Software:

- Show the information of the tags that the user gathered from the database.
- Trace the route of specific tags

Database :

- User Table : keep information about the tag users
- Readers Table: keep information about the readers and their location
- Detects Table: keep information about the tags, their users and the times when they were detected.
- Datasheet Table: keep track of location for each tag.

B. Hardware

The framework requires the use of Radio Frequency dynamic peruses with an Omni-directional receiving wire that gives a round scope zone of not less than 30 meters, or uni-directional reception apparatus that gives a straight scope region of 100 meters. The framework requires the utilization of RF dynamic peruses capable 3 diverse correspondence mediums: serial link, LAN link, and WLAN reception apparatus.

Readings: Reader should detect Reader ID and Tag Id at anytime.

Detecting the Child: if the child is in the coverage area of a reader, then the system is able to detect the child.

C. Tracing route and Detecting User

The system able to show the route taken by the child by plotting the last detected positions and connecting them. By specifying tag number we can delete the user who is leaving.

III. IMPLEMENTATION

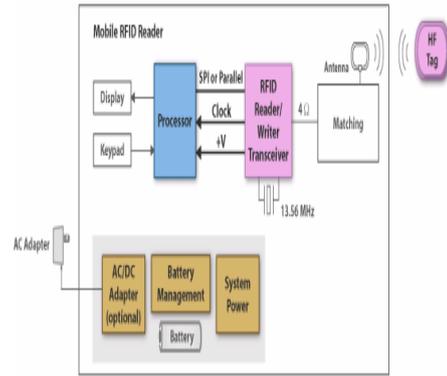
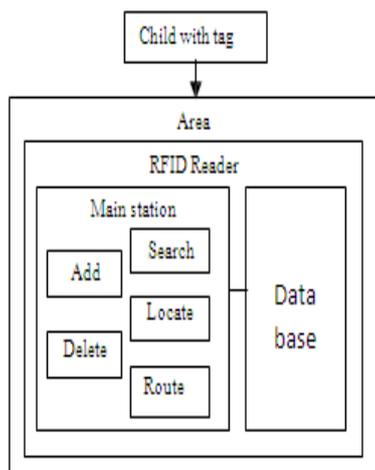


Fig2: Block diagram of RFID

At the passageway every guardian needs to fill the application form of children and guardian subtle element. All data is gone into database. Each child is given a RFID tag to wear it. At the point when the kid is missing, guardians need to report it to the security station. The label number relating to the child is recovered from the database and the label peruses are requested to find the lost child. As number of peruses are fix around the range and every peruse transmit the sign, by utilizing covering signals we can track the lost child.

IV. CONCLUSION

System will successfully work in crowdie area to track lost kids.

REFERENCE

[1] http://www.google.co.in/search?hl=en&source=hp&q=RFID+operation+frequency+ranges+&meta=&aq=f&aqi=&aql=&oq=&gs_rfai=

[2] <http://focus.ti.com/docs/solution/folders/print/447.html>

[3] Karen Coyle, "Management of RFID in Libraries," Journal of Academic Library.

Artificial Intelligence in Healthcare

Guided by: Ms. Madhulika B

Parikshit S Bhardwaj
TIMSCDR

Vivek Chaube
TIMSCDR

Sandeep Chauhan
TIMSCDR

Abstract—Artificial Intelligence is the science of making intelligent machines with the help of engineering, especially an intelligent computer programs. Artificial intelligence (AI), it involves two basic ideas. Firstly, it involves studying the thoughts processes of human a being. Secondly, it deals with by representing those processes via the machines (like computers, etc.). Artificial intelligence (AI) is the behavior of a machine, which, if performed by any human being, than would be stated as an intelligent work. The program makes the machines smarter and more useful, and it is less expensive rather than natural intelligence.

The medical field has already started to maintain their records digitally. Medical science has begun to digitalize their work & records. There can be many features of AI in medical such as a remote monitoring, helping high risk patients with remote AI, doctor's suggestions & guidance through AI voice feature.

I. INTRODUCTION

When we say Artificial Intelligence (AI), we think it as a machine which is capable or smart enough to take actions on its own and are as smart & responsive as a human being. Which can perform a task better than a human and without the help of a human? The major AI researchers define the field as "the study and the design of intelligent agent's", in which the intelligent agent is a system that perceives its environment and takes actions that maximize its chances of success. John McCarthy, who introduced the term in 1955, defines it as "The science and engineering of making intelligent machines".

The Natural Language Processing (NLP) refers to the artificial intelligence methods of communicating with computers in natural languages like English. The main purpose of NLP programs is to understand input & initiate actions.

II. ARTIFICIAL INTELLIGENCE

It's the science & engineering of making intelligent machines, especially an intelligent computer program. It is related to the similar task of using the computers to understand human intelligence, but

AI does not have to confine itself to the methods that are biologically observable.

Artificial Intelligence, intelligence However cannot be defined but AI can be described as a branch of computer science dealing with the simulation of machines exhibiting the intelligent behavior.

The Things involved in Artificial Intelligence is:

- The Ability to interact in the real world.
- Speech & Voice Recognition within a system.
- Learning and Adaption of things.

III. HEALTHCARE

Many medical facilities have adapted to digitalize their system & working functions as well as methods. All the primary activities whose primary function is to promote, maintain & restore a health.

Why do we need Healthcare systems? It can help us to deliver better treatments & prevention programs more effectively. If the priorities are set to the cost effective care, the overall population health may improve.

IV. CLAIMS

A system for remote monitoring of at least one patient by using telephone lines & artificial intelligence, through a monitoring device which is used by the patient, without the patients having to read the monitoring device said system comprising:

Record the patient's information including patients data, history & other relevant data for the treatment of the patient.

The plurality of physician's generated questions recorded in the voice of the patient's physician's voice are the questions which could be selectively asked to the patient in the physician's voice using voice technology through telephone line without a modem.

A programmed read only memory is programmed using artificial intelligence to select applicable questions & to parse the selected questions in a

sequence from the physician generated questions, based on the patient's response.

Without using a modem for obtaining and transmitting signals from said monitoring device over telephone lines to a computer integrated telephone system, where in by comparison of the transmitted signals with criteria predetermined by the physician, suitable intervention and action can be initiated.

In a method of remote monitoring by a physician through a monitoring device which is used by the patient, without the patient having to read the monitoring device. The improvement comprising of:

- a. Recording the patient's information including the patient data, its medical history & other information's relevant to treatment of the concerned patient.
- b. Recording physician's generated questions in the physician's own voice, which generated questions could be selectively asked of a patients being monitored.
- c. Using voice technology & creating a telephone monitoring encounter with the said patient to be monitored directly through the telephone lines without the modem to present to said patient selected ones of said physician generated questions.
- d. Using artificial intelligence (AI) for parsing said monitoring device using telephone lines selection of parsed questions based on progressive patient's response.
- e. Obtaining a reading signal from said monitoring device using telephone lines without a modem but using a telephone cord splitter.
- f. Applying physician's predetermined criteria for evaluating said progressive patient response. A method of remote monitoring at least one patient by using telephone & artificial intelligence through some monitoring device, comprising with the steps of:
- g. Recording the patient's information including the patient data, its medical history & other information's relevant to treatment of the patients concerned.

- h. Generating several questions which could be asked to the patients by an attending physician & which questions are of the physician's own choice.
- i. Creating a telephone monitoring encounter with the patient to be monitored in the doctor's own voice without the use of a modem.
- j. Using the Artificial Intelligence (AI) for the parsing described questions to create a selection & sequence of selected chosen questions by AI from said several questions based on progressive patients response.

At the end of said sequence of selected questions, obtaining a reading signal from the monitoring device through telephones lines without a modem & by using a telephone cord splitter.

Applying the physician's laid down criteria for evaluating said progressive patients response and the obtained reading signals to determine any course of actions necessary.

V. BACKGROUND OF INVENTIONS

Sometimes the patient's condition becomes a necessity to assure timely interventions by a healthcare practitioner or a physician to begin the right medical procedure or administer the required medication in a timely manner. Situations with high risk patients in the area of cardiology, obstetrics, neurology, psychology are some examples where artificial intelligence can help to take patients care necessity.

The patients care services are sometimes offered by health maintenance organizations. Statistics indicate that corporate healthcare benefits represented about the payroll. The part of corporate healthcare benefits are patients monitoring costs, & reducing such costs will be a very attractive proposition. In any event it is important to note that avoidance of clinically unnecessary outpatient visits can be a key to reducing healthcare costs, if insurance liability can be reduced and the clinical integrity preserved.

It is known in prior art to generate signals representing a patient's condition and record them for later scrutiny by a physician. An example can be the well known Holter heart monitor, wherein a continuous 24 hour cardiogram of a patient is recorded by a monitor which is worn by the patient. Functionally it is not alerted by any emergency situations which could occur during the duration of the monitoring.

The underlying considerations like hearing the doctor's voice during the monitoring service for the high risk patients include the facts that:

- The high risk patient's like to hear the doctor's voice during the monitoring interactions.
- The doctor has the responsibility to determine in a timely manner as to whether an appointment is necessary.
- The doctor should have the opportunity of determining if the high risk patient is in need of any urgent medication or any change in the treatment.
- It is often undesirable for monitoring to be done by patient intervention especially in situations like acute hypertension cases or serious non-ignorable cases.
- It is desirable to minimize the continuous use of a nurse or other healthcare worker to personally attend to the patient and do the monitoring.
- It is desirable if a physician could attend simultaneously and remotely to large number of high risk patients so that doctor intervention or other medical remedial measure could be initiated when necessary & no sooner.

Physicians who would use the inventive artificial Intelligence monitoring system would easily view the system as a superior clinical tool, since it frees them from being tied to one high risk patient at a time & also since they are not relying upon a patient to read a monitoring device. The system allows a physician user to be warned of critically ill patients automatically according to the physician's own choice guidelines as controlled by the AI.

VI. DESCRIPTION OF PREFERRED EMBODIMENT

The patient whose condition is monitored by monitoring team with artificial intelligence, which as aforesaid could take any form such as blood pressure cuff, pulse monitor, uterine activity strips, glucometer or electroencephalograph. If an infant is being monitored according to some artificial intelligence, the illustrated embodiment will need to be modified to the extent that a "conversation" with the patient may not be practical, but the system would still function to enable & ensure timely intervention by a healthcare person or a physician.

As developed, an Artificial Intelligence (AI) program directs the physician patient conversation through voice technology, measuring physical parameters near the end of the conversation. The doctors previously would have recorded questions

typically asked during a patient's examination; the AI program parses the questions and selects specific follow-up questions, depending upon the patient's response. The automatic monitoring (note, the patients is not required to read the device) may take place at the end of the conversations, where the AI system commands the voice system to play the physicians recorded request to place the monitor in the position and the recording function takes place. The patients monitoring device might use AT&T's dual tone matrix frequency standard for touch tone telephones, which DEC voice hardware (the voice Recognition technology) can recognize. An inexpensive encoding device is used to translate the analog signal coming from the home monitoring instruments to the telephone.

The Artificial Intelligence system is driven by a simple to use Natural Language interface which directs the voice system to speak, appropriate questions, recognize (listen) the patient's answers, update the patient's database. The patient's calls is available to the medical practitioner on both a real-time basis when the calls are being made, or on an ad-hoc basis after the data's are logged.

REFERENCES

- [1] https://scholar.google.co.in/scholar?q=Artificial+Intelligence+in+health&btnG=&hl=en&as_sdt=0%2C5
- [2] <https://www.google.co.in/#q=artificial+intelligence+in+healthcare+ppt>
- [3] <https://www.google.com/patents/US5357427>
- [4] https://www.google.co.in/search?biw=1280&bih=835&q=abstract+on+artificial+intelligence+in+ieeee+format&revid=462098419&sa=X&ved=0CHEQ1QIoAmoVChMIzBzY_OPuyAIV5S2mChIaCgEl#q=intro+to+health+care
- [5] http://cugh.org/sites/default/files/content/resources/modules/To%20Post%20Both%20Faculty%20and%20Trainees/34_Introduction_To_Comparative_Health_Care_Systems_FINAL.pdf
- [6] http://www.patientnavigatortraining.org/healthcare_system/index.htm
- [7] http://www.allianceforwaterefficiency.org/Medical_and_Health_Care_Systems_Introduction.aspx

Cloud Computing and DataBase

Guided by: Ms.Puja Agarwal

Vikas Pandey
TIMSCDR

Abstract— The Cloud is used as a Storage Location. Cloud database management system is a database that delivers computing as a service. What's more, the database can be gotten to and processed from any place. The substantial number of web application makes the utilization of dispersed stockpiling arrangement with a specific end goal to scale up. It empowers client to outsource the asset and administrations to the outsider server. It is sharing of web foundation for assets, programming and data over a system. In this paper, we talk about the late pattern in cloud administration in light of database administration framework and offering it as one of the administrations in cloud. We also proposed architecture of cloud based on database management system.

Keywords— *Cloud computing, DBMS, Database Management System, Databases in cloud computing, Different Models*

I. INTRODUCTION

Cloud Computing saves managing cost and time for business organizations. Industries, such as banking, healthcare and education are moving towards the cloud due to the efficiency of services provided by the pay-per-use pattern based on the resources such as processing power used, transactions carried out, bandwidth consumed, data transferred, or storage space occupied etc. Cloud computing is a totally web subordinate innovation where customer information is put away and keep up in the server from of a cloud supplier like Google, Amazon, Salesforce.com and Microsoft and so on. Distributed computing is design for giving registering administration by means of the web on interest and pay per use access to a pool of shared assets in particular systems, stockpiling, servers, administrations and applications, without physically procuring them. Constrained control over the information may acquire different security issues and dangers which incorporate information spillage, unreliable interface, sharing of assets, information accessibility and inside assaults. There are different examination challenges additionally there for receiving distributed computing, for example, very much oversaw administration level understanding (SLA), security, interoperability and unwavering quality.

The main idea of cloud computing is to deliver both software and hardware as services. In general cloud providers offer three types of services i.e. Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). There are various reasons for organizations to move towards IT solutions that include cloud computing as they are just required to pay for the resources on consumption basis. In addition, organizations can easily meet the needs of rapidly changing markets to ensure that they are always on the leading edge for their consumers.

Software as a Service (SaaS): SaaS can be described as a process by which Application Service Provider (ASP) provide different software applications over the Internet. This makes the customer to get rid of installing and operating the application on own computer and also eliminates the tremendous load of software maintenance. Examples of SaaS includes: Salesforce.com, Google Apps, YouTube, and Facebook.

Platform as a Service (PaaS): “PaaS is the delivery of a processing stage and arrangement stack as an administration without programming downloads or establishment for designers, IT chiefs or end-clients. It gives a base an abnormal state of combination keeping in mind the end goal to execute and test cloud applications. Case of PaaS incorporates: Force.com, Google App Engine and Microsoft Azure.

Infrastructure as a Service (IaaS): Infrastructure as a service (IaaS) refers to the sharing of hardware resources for executing services using Virtualization technology. Its main objective is to make resources such as servers, network and storage more readily accessible by applications and operating systems. Along these lines, it offers fundamental base on-interest administrations. A case of IaaS incorporates Amazon Flexible Distributed computing (EC2), Amazon S3, and Go Network.

II. KEY SECURITY ISSUES IN CLOUD COMPUTING

- a. Transmission of data.
- b. Security of data.
- c. Integrity of data.
- d. Location of data.
- e. Privacy of data.
- f. Reliability.
- g. Complexity.
- h. Secure Network.

III. MODELS OF CLOUD COMPUTING

Models of Cloud Computing are the one which allows Cloud Database to runs on Cloud Computing platform. Models are further divided into two common models:

A. Deployment Model

There are two important and primary methods to run a database on the cloud and they are:

- Database as a service (DBaaS) - Some Cloud stage offers database as an administration .In this Strategy, application proprietors don't need to introduce and keep up the database all alone yet rather the database administration supplier assumes liability for introducing and keeping up the databases, and application proprietors pay as indicated by their use. case, Amazon Web Administrations gives three database administrations as a major aspect of its cloud offering, Basic DB, Amazon Social Database Benefit, a SQL-based database administration with a MySQL interface, and Element DB. Additionally, Microsoft offers the Purplish blue SQL Database administration as a feature of its cloud advertising.
- Virtual machine Image – In this method of Deployment model cloud platform allow users to purchase virtual machine instances for a limited time. It is simple and conceivable to run a database on these virtual machines with those examples. Clients can either transfer their own machine picture with a database introduced on it, or use instant machine pictures that as of now incorporate establishment of a database. For example, oracle provides a ready-made machine image with an installation of Oracle Database 11g Enterprise Edition on Amazon EC2 and on Microsoft Azure.
- In this a third strategy which is not that critical but rather mulled over in which an alternative is overseen for database facilitating on the cloud, where the database is not offered as an administration, but rather the cloud database

seller has the database and oversees it for the application proprietor's benefit. For example, Cloud provider Rackspace offers managed hosting for MySQL on dedicated and cloud architectures and No SQL databases via Object Rocket's managed Mongo DB service. Similarly, on Azure, and Amazon Web Services, Mongo Lab provides Mongo DB-as-a-Service.

Architecture and common characteristics

- Most database administrations give electronic consoles, which the end client can use to procurement and set up together database object. For instance, the Amazon Web Administrations web console empowers clients to dispatch database cases, make previews (like reinforcements) of databases, and screen database measurements.
- Database services be formed of a database manager component, which control the underlying database object using a service API. The administration Programming interface is investigated to the end client, and license clients to perform support and scaling operations on their database objects. For instance, the Amazon Social Database's administration Programming interface empowers making a database occasion, changing the assets accessible to a database occurrence, erasing a database example, making a preview (like a reinforcement) of a database, and reestablishing a database from a depiction.
- Database services make the underlying software heap transparent to the user - the heap typically includes the operating system, the database and third-party software used to manage the database. The service provider (e.g. MongoLab or ObjectRocket) has the responsible for installing, patching and updating the underlying software heap and ensuring the overall health and performance of the database.
- Database administrations are capable of adaptability and high accessibility of the database. Adaptability highlights has diverse perspectives between merchants - some offer auto-scaling, others empower the client to scale up utilizing a Programming interface, yet don't scale consequently. There is normally procurement for a specific level of high accessibility.

B. Data Model

It is also important to differentiate between cloud databases which are relational as opposed to non-relational or NoSQL.

- SQL databases, such as PostgreSQL, NuoDB, Oracle Database, Microsoft SQL Server, and MySQL, are one type of database which can run in the cloud (either in a Virtual Machine Image or as a service, depending on the vendor). SQL databases are hard proportional, which means they are not locally suited to a cloud situation despite the fact that cloud database administrations in light of SQL have begun to address this test.
- NoSQL databases, such as Apache Cassandra, CouchDB and MongoDB, are another type of database which can run on the cloud. NoSQL databases are worked to administration substantial read/compose stacks and can scale all over effectively, and in this way they are all the more locally suited to running on the cloud. Be that as it may, most contemporary applications are worked around a SQL information model, so working with NoSQL databases frequently requires a complete change of utilization code.

IV. VENDORS

The following table lists notable database vendors with a cloud database offering, classified by their deployment model – machine image vs. database as a service – and data model, SQL vs. NoSQL.

	Virtual Machine Deployment	Database as a Service
SQL Data Model	<ul style="list-style-type: none"> • IBM DB2 • Ingres (database) • MySQL • NuoDB • Oracle Database • PostgreSQL 	<ul style="list-style-type: none"> • IBM dashDB data warehouse as a service • Amazon Relational Database Service • Clustrix Database as a Service • EnterpriseDB Postgres Plus Cloud Database • Heroku PostgreSQL as a Service (shared and dedicated database options) <ul style="list-style-type: none"> • Microsoft Azure SQL Database (MS SQL) Xeround Cloud Database* - MySQL front-end (*service)

		no longer available)
NoSQL Data Model	<ul style="list-style-type: none"> • Apache Cassandra on Amazon EC2 • Clusterpoint Database Virtual Box VM • CouchDB on Amazon EC2 • Hadoop on Amazon EC2 or Rackspace 	<ul style="list-style-type: none"> • Amazon DynamoDB • Amazon SimpleDB • Cloudata Data Layer (CouchDB) • Google App Engine Datastore MongoDB Database as a Service (several options)

A. Hadoop

- Is an entire ecosystem of integrated distributed computing tools, at the core of which file systems (HDFS) are and a programming framework.
- The Vanilla hadoop consists of a Distributed File System (DFS) at the core and libraries to support Map Reduce model to write programs to do analysis.
- It takes care of chunking data into multiple nodes in a multi node cluster so that Map Reduce can work on individual chunks of data available nodes thus enabling parallelism.

B. Cassandra

- Is a NoSQL data store based on a key-value pairing system, where value is then further structured into a columnar like store.
- Cassandra is a highly scalable, eventually consistent, distributed, structured key-value store. It is not a conventional database but is more like Hashtable or HashMap which stores a key/value pair.
- Cassandra works on top of HDFS and makes use of it to scale. Both Cassandra and HBase are implementations of Google's BigTable.
- BigTable makes utilization of a String Sorted Table (SSTable) to store key/esteem sets. SSTable is only a Record in HDFS which stores key took after by quality. Besides BigTable keeps up a file which has key and balanced in the

Document for that key which empowers perusing of worth for that key utilizing just a look to the counterbalance area. SSTable is viably changeless which implies subsequent to making the Document there is no adjustments should be possible to existing key/esteem sets. New key/esteem sets are added to the record. Redesign and Erase of records are affixed to the document, overhaul with a more current key/quality and erasure with a key and headstone esteem. Copy keys are permitted in this document for SSTable. The list is likewise adjusted with at whatever point upgrade or erase occur so that counterbalance for that key focuses to the most recent worth or gravestone esteem.

V. CONCLUSION

Database Management Systems runs as a scalable, elastic service available on a cloud infrastructure. It is less costly as compared to other computing systems Cloud DBMSs will have an impact for vendors desiring a less expensive platform for development.

- In this paper, we presented the idea of DBMS in the cloud, the possibilities to be offered as one of the services offered by promising capability of cloud computing, that is to be a DBMS as a Service.

- In this paper we have compared Appache/Casandra and Hadoop.
- Thus you can see Cassandra's internal allow fast read/write which is crucial for real time data handling. Whereas Vanilla Hadoop with Map Reduce can be used to process batch oriented passive data.

REFERENCE

- [1] [https:// www.Cloud_database.pdf](https://www.Cloud_database.pdf)
- [2] R. Marty, "Cloud Application Logging for Forensics", Proceedings of the ACM Symposium on Applied Computing, 2011
- [3] G. Sibiya, H. Venter and T. Fogwill, "Digital forensic framework for a cloud environment", Proceedings of the 2012 Africa Conference, 2012
- [4] A. Patrascu and V. Patriciu, "Beyond Digital Forensics. A Cloud Computing Perspective Over Incident Response and Reporting", International Symposium on Applied Computational Intelligence and Informatics, 2013

Text Mining in Healthcare

Guided by: Ms.Puja Agarwal

Mamta Patel
TIMSCDR

Vandana Patel
TIMSCDR

Abstract - This paper has talked about information digging for security applications. We initially began with an exchange of information digging for digital security applications and afterward gave a brief outline of the apparatuses we are creating. Information digging for national security and also for digital security is an extremely dynamic exploration area. Various information mining methods including join investigation and affiliation guideline mining are being investigated to distinguish anomalous examples. In view of information mining, clients can now make a wide range of connections. This additionally raises security concerns. One of the ranges we are investigating for future examination is dynamic guard. Here we are researching approaches to screen the enemies. For such observing to be powerful, the screen must maintain a strategic distance from location by the static and element examinations utilized by standard hostile to malware bundles.

Keywords: Text mining, Models of text mining, Application area in text mining.

I. INTRODUCTION

Industries today experience structured information (both semi-organized and unstructured) while maintaining their everyday business. The wellspring of the information could be electronic content, call focus logs, online networking, corporate reports, research papers, application shapes, administration notes, messages, and so forth. This information might be open yet stays undiscovered because of the absence of consciousness of the data riches an association has or the absence of philosophy or innovation to investigate this information and get the valuable knowledge.

The term text analytics describes a set of linguistic, statistical and machine learning techniques that model and structure the information content of textual sources for business intelligence, exploratory data analysis, research, or investigation. The term is roughly synonymous with text mining; indeed, Ronen Feldman modified a 2000 description of "text mining" in 2004 to describe "text analytics. The latter term is now used more frequently in business settings while "text mining" is used in some of the earliest application

areas, dating to the 1980s, notably life-sciences research and government intelligence.

The term content investigation likewise depicts that use of content examination to react to business issues, whether freely or in conjunction with question and investigation of handled, numerical information. It is a cliché that 80 percent of business-significant data starts in unstructured structure, principally message. These procedures and procedures find and present learning – realities, business guidelines, and connections – that is generally secured printed structure, impervious to mechanized preparing.

II. TEXT MINING

Essentially put content mining is the learning disclosure from literary information or printed information investigation to reveal valuable yet shrouded data. Be that as it may, numerous individuals have characterized content mining somewhat in an unexpected way. The accompanying is a couple of definitions:

“The objective of Text Mining is to exploit information contained in textual documents in various ways, including ...discovery of patterns and trends in data, associations among entities, predictive rules, etc.”

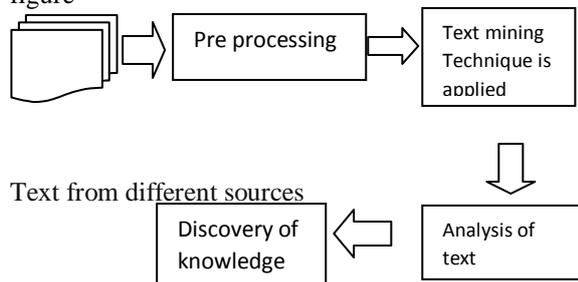
“Another way to view text data mining is as a process of exploratory data analysis that leads to heretofore unknown information, or to answers for questions for which the answer is not currently known.”

Text mining also known as text data mining or text analytics is the process of discovering high quality information from the textual data sources. The application of text mining techniques to solve specific business problems is called business text analytics or simply text analytics. Text mining techniques can facilitate organizations derive valuable business insight from the wealth of textual information they possess.

Text mining changes printed information into organized arrangement using a few systems. It includes recognizable proof and accumulation of the literary information sources, NLP procedures like grammatical feature labeling and syntactic parsing, substance/idea extraction which distinguishes named highlights like individuals, spots, associations, and so forth., disambiguation, building up relationship between various elements/ideas, example and pattern examination and perception systems.

TEXT MINING PROCESS

Text mining process is as shown in following figure



Text mining process starts with a document collection from various resources. Text mining tool would retrieve a particular document and pre-process it by checking format and character sets. Then document would go through a text analysis phase. Text analysis is semantic analysis to derive high quality information from text. Many text analysis techniques are available; depending on goal of organization combinations of techniques could be used. Sometimes text analysis techniques are repeated until information is extracted. The resulting information can be placed in a management information system, yielding an abundant amount of knowledge for the user of that system. Text mining process Rest of this paper presents challenging issues, merits and demerits, methods, and techniques of text mining.

III. METHODS AND MODELS USED IN TEXT MINING

Traditionally there are so many techniques developed to solve the problem of text mining that is nothing but the relevant information retrieval according to user's requirement. According to the information retrieval basically there are four methods used:

- A. Term Based Method (TBM).
- B. Phrase Based Method (PBM).
- C. Concept Based Method (CBM).
- D. Pattern Taxonomy Method (PTM).

A. Term Based Method (TBM).

Term in archive is word having semantic significance. In term construct technique archive is broke down in light of the premise of term and has preferences of productive computational execution and also develops hypotheses for term weighting. These procedures are risen in the course of the last couple of decades from the data recovery and machine learning groups. Term based techniques experience the ill effects of the issues of polysemy and synonymy. Polysemy implies a word has different implications and synonymy is various words having the same importance. The semantic significance of numerous found terms is questionable for noting what clients need. Data recovery gave numerous term-based techniques to explain this test.

B. Phrase Based Method

Phrase carries more semantics like information and is less ambiguous. In phrase based method document is analyzed on phrase basis as phrases are less ambiguous and more discriminative than individual terms. The likely reasons for the daunting performance include:

- Phrases have inferior statistical properties to terms,
- They have low frequency of occurrence, and
- Large numbers of redundant and noisy phrases are present among them.

C. Concept Based Method

In concept based terms are analyzed on sentence and document level. Text Mining techniques are mostly based on statistical analysis of word or phrase. The statistical analysis of the term frequency captures the importance of word without document. Two terms can have same frequency in same document, but the meaning is that one term contributes more appropriately than the meaning contributed by the other term. The terms that capture the semantics of the text should be given more importance so, a new concept-based mining is introduced This model included three components. The first component analyzes the semantic structure of sentences. The second component constructs a conceptual ontological graph (COG) to describe the semantic structures and the last component extract top concepts based on the first two components to build feature vectors using the standard vector space model. Concept-based model

can effectively discriminate between non important terms and meaningful terms which describe a sentence meaning. The concept-based model usually relies upon natural language processing techniques. Feature selection is applied to the query concepts to optimize the representation and remove noise and ambiguity.

D. Pattern Taxonomy Method

In pattern scientific categorization strategy records are broke down on example premise. Examples can be organized into scientific classification by utilizing is-a connection. Design mining has been broadly concentrated on in information digging groups for a long time. Examples can be found by information mining methods like affiliation principle mining, visit thing set mining, consecutive example mining and shut example mining. Utilization of found learning (designs) in the field of content mining is troublesome and insufficient, on the grounds that some helpful long examples with high specificity need in backing (i.e., the low-recurrence issue). Not all successive short examples are valuable subsequently known as misinterpretations of examples and it prompts the inadequate execution. In examination work, a viable example revelation procedure has been proposed to conquer the low-recurrence and distortion issues for content mining. The example based procedure utilizes two procedures design sending and example developing. This strategy refines the found examples in content reports. The trial results demonstrate that example based model performs superior to not just other immaculate information mining-based techniques and the idea based model, additionally term-based models.

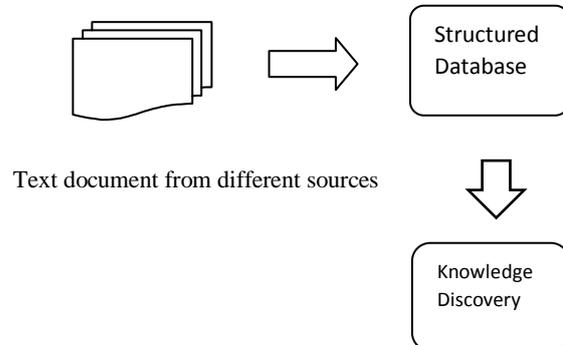
IV. TECHNIQUES USED IN TEXT MINING

To show how to examine, comprehend and create content, innovations are delivered by normal dialect preparing. The advancements like data extraction, trashing, arrangement, bunching and data perception are utilized as a part of the content mining process. In the accompanying areas we will talk about each of these innovations and the part that they play in content mining. The sorts of circumstances where every innovation might be valuable with a specific end goal to help clients are additionally talked about.

A. Information Extraction

Data extraction is beginning stride for PC to dissect unstructured content by distinguishing key expressions and connections inside content. To do this errand procedure of example coordinating is utilized to search for predefined arrangements in

content. Data extraction undertaking incorporates tokenization, distinguishing proof of named elements, sentence division, and grammatical form task. Firstly expressions and sentences are parsed and semantically translated then required bits of data went into the database. General information extraction process is as shown in figure



The most accurate information extraction systems involve handcrafted language processing modules, substantial progress has been made in applying data mining techniques to a number of these steps. This technology can be very useful when dealing with large volumes of text. For many applications challenging is electronic information is in the form of free natural language documents rather than structured databases like relational databases. Information extraction solves this problem of transforming a corpus of textual documents into a more structured database. For further mining of knowledge database constructed by an information extraction module can be provided to the KDD module.

B. Categorization

Classification is managed learning technique since it depends on information yield case to group new reports. Predefined classes are appointed to the content archives in light of their substance. The common content arrangement process comprises of pre-preparing, indexing, dimensionally diminishment, and grouping. The objective of order is to prepare classifier on the premise of known cases and after that obscure cases are arranged naturally. Factual order strategies like Innocent Bayesian classifier, Closest Neighbour classifier, Choice Tree, and Bolster Vector Machines can be utilized to arrange content.

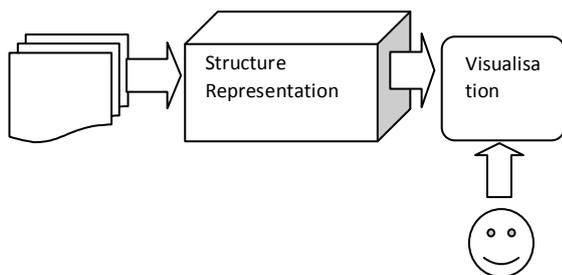
C. Clustering

Clustering method can be used in order to find groups of documents with similar content. The outcome of clustering is typically a partition called clusters P and each cluster consists of a number of

documents d. The contents of the documents within one cluster are more similar and between the clusters more dissimilar then the quality of clustering is considered better. Even though clustering technique used to group similar documents it differs from categorization because in clustering documents are clustered on the fly instead of use of predefined topics. As documents can appear in multiple subtopics clustering ensures that a useful document will not be omitted from search results. In data mining K-means is frequently used clustering algorithm; in text mining field also it obtains good results. A basic clustering algorithm creates a vector of topics for each document and measures the weights of how well the document fits into each cluster. The organization of management information systems makes use of clustering technology as organizational database contain thousands of documents.

D. Visualization

In text mining visualization methods can improve and simplify the discovery of relevant information. To represent individual documents or groups of documents text flags are used to show document category and to show density colours are used. Visual text mining puts large textual sources in a visual hierarchy. The user can interact with the document by zooming and scaling. Information visualization is applicable to government to identify terrorist networks or to find information about crimes. Following figure shows steps involved in visualization process.

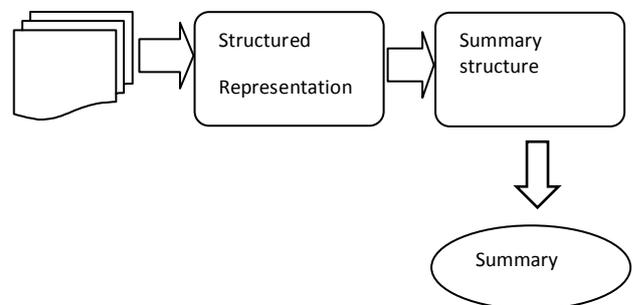


The goal of information visualization divided into three steps:

- Data preparation step includes deciding and obtaining original data of visualization and form original data space.
- The process of analyzing and extracting visualization data needed from original data and to form visualization data space is known as Data analysis and extraction.
- Visualization mapping step employ certain mapping algorithm to map visualization data space to visualization target.

E. Summarization

Text summarization is to lessen the length and subtle element of a report while holding most essential focuses and general significance. Content synopsis is useful for to make sense of regardless of whether a long report addresses the client's issues and merits perusing for additional data henceforth rundown can supplant the arrangement of records. In the time taken by the client to peruse the primary section content rundown programming forms and compresses the huge content record. It is difficult to teach software to analyze semantics and to interpret meaning of text document even though computers are able to identify people, places, and time. Humans first reads entire text section to summarize then try to develop a full understanding, and then finally write a summary, highlighting its main points.



Summarization process include following steps:

1. Pre-processing obtain a structured representation of the original text.
2. To transform summary structure from text structure algorithm is applied in next processing step.
3. In the invention step the final summary is obtained from the summary structure.

V. TEXT MINING APPLICATION IN HEALTHCARE

Several research studies have focused on the processing of textual information available in healthcare datasets.

A brief overview of studies that highlight the significance of textual data and its suitability in research settings is presented here.

One striking examination activity in was performed at the Venderbilt Clinic, New York [1]. The goal was to figure out whether a characteristic dialect handling program (NLP) could naturally code practical status data as per the International Classification of Functioning, Disability, and Health (ICF) necessities. Mechanized coding is a

conspicuous decision for these sorts of activities. Coding is critical for repayment purposes and record keeping; be that as it may, it is additionally an exceptionally monotonous and tedious procedure. In the event that this could be expert precisely with innovation it would spare the therapeutic offices a considerable measure of assets. The researchers extended the existing NLP MedLEE to code rehabilitation discharge summaries. Ten ICD-9 codes were pre-selected for their known relationship to changes in functional status. Evaluations were performed by the NLP system, expert coders, and non-expert coders. They found that the NLP system coded with similar results to the human coders. This is a promising finding for research into automated coding for ICD-9 codes, which are the main basis for reimbursement, in majority of healthcare settings. A study, directed at the University of Utah [2], utilized an adjusted rendition of MedLEE and additionally an expression coordinating calculation to concentrate information for examination activities. Most electronic records are managed in a story structure and physically recovering particular information for examination can be tedious and costly. The reason for this study was to concentrate information identified with unfavourable occasions associated with focal venous catheter arrangement. Adverse events can be things such as infections, complications from misplacement, and pneumothorax (a collapsed lung). Tests were conducted using each method individually and then using them together on a sample of records that had been manually reviewed beforehand. The trials using the individual methods were unsuccessful. The phrase-matching algorithm was not specific enough and the NLP system was not sensitive enough. They created positive forecast estimations of 6.4 and 6.2% separately. Be that as it may, when utilized together the outcomes were promising. They yielded a 72.0% affectability and a 80.1% specificity which are satisfactory qualities. This study indicates potential for utilizing NLP frameworks to robotize research information extraction. Occasion identification is another noteworthy range of examination. Hazlehurst et. al [3] preformed a study to recognize antibody responses for the Vaccine Safety Datalink Project (VSD). The VSD is an association between the CDC and eight substantial HMO's to examine unfavorable occasions taking after inoculation. They are endeavoring to do this by dissecting medicinal consideration databases and patient restorative records. In this study, a changed rendition of the NLP framework MediClass which had been prepared with the learning important to

identify conceivable immunization responses was utilized. It achieved both a high sensitivity and specificity percentage. Compared with methods that are used by clinicians this system significantly improved the positive predictive value. Studies such as these are especially important because the ultimate goal is to migrate to a system that can predict such occurrences in future.

Recently, text mining tools have been utilized in healthcare research, e.g. Cerrito and Cerrito [4] analyzed the electronic medical records from the emergency department of a hospital over a six month period, using text mining. They found that similar complaints were treated differently depending on the physician on call. Such differences can affect care quality and costs. Therefore text mining of prior expert treatment can provide physicians on call with an optimized treatment plan. It can also lead to development of protocols to alleviate disparity in treatment.

VI. CONCLUSION

This paper has presented overview techniques, methods and challenging issue in text mining. The focus has been given on fundamental methods for conducting text mining. The paper also addressed the most challenging issue in developing text mining systems. Four methods of text mining term based, phrase based, concept based and pattern taxonomy model discussed. Term based approach suffer from polysemy and synonymy while phrase based approach performs better as phrase carries more semantics like information and is less ambiguous. Two terms can have same frequency from statistical analysis this problem can be solved by concept based approach by finding term contributing more meaning. In pattern based approach pattern taxonomy is formed to solve low frequency problem and misinterpretation problem. Then in next half Text Mining is discussed with its various techniques and usages. To extract structured information from the unstructured text Information Extraction is used. In information extraction data mining techniques can be applied for getting useful patterns or knowledge from the documents. To produce the relevant information from the corpus is known as summarization. Classification is a supervised technique because before it can be used to classify the newly arrived document it has all the input output patterns which are used to train the model. Clustering is unsupervised learning technique because no pre-defined input-output patterns are there. According to summery of documents text is clustered, this process is known as clustering. To provide improved understandable information for mining the documents Graphical Visualization is used.

REFERENCES

- [1] Evans David. 2000. "Advanced text mining tools help CRM systems "see" customers". Unisys World.
- [2] Landauer T. K., Laham D., Foltz P. W. 1998.
- [3] "Learning human-like knowledge by singular value decomposition: a progress report". Advances in Neural Information Processing Systems 10, M. I. Jordan,
- [4] M. J. Kearns, and S. A. Solla, Eds., Cambridge: MIT Press, 1998, pp. 45-51.
- [5] Landauer T. K., Foltz P. W., Laham D. (1998). "Introduction to Latent Semantic Analysis". Discourse Processes, 25, pp. 259-284.
- [6] Dumais S. 1990. "Enhancing Performance in Latent Semantic Indexing (LSI) Retrieval". TM-ARH-017527 Technical Report, Bellcore. <http://citeseer.nj.nec.com/dumais92enhancing.html>

Digital Immortality

Guided by: Mr. Pankaj Mudholkar

Protima Paul
TIMSCDR

Nikhil Pednekar
TIMSCDR

Abstract: Progressed heavenliness or "virtual endlessness" is the hypothetical thought of securing or trading a man's personality in more solid media, i.e., a PC, and allowing it to talk with people later on. The result may take after an image continuing, reacting, and having a comparative viewpoint as a man on the reason of that individual's propelled document. After the death of the individual, this image could stay static or continue to learn and develop self-rulingly. A broad piece of transhumanists and singularitarians place unprecedented trust into the conviction that they may get the chance to be wearisome by the year 2045, by making one or various non-natural useful copies of their brains, in this way leaving their "normal shell". These copies may then "live everlastingly" in an adjustment of cutting edge "heaven" or paradise.

I. INTRODUCTION

Humans have been chasing immortality for millennia. In some cultures, you attain a kind of immortality by doing great deeds, which people will talk about long after you pass away. Several religions feature some concept of immortality -- the body may die but some part of you will exist forever. But what if science made it possible to be truly immortal? What if there were ways for you to live forever? That is the basic concept behind digital immortality. A few futurists, maybe most prominently creator Ray Kurzweil, trust that we will reveal an approach to augment the human lifespan inconclusively. They have recognized a few potential ways that could prompt this destination. Maybe we will recognize the qualities that administer maturing and change them so that our bodies quit maturing once they achieve development. Perhaps we will make new systems for making simulated organs that join natural matter with innovation and after that supplant our unique parts with the as good as ever forms. On the other hand possibly we will simply dump our recollections, considerations, sentiments and everything else that makes us who we are into a PC and live in the internet.

II. THE REALISM OF THE CONCEPT

The futurist Ian Pearson trusts that people will accomplish a sort of virtual everlasting life by sparing into PCs by the year 2050. The National Science Foundation has recompensed a half-million-dollar stipend to the colleges of Central Florida at Orlando and Illinois at Chicago to investigate how scientists may utilize counterfeit consciousness, documenting, and PC imaging to make persuading, computerized forms of genuine individuals, a conceivable initial move toward virtual everlasting life. The Digital Immortality Institute investigates three elements fundamental for advanced godlikeness. To start with, at whatever level of usage, symbols require ensured Internet openness. Next, symbols must be what clients indicate, and they should remain so. At long last, future representations must be secured before the living clients are no more. The point of Dmitry Itskov's 2045 Initiative is to "make innovations empowering the exchange of an individual's identity to a non-natural bearer, and broadening presence, including to the point of eternity".

III. METHODS

Reaching digital immortality is a three-step process:

- A. archiving and digitizing people,
- B. making the avatar live
- C. calibration process

A. Archiving and digitizing people

As per Gordon Bell and Jim Gray from Microsoft Research, holding each discussion that a man has ever heard is as of now sensible: it needs not exactly a terabyte of capacity (for satisfactory quality). The discourse or content acknowledgment innovations are one of the greatest difficulties of the idea.

A second possibility would be to archive and analyze social Internet use to map the personality of people. By analyzing social Internet use during 50 years, it would be possible to model a society's culture, a society's way of thinking, and a society's interests.

Rothblatt envisions the creation of "mind files" -- collections of data from all kinds of sources, including the photos we upload to Facebook, the

discussions and opinions we share on forums or blogs, and other social media interactions that reflect our life experiences and our unique self.

Richard Grandmorin compressed the idea of computerized everlasting life by the accompanying condition: "semantic investigation + social web use + Artificial Intelligence = godlikeness".

Some find that photos, videos, sound clips, social media posts and other data of one could already be regarded as such an archiving

B. Making the avatar alive

Defining the avatar to be alive allows it to communicate with the future in the sense that it continues to learn, evolve and interact with people, if they still exist. Technically, the operation exists to implement an artificial intelligence system to the avatar. This artificial intelligence system is then assumed to think and will react on the base of the archive.

Rothblatt proposes the expression "mind product" for programming that is being created with the objective of creating cognizant AIs. Such programming would read a persons' "psyche document" to produce a "brain clone". Rothblatt additionally proposes a specific level of administrative endorsement for psyche product, similar to a FDA accreditation, to guarantee that the subsequent personality clones are well made.

C. Calibration process

Amid the alignment procedure, the natural individuals are inhabiting the same time as their relic in silicon. The ancient rarity in silicon is aligned to be as close as could be expected under the circumstances to the individual being referred to. During this process ongoing updates, synchronization, and interaction between the two minds would maintain the twin minds as one.

IV. THE SINGULARITY IS NEAR: MIND UPLOADING BY 2045

By 2045, "based on conservative estimates of the amount of computation you need to functionally simulate a human brain, we will be able to expand the scope of our intelligence a billion-fold," Kurzweil said. Itskov and other so-called "transhumanists" interpret this impending singularity as digital immortality. Specifically, they believe that in a few decades, humans will be able to upload their minds to a computer, transcending the need for a biological body. The idea sounds like sci-fi, and it is — at least for now. The reality, however, is that neural engineering is making significant strides toward modeling the brain and developing technologies to restore or replace some of its biological functions.

A. Brain prostheses

Significant accomplishments have been made in the field of mind PC interfaces, or BCIs (additionally called cerebrum machine interfaces). The cochlear insert — in which the mind's cochlear nerve is electronically fortified to reestablish a feeling of sound to somebody who is almost deaf — was the primary genuine BCI. Numerous gatherings are currently creating BCIs to reestablish engine aptitudes, taking after harm to the sensory system from a stroke or spinal string damage. José Carmena and Michel Maharbiz, electrical architects at the University of California, Berkeley, are attempting to create best in class engine BCIs. These gadgets comprise of pill-size anode exhibits that record neural signs from the cerebrum's engine ranges, which are then decoded by a PC and used to control a PC cursor or prosthetic appendage, (for example, an automated arm). Carmena and Maharbiz talked about the test of making a BCI that works steadily after some time and does not require being fastened to wires. Theodore Berger, a neural architect at the University of Southern California in Los Angeles, is taking BCIs to another level by building up a memory prosthesis. Berger plans to supplant part of the cerebrum's hippocampus, the area that proselytes transient recollections into long haul ones, with a BCI. The gadget records the electrical action that encodes a straightforward fleeting memory, (for example, pushing a catch) and changes over it to a computerized signal. That flag is gone into a PC where it is numerically changed and after that bolstered over into the mind, where it gets fixed in as a long haul memory. He has effectively tried the gadget in rats and monkeys, and is currently working with human patients.

B. Mind uploading

The conference took a surreal turn when Martine Rothblatt — a lawyer, author and entrepreneur, and CEO of biotech company United Therapeutics Corp. — took the stage. Even the title of Rothblatt talk was provocative: "The Purpose of Biotechnology is the End of Death." Rothblatt introduced the concept of "mind clones" — digital versions of humans that can live forever. She described how the mind clones are created from a "mind file," a sort of online repository of our personalities, which she argued humans already have (in the form of Facebook, for example). This psyche record would be keep running on "brain product," a sort of programming for awareness. "The main organization that creates mind product will have [as much achievement as] a thousand Google's," Rothblatt said. Be that as it may, would such a psyche clone be alive? Rothblatt thinks so. She referred to one definition as a self-reproducing code that keeps up itself against turmoil. A few faultfinders have evaded what Rothblatt called

"spooky Cartesian dualism," contending that the psyche must be implanted in science. Actually, programming and equipment are in the same class as wet product, or organic materials, she contended. Rothblatt went ahead to talk about the ramifications of making brain clones. Continuity of the self is one issue, because your persona would no longer inhabit just a biological body. Then, there are mind-clone civil rights, which would be the "cause célèbre" for the 21st century, Rothblatt said. Even mind clone procreation and death were mentioned.

C. The quantum world

In parallel with the talk of brain technologies and mind-uploading, much was said about the nature of consciousness in the universe. Physicist Roger Penrose of the University of Oxford and others disagree with the interpretation of the brain as a mere computer. Penrose argued that consciousness arising from the fabric of the universe. Those of the "Penrose school" think uploading the brain would have to involve quantum computers — a development unlikely to happen by 2045. But Itskov thinks otherwise. The 32-year-old president of the Global Future 2045 Congress is dead set on living forever.

V. CONCLUSION

These predictions may sound like they belong in a science-fiction film, but there are people all over the world who are contributing work that may one day allow humans to live forever. Some of the work relates directly to the goal -- scientists at Harvard have succeeded in altering genes that regulate aging in mice, for example [source: **Sample**]. If it's possible for scientists to apply these techniques to humans, we may be able to prevent our bodies from aging and live to ages far beyond the average lifespan of today. Different tasks like Blue Brain may in the long run give different bits of knowledge into accomplishing

advanced everlasting life. The Blue Brain's task will likely figure out the human cerebrum and make a virtual model that would permit neuroscientists to test strategies and medications on an advanced mind to perceive how a genuine human mind may respond [source: The Blue Brain Project]. Maybe through this work we'll accomplish a more noteworthy comprehension of how the human cerebrum functions and its relationship to what we call the brain. Is it conceivable to figure out the procedure of supposing by building a complex virtual model of a mind?

Ray Kurzweil has written extensively about this concept. He foresees a diverse set of disciplines approaching the problem through different methods, some of which may ultimately converge and convert digital immortality from a concept into reality. He identified three bridges that could provide us a way to live as long as we care to.

REFERENCES

- [1] https://en.wikipedia.org/wiki/Digital_immortality
- [2] electronics.howstuffworks.com/gadgets/high.../digital-immortality.htm
- [3] http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1252799&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D1252799
- [4] www.bbc.com/future/story/20150122-the-secret-to-immortality
- [5] <http://betanews.com/2016/02/26/digital-immortality-how-to-create-an-eternal-copy-of-yourself/>
- [6] <http://knowledge.wharton.upenn.edu/article/digital-immortality-cyber-consciousness-mind-clones-and-the-future-of-humanity/>
- [7] <http://knowledge.wharton.upenn.edu/article/digital-immortality-cyber-consciousness-mind-clones-and-the-future-of-humanity/>
- [8] <http://www.livescience.com/37499-immortality-by-2045-conference.html>
- [9] <http://www.livescience.com/37499-immortality-by-2045-conference.html>
- [10] <http://www.amazon.in/Virtually-Human-Promise-Digital-Immortality/dp/1491532912>

Use of Big Data in Social Networking

Guided by: Mr. Pankaj Mudholkar

Abhinav Singh
TIMSCDR

Randhir Singh
TIMSCDR

Rakhi Singh
TIMSCDR

Abstract: We have decided to go with this topic to learn more about emerging technology “BIG DATA” and to see how this thing is playing important part in our daily social life. According to a recent report, the amount of digital content on the Internet is now close to five hundred billion gigabytes. This number is expected to double within a year. The explosion of mobile networks, cloud computing and new technologies have given rise to incomprehensibly large worlds of information, often described as “Big Data”. We describe how the analysis of Twitter content can reveal mood changes in entire populations, how the political relations among Indian leaders can be extracted from social networking platform, how we can determine what news people really want to read.

Keywords: big data, social, analysis, process, node XL.

I. INTRODUCTION

The ready availability of masses of data and the means to exploit them is changing the way we do science in many domains.

What do you mean by big data?

Extract insight from a high volume, variety, velocity and veracity of data in a timely and cost-effective manner.

Variety: Manage and benefit from diverse data types and data structures.

Velocity: Analyze streaming data and large volumes of persistent data

Volume: scale from terabytes to zetta bytes.

Veracity: Establish confidence in data, information and solutions.

How do we measure the value of big data? When we set out to size the potential of big data to create value, we considered only those actions that essentially depend on the use of big data—i.e., actions where the use of big data is necessary (but usually not sufficient) to execute a particular lever. We did not include the value of levers that consist only of automation but do not involve big data.

Face book, twitter runs the largest social networking platform that serves hundreds of millions users at peak times using tens of thousands of servers located in many data centers around the world. There are strict operational requirements on their platform in terms of execution, unwavering

quality and proficiency, and to bolster ceaseless development the stage should be very versatile. Managing disappointments in a base included a great many parts is our standard method of operation; there are dependably a little however huge number of server and system segments that are falling flat at any given time. Thusly, the product frameworks should be developed in a way that regards disappointments as the standard as opposed to the special case.

II. LITERATURE OF SURVEY

Measuring the current public mood is a challenging task. The traditional approach would require questioning a large number of people about their feelings. Social media, such as Twitter or Face book, can easily become a valuable source of information about the public due to the fact that people use them to express their feelings in public.

As showed in our study it is practical to catch the general population temperament by observing the flood of Twitter information. The dataset that was dissected was involved millions tweets that were created by a large number of clients on everyday schedule. The information is gathered from the everywhere throughout the world. We concentrated on following four mind-sets which are "Trepidation", "Satisfaction", "Annoyance" and "Misery". For every state of mind, we track a considerable rundown of related words and we check the frequencies that these words show up in tweets. This process generates one timeline of the volume of related tweets for each emotion. The further investigation of these timetables uncovers that each of the four feelings changes after some time in a somewhat unsurprising way. To give a case, we found an intermittent top of happiness around Christmas and an occasional crest of apprehension around Halloween. More surprisingly, we found that negative moods started to dominate the Twitter content after India lost the semi final against Australia Also there was a significant increase in happiness in the weeks before when India won the match against Pakistan.

III. BIG DATA TECHNIQUES AND TECHNOLOGIES

A wide assortment of systems and advancements has been created and adjusted to total, control, break down, and imagine huge information. These methods and advancements draw from a few fields including insights, software engineering, connected arithmetic, and financial matters.

A. Techniques for analyzing big data

There are many techniques that draw on disciplines such as statistics and computer science (particularly machine learning) that can be used to analyze datasets.

A/B testing : A system in which a control gathering is contrasted and an assortment of test gatherings to figure out what medications (i.e., changes) will enhance a given target variable

Crowd sourcing: A technique for collecting data submitted by a large group of people or community (i.e., the “crowd”) through an open call, usually through networked media such as the Web. This is a type of mass collaboration and an instance of using Web 2.0.

Machine learning: A subspecialty of computer science (within a field historically called “artificial intelligence”) concerned with the design and development of algorithms that allow computers to evolve behaviors based on empirical data. A major focus of machine learning research is to automatically learn to recognize complex patterns and make intelligent decisions based on data. Natural language processing is an example of machine learning.

Data fusion and data integration: A course of action of systems that arrange and analyze data from different sources with a particular finished objective to make bits of learning in ways that are more capable and conceivably more correct than if they were delivered by researching a singular wellspring of data.

Data mining: A set of techniques to extract patterns from large datasets by combining methods from statistics and machine learning with database management. These techniques include association rule learning, cluster analysis, classification, and arrested development.

Neural networks: Computational models, roused by the structure and Synonyms (Ordered by Estimated Frequency) of noun working of natural neural organization (i.e., the cells and connation inside a cerebrum), that discover designs in data . Neural organization is appropriate for finding nonlinear examples. They can be utilized for

example acknowledgment and improvement. Some neural system applications include directed encyclopedias and others include unsupervised learning.

Network analysis: A set of techniques used to characterize relationships among discrete nodes in a graph or a network. In social network analysis, connections between individuals in a community or organization are analyzed.

Predictive modeling: A set of techniques in which a mathematical model is created or chosen to best predict the probability of an outcome.

Simulation: Demonstrating the conduct of complex frameworks, frequently utilized for anticipating, foreseeing and situation arranging. Monte Carlo recreations, for instance, are classes of calculations that depend on rehashed arbitrary inspecting, i.e., running a huge number of reenactments, each in light of various presumptions. The outcome is a histogram that gives a likelihood dissemination of results. One application is surveying the probability of meeting money related targets given instabilities about the accomplishment of different activities.

Time series analysis: Set of techniques from both statistics and signal processing for analyzing sequences of data points, representing values at successive times, to extract meaningful characteristics from the data.

Visualization: Procedures utilized for making pictures, outlines, or movements to convey, comprehend, and enhance the aftereffects of huge information examinations.

B. Big data technologies

Big Table: Proprietary distributed database system built on the Google File System. Inspiration for HBase.

Business intelligence (BI): A sort of utilization programming intended to report, dissect, and exhibit information. BI instruments are frequently used to peruse information that have been already put away in an information distribution center or information shop. BI devices can likewise be utilized to make standard reports that are created on an occasional premise, or to show data on continuous administration dashboards, i.e., incorporated showcases of measurements that gauge the execution of a framework.

Big Data: An open source (free) database management system designed to handle huge amounts of data on a distributed system. This system was originally developed at Facebook and

is now managed as a project of the Apache Software foundation.

Cloud computing: A computing paradigm in which highly scalable computing resources often configured as a distributed system. They are provided as a service through a network.

Dynamo: Proprietary distributed data storage system developed by Amazon.

Extract, transform, and load (ETL): Programming instruments used to concentrate information from outside sources, change them to fit operational needs, and load them into a database or information distribution center.

Hadoop: An open source (free) software framework for processing huge datasets on certain kinds of problems on a distributed system. Its development was inspired by Google's Map Reduce and Google File System. It was originally developed at Yahoo! and is now managed as a project of the Apache Software Foundation.

HBase: An open source (free), disseminated, non-social database demonstrated on Google's Big Table. It was initially created by Power set and is presently overseen as a task of the Apache Software establishment as a major aspect of the Hadoop.

MapReduce: A software framework introduced by Google for processing huge datasets on certain kinds of problems on a distributed system.³² Also implemented in Hadoop.

Mashup: An application that uses and combines data presentation or functionality from two or more sources to create new services. These applications are often made available on the Web, and frequently use data accessed through open application programming interfaces or from open data sources.

R programming: An open source (free) programming dialect and programming environment for factual registering and illustrations. The R dialect has turned into an accepted standard among analysts for creating measurable programming and is generally utilized for factual programming improvement and information examination. R is a piece of the GNU Project, cooperation that backings open source ventures.

Visualization: Presenting information in such a way that people can consume it effectively is a key challenge that needs to be met if analyzing data is to lead to concrete action.

Tag cloud: This graphic is a visualization of the text of this report in the form of a tag cloud, i.e., a weighted visual list, in which words that appear most frequently are larger and words that appear less frequently smaller.

Clustergram: A clustergram is a visualization technique used for cluster analysis displaying how individual members of a dataset are assigned to clusters as the number of clusters increases.

History flow: History flow is a visualization technique that charts the evolution of a document as it is edited by multiple contributing authors. Time appears on the horizontal axis, while contributions to the text are on the vertical axis.

IV. WHY WE USE BIG DATA RATHER THAN TRADITIONAL DATABASE TECHNIQUES?

Big data as is defined is build for volume, variety and velocity. So, rightfully, any data base involving Big data will be build to handle this. This is different from traditional warehouse which are relational, have fixed data type formats and have limitation on the data types they can handle and have limitation in terms of the volume they can handle. Today, the problem is very different. Data sources are unpredictable, multi-structured (emanating from organized systems) and massive. Many are external to the enterprise. The techniques for mining data from these sources, and even the platforms most appropriate for doing so, are now somewhat in question. With the entry of Hadoop to the market – an approach that is separate from the relational data warehouse – the issue facing decision makers today is where and when to deploy these technologies for performing useful analytics.

V. SOCIAL NETWORK DATA ANALYSIS USING BIG DATA

As per Wikipedia, informal community examination is "the investigation of interpersonal organizations. Interpersonal organization examination sees social connections regarding system hypothesis, comprising of hubs (speaking to individual performing artists inside the system) and ties. These systems are frequently delineated in an informal community graph, where hubs are spoken to as focuses and ties are spoken to as lines."

A. Introduction of Open Source Big Data Synthesize - NodeXL

Marc Smith is a sociologist who specializes in the social organization of online communities and computer-mediated interaction. Marc has pioneered a revolutionary graphics program, NodeXL, which synthesizes and clusters social network data. Instead of a complicated listening platform,

NodeXL is able to synthesize, for example, Twitter feeds, and produce a relevant graphic and report. NodeXL creates maps that make sense of social media, and that is just the beginning.

B. Social Network Analysis

Social network analysis views social relationships in terms of network theory, consisting of nodes (representing individual actors within the network) and ties (which represent relationships between the individuals, such as Facebook friendships, email correspondence, hyperlinks, or Twitter responses) These networks are often depicted in a social network diagram, where nodes are represented as points and ties are represented as lines. Visualization of social networks is now coming online to make sense of network data and convey the results of analyses. NodeXL is one of the leading, open-source analytic software modules for network visualization.

This is valid for Social Network Analysis. With a specific end goal to channel the brand discussion, the primary wellsprings of information are Twitter, Facebook, YouTube, and others. Each contains one or more informal organizations.

Online networking system maps are an extraordinary approach to get your very own diagram social networking themes and the related ones that matter to you. Maps offer a fast approach to evaluate the kind of social structure that has developed around a subject. These maps make it simple to think about the courses bunches in your systems use URLs, hashtags, words, @usernames, and phrases.

C. Deciphering a Network Graph

A network graph can be summarized by a number of metrics. NodeXL reports these measures for each network. These values can be a useful way to compare networks. Marc Smith, the creator of NodeXL, offers this summary of key metrics.

Vertices: the count of unique entities in the network (this is the count of users or accounts)

Unique Edges: the count of connected pairs of vertices. VII.

Edges With Duplicates: the count of connections that are repeated

Total Edges: the count of all connections

Self-Loops: the count of connections that start and end in the same person (these represent tweets that have no other person's name in them)

Reciprocated Vertex Pair Ratio: the proportion of vertices or users who have a connection returned to them.

Reciprocated Edge Ratio: The fraction of edges that are matched by an edge in the opposite direction.

Connected Components: the number of distinct, isolated, sets of connected vertices.

Single-Vertex Connected Components: the count of isolate vertices, users who do not mention anyone

Maximum Vertices in a Connected Component: the count of vertices within the largest set of connected vertices.

Maximum Edges in a Connected Component: the count of connections within the largest set of connected vertices.

There is quite a bit of information to digest. However, the marketing researcher, the corporate client, and the C-Suite executive need only remember a few of these measures. These are summarized below.

D. Interpretation of measures

Centrality of measure

Degree => How many people can this person reach directly?

Betweenness => How likely is this person to be the most direct route between two people in the network?

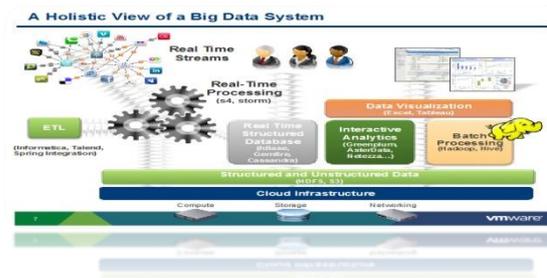
Closeness => How fast can this person reach everyone in the network?

Eigenvector => how well is this person connected to other well-connected people?

Social network maps are made up of four major factors: Sub-groups. Bridges. Island, and clusters.

1. Hubs
2. Bridges
3. Island
4. Clusters/Crowd.

VI. BIG DATA SYSTEM



The diagram above illustrates this framework and shows that some components, or potentially the entire big data system, can run on a cloud infrastructure, which can make the system elastic, highly available, and multi-tenant. With resource sharing, we can ultimately bring the benefits of cloud computing to big data and keep budgets in check.

Huge information frequently brings four more current and altogether different contemplations in an endeavor design: Data sources have an alternate scale – while the most self-evident, numerous organizations work in the multi-terabyte and even pet byte coliseum.

Speed is critical – nightly ETL (extract-transform-load) batches are insufficient and real-time streaming from solutions like s4 and Storm are required.

Storage models are changing – solutions like HDFS (Hadoop Distributed File System) and unstructured data stores like Amazon S3 provide new options.

Multiple analytics paradigms and compute methods must be supported:

Real-time database and analytics: These are commonly in-memory, scale-out motors that give low-dormancy, cross-server farm access to information, and empower appropriated handling and occasion era abilities.

Interactive analytics: Includes distributed MPP (massively parallel processing) data warehouses with embedded analytics, which enable business users to do interactive querying and visualization of big data.

Batch processing: Hadoop as a circulated handling motor that can examine a lot of information and apply calculations that range from the straightforward (e.g. accumulation) to the complex (e.g. machine learning).

Using large databases to improve the delivery of government services and to monitor for threats to national security.

Large databases also open up all sorts of new business opportunities. “Now-casting” is helping companies understand the real-time dynamics of certain areas of life—from the diffusion of diseases to consumer purchases to night-life activity—which will have many long-term reverberations on markets. New sorts of information middle people are likewise liable to emerge to individuals understand a something else stupefying surge of data. Surely, information middle people and translators could speak to a blossoming portion of the data innovation segment in the years ahead. However, Big Data additionally introduces numerous imposing difficulties to government and natives accurately in light of the fact that information advancements are turning out to be so pervasive, nosy and hard to get it. In what manner might society secure itself against the individuals who might abuse or mishandle huge databases?

What new regulatory systems, private-law innovations or social practices will be capable of controlling anti-social behaviors—and how should we even define what is socially and legally acceptable when the practices enabled by Big Data are so novel and often arcane? These are a portion of the vital open inquiries postured by the ascent of Big Data. This report proposes a portion of the more notable issues that ought to be tended to. In the coming years, government, business, buyers

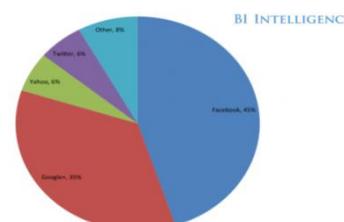
and resident gatherings should dedicate much more noteworthy consideration regarding the monetary, social and individual ramifications of expansive databases. Somehow, our general public should take some inventive, innovative jumps to guarantee that database advancements and strategies are utilized successfully and dependably.

VIII. SOCIAL NETWORKS WITH MARKETING

For marketers this is an opportunity to better understand, target, approach clients, prospects, suspects, to sell them more, for better lead communities, to innovate, differentiate themselves from the competition and develop a competitive advantage. Tweets and products show the popularity of the brand thereby attracting consumers for the same and enhancing sales promotion to manufactures. By analyzing the social connections, the enterprise will probably discover that the existing organization structure is partially bypassed so that the individuals can work with others that are not part of their group or department. The enterprise may be able use the connection information to reorganize their internal organization structure to reflect the social network connections and thereby improve enterprise productivity. If you share your blog post in email or on Twitter, Facebook and other social channels to Check the Availability of Your Business Name on Social Networks. We can track the content and analyze the blog in variety of channels (Email, Facebook, Google Plus, LinkedIn, and Twitter) for viewing the information and monitor user visitor’s activity. It also provides the option to find the competitors ranks.

IX. SOCIAL NETWORKS ANALYTICS

The enterprise, with a much smaller number of internal social network users than the public social networks like Facebook, is still generating big data, and enterprises should therefore consider what to collect and how to use the collected data to improve their internal operations. Analyzing internal social networks can provide business intelligence. Google Analytics is a great tool to monitor the performance of your blog, but you may not look at it regularly enough. One way to get around this is to set up a weekly report for delivery to you via email.



X. CONCLUSION

Big data provides a efficient and effective way to deal with the data generated from the social networks on daily basis which produces useful information for decision making in many areas.

REFERENCES

- [1] Promise and Peril of Big Data , David Bollier, Rapporteur
- [2] Visualizing Big Data Social Network Analysis , Michael Lieberman.
- [3] Hansen, Derek, Ben Shneiderman, and Marc A Smith. (2001). *Analyzing Social Media Networks WithNodeXL: Insights from a connected world.* Morgan Kaufmann, Marc A Smith, Kollock, Peter (1998). *Communities in Cyberspace*, Routledge .
- [4] Barnes, John. (1954). "Class and Committees in a Norwegian Island Parish." *Human Relations*, (7): 39-58.
- [5] <http://blogs.vmware.com/vfabric/2012/08/4-key-architecture-considerations-for-big-data-analytics>

ARTICLES

Role of Computer in Global Warming

Guided by: Rupali Jadhav

Kumbhar Harish
TIMSCDR

Masdekar Mohanish
TIMSCDR

Mishra Manishkumar
TIMSCDR

Abstract - In the past few years, the world has experienced devastating natural disasters on a level that hasn't been seen for decades. There is much speculation that these especially strong phenomena are due to global climate change, brought on by Global Warming. The term 'Global Warming' refers to the rising temperature of the earth due to an increased amount of greenhouse gases. The green house gases increased due to excess and non-efficient use of computers. The computer spews Tons of CO₂. Computer doesn't necessarily run on carbon fuels however the electricity it uses to run daily does and the amount used has been piling up over the past few decades now as more people buy home PC units to enjoy the world of computing. There is solution to reduce growing problem & many ways to help lessen the burden on our mother earth. We focused on the above problem & try to give out some simple solutions that can be implemented with very less efforts.

Keywords: *Global warming, green house gases, power saver*

I. INTRODUCTION

'Global warming' is a phrase that refers to the effect on the climate of human activities, in particular the burning of fossil fuels (coal, oil and gas) and large-scale deforestation, which cause emissions to the atmosphere of large amounts of 'greenhouse gases', of which the most important is carbon dioxide. Such gasses retain infrared radiation discharged by the World's surface and go about as covers over the surface keeping it hotter than it would somehow be. Connected with this warming are changes of atmosphere. The fundamental exploration of the 'nursery impact' that prompts the warming is surely knew. An unnatural weather change is the most imperative ecological issue the world countenances.

Adjustment to the inescapable effects and Improvement to diminish their extent are both important. Worldwide move is being made by the world's logical and political groups. In view of the requirement for pressing activity, the best test is to move quickly to quite expanded vitality productivity and to non-fossil-fuel vitality sources utilizing cutting edge advances.

II. GLOBAL WARMING

A. What is Global Warming?

Global Warming is when the earth heats up. It happens when the green house gases trap heat & light from sun in the earth's atmosphere, which increases the temperature. This is called Green house effect. An Earth-wide temperature boost results from an expanded nursery impact in the climate. As global temperatures increase, the climate changes.[1]

B. Causes/Reasons for Global Warming

- Excess use of Computer
- Automotive Applications
 - Use of Automobiles has the major share.
 - Tail end emissions increase the green house gases.
 - Excessive use of fuels demanding more resources
- Thermal Power Generation
 - Usage of coal reducing the fossil resources.
 - Burning of coal increasing the pollution.
- Chimney Top emissions in Factories [3]

C. Green House Gases :

- CO₂ Carbon Dioxide
- CFC's Chloro fluoro carbons
- CO Carbon Monoxide
- SO₂ Sulfur dioxide

D. Green House Effect:

- Certain gasses in the climate are asserted to assimilate heat vitality and decrease the sum that getaways into space
- This 'trapping' of heat from our only source of warmth – the Sun – by the atmosphere is known as the 'Greenhouse Effect'
- This gives us a global average temperature of +15°C rather than -18°C

- if we had no Greenhouse Effect
- An hypothesis unsupported by experiment – could be cloud cover alone...the jury is out .
- Water vapors is the most important greenhouse gas - 60 - 98% of the greenhouse effect is due to water vapor in the atmosphere (jury still out)
- CO2 and other minor gasses account for the remainder[1][7]

III. PROBLEMS

A. Fuel Consuming Computer

The computer doesn't necessarily run on carbon fuels however the electricity it uses to run daily does and the amount used has been piling up over the past few decades now as more people buy home PC units to enjoy the world of computing. [6]

Computer spews Tons of CO₂:

The Met Office has caused a storm of controversy after it was revealed their 30million supercomputer designed to predict climate change is one of Britain's worst polluters. ...

It is fit for 1,000 billion figuring consistently to encourage information to 400 researchers and utilizations 1.2 megawatts of vitality to run - enough to power more than 1,000 homes.

The machine was hailed as the 'future of weather prediction' with the ability to produce more accurate forecasts and produce climate change modeling. However the Met Office's HQ has now been named as one of the worst buildings in Britain for pollution - responsible for more than 12,000 tones of carbon dioxide a year.

It says 75 for every penny of its carbon impression is created by the super PC meaning the machine is authoritatively one of the nation's slightest green machines.

The carbon-spewing energy hog does not even predict the weather accurately.

B. What will happen if global warming continues?

There are already some changes happening because of global warming. Sea level is rising and some animals are already moving to new homes. It's already too late to stop global warming completely. If the warming gets worse, as scientists expect, there may be some kinds of plants and animals that become extinct (disappear completely) because they can't move to new homes. There might be more tempests and surges. Ocean level may rise so

much that individuals need to move far from the coasts. A few ranges may turn out to be excessively dry for cultivating.

IV. SOLUTIONS

So you are most likely pondering what arrangement there might be to decrease your commitment to this developing issue. There are numerous approaches to diminish the weight on our mom earth:

Turn Off Computer When Not in Use Adjust Your Hibernate Mode to Idle at a Lesser Time Install the Free Co2 Saver from Snap Onto Your Computer Explore Alternative Energy for Your Home (Solar, Wind, etc.)[6]

A. Free Energy Saving Program Installs

The Co2 Saver from Snap is a cool install application that is free of spyware that allows you to monitor your computer's usage and the impact it is making on the environment in relation to carbon dioxide. It is a real nifty little tool as it lets you see what you have saved as well as how much you have consumed. It automatically can adjust your computer to the most environmentally friendly settings in your contribution to battle global warming.

Another cool install does about the same thing exact thing except it also shows you what natural resources you have saved and what you have consumed in relation to your computer's energy consumption. The program is offered by a company named ni Blue and is called Local Cooling. Both programs are entirely free, so it might be fun to see just how much of a dent you can put in the global warming trend.

While an unnatural weather change and contamination may appear like a stylish theme regularly connected with governmental issues, nonconformist parades, or MTV the point is nothing to disregard or simply indicate sympathy toward one day out of a year. The truth of life is that we as a whole have recently this one planet for some eras to come. This planet is the thing that gives us life, sustenance, air, and water. We can't hope to appreciate the advantage of living in the event that we don't secure our human environment.

TopSavings.Net - High Speed Internet Service can be compared by region online at TopSavings.Net over the Internet. Compare and save with no sales calls just by entering email and phone number details. NOTE: This article may not be reproduced without this information intact with working links.

B. Shut off your computer

A screen saver is not an energy saver. According to the U.S. Department of Energy, 75% of all the electricity consumed in the home is standby power used to keep electronics running when those TVs, DVRs, computers, monitors and stereos are "off." The average desktop computer, not including the monitor, consumes from 60 to 250 watts a day. Compared with a machine left on 24/7, a computer that is in use four hours a day and turned off the rest of the time would save you about \$70 a year. The carbon impact would be even greater. Shutting it off would reduce the machine's CO2 emissions 83%, to just 63 kg a year.

C. Power saver Search engine:

www.powersaversearch.com :By using this search engine, we can save energy upto 10% of India's annual power shortage.

This special search engine (like google,yahoo) saves power consumed by your monitor while you are searching the net. This way, 750 Megawatt energy can be saved annually, which is more than 10% of India's total power shortage.

The site uses dark color themes which consume less power, and hence, saves tremendous amount of energy. Also, the site reminds us of doing our bit each time it's loaded. So, set <http://www.Powersaversearch.com> as your browser homepage, and start saving power today! Next time whenever you think of searching something, please use <http://www.powersaversearch.com>. You get the same results as Google, so there is no difference in output also. It's like a black Google that saves the power for us. [5]

Display Colors: White and bright colors (especially in backgrounds) can use up to 20% more power than black or dark colors. Look to the right to see the power usage (in Watts) of a sample monitor with different screen backgrounds. Unfortunately, e-mail and word processors tend to use white backgrounds, so your workstation uses considerable power while you are in these programs, which you are during much of the day. Because black-on-white is the most familiar (it's just like the newspaper), selecting alternate combinations may not be appealing. However, you can change your desktop background to something efficient. by selecting Start, Settings, Control Panel, Display, and Appearance tab. The "Item" field Should say "Desktop" . Under colour select one of the color at the right that has rating below 65W& click [5]

White - 74W	Fuchsia - 69W	Yellow - 69W
Aqua - 69W	Silver - 67W	Blue - 67W
Red - 67W	Lime - 67W	Gray - 62W
Olive - 61W	Purple - 61W	Teal - 61W
Green - 60W	Maroon - 59W	Black - 59W (Least Power Consumption)

Table 1. Color and power usage

Blackle inquiries are fueled by Google Custom Search. Blackle spares vitality in light of the fact that the screen is dominantly dark. "Picture showed is basically an element of the client's shading settings and desktop illustrations, and additionally the shading and size of open application windows; a given screen requires more energy to show a white (or light) screen than a dark or dull screen In January 2007 a blog entry titled Black Google Would Save 750Megawatt-hours a Year proposed the theory that a black version of the Google search engine would save a fair bit of energy due to the popularity of the search engine. [4]

V. CONCLUSION

"Let's be accountable in dropping the rate of global warming by following the preventive steps and make earth a better place to LIVE in."

REFERENCES

- [1] en.wikipedia.org/wiki/Global_warming
- [2] environment.nationalgeographic.com/...
- [3] <http://timeforchange.org/cause-and-effect-for-global-warming>
- [4] <http://www.blackle.com/about/>
- [5] <http://www.ynotshare.com/save-your-pc-power-usage-by-25-powersaversearch-com/22836/>
- [6] <http://ezinearticles.com/?GlobalWarming-and-Yourcomputer&id=1070832>
- [7] www.globalwarming.org.in

Introduction to Machine Learning

Guided by: Ms. Priya Sinha

Trishul Mody
TIMSCDR

Rohit Maurya
TIMSCDR

Matsyendra Mishra
TIMSCDR

Abstract- in the advancing field of computing, the features of algorithmic learning, intuitiveness and consciousness can be achieved via Machine Learning. This paper gives a brief introduction about Machine Learning and its different aspects like – algorithms, problems, applications, uses and future development. This paper analyzes the different real life applications and uses of Machine Learning and also deals with its shortcomings and lists the unsolved problems that could be solved in the future with the help of better and refined Machine Learning approaches. The conclusion denotes that Machine learning and other similar approaches will be used extensively in the future of computing.

I. INTRODUCTION

Machine learning is a kind of artificial intelligence (AI), the ability to provide computers and learning are not explicitly programmed. Machine learning focuses on study when exposed to new data growth and change in the development of computer programs.

Machine learning and data mining process is similar. Both systems perform a search in any data to look for patterns. However, instead of human understanding of the extracted data - as in the case under the application of data mining - machine learning aims to use the data to detect patterns in the data, and adjust the plan of action accordingly. Machine learning algorithms are often categorized as being supervised or unsupervised. Supervised algorithms aim to apply what has been understood in the past to new data. Unsupervised algorithms aim to draw inferences from the datasets.

Machine learning is the process to enable all of the computers to act without being explicitly programmed. In the last decade, machine learning has introduced us to self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many researchers also think it is the best way to make progress towards human-level AI. [1] [2]

II. TYPES OF MACHINE LEARNING PROBLEMS

There are common classes of problem in Machine Learning. The problem classes below are archetypes for most of the problems we refer to when we are *doing* Machine Learning.

- **Classification:** Data is labeled meaning it is assigned a class, for example spam/non-spam or fraud/non-fraud. The decision being modeled is to assign labels to new unlabeled pieces of data. This can be thought of as a discrimination problem, modeling the differences or similarities between groups.
- **Regression:** Data is labeled with a real value (think floating point) rather than a label. Examples that are easy to understand are time series data like the price of a stock over time, the decision being modeled is what value to predict for new unpredicted data.
- **Clustering:** Data is not labeled, but can be divided into groups based on similarity and other measures of natural structure in the data. An example from the above list would be organizing pictures by faces without names, where the human user has to assign names to groups, like iPhoto on the Mac.
- **Rule Extraction:** Data is used as the basis for the extraction of propositional rules (antecedent/consequent aka *if-then*). Such rules may, but is typically not directed, meaning that the methods discover statistically supportable relationships between attributes in the data, not necessarily involving something that is being predicted. [3] [9]

III. USES OF MACHINE LEARNING

There are numerous applications of machine learning. It's actually hard to realize how much machine learning has achieved in real world applications. Machine learning is generally just a

way of fine tuning a system with tunable parameters. It is a way of making a system better with examples, usually in a supervised or unsupervised manner. Machine learning is normally applied in the offline training phase. Thus machine learning is used to improve the following applications.

- **Face detection:** The face detection Feature in mobile cameras is an example of what machine learning can do. Cameras can automatically snap a photo when someone smiles more accurately now than ever before because of advances in machine learning algorithms.
- **Face recognition:** This is where a computer program can identify an individual from a photo. You can find this feature on Face book for automatically tagging people in photos where they appear. Advances in machine learning mean more accurate auto-face tagging software.
- **Image classification:** A good example is the application of deep learning to improve image classification or image categorization in apps such as Google photos. Google photos would not be possible without advances in deep learning.
- **Speech recognition:** Another good example is Google now. Improvements in speech recognition systems has been made possible by, you guessed right, machine learning specifically deep learning.
- **Google:** Google defines itself as a machine learning company now. It is also a leader in this area because machine learning is a very important component to its core advertising and search businesses. It applies machine learning to improve search results and search suggestions.
- **Anti-virus:** Machine learning is used in Anti-virus software's to improve detection of malicious software on computer devices.
- **Anti-spam:** machine learning is also used to train better anti-spam software systems.
- **Genetics:** Classical data mining or clustering algorithms in machine learning such as agglomerative clustering algorithms are used in genetics to help find genes associated with a particular disease.
- **Signal denoising:** Machine learning algorithms such as the K-SVD which is just a generalization of k-means clustering are used

to find a dictionary of vectors that can be sparsely linearly combined to approximate any given input signal. Thus such a technique is used in video compression and denoising.

- **Weather forecast:** Machine learning is applied in weather forecasting software to improve the quality of the forecast. ^[4]

IV. REAL LIFE EXAMPLES OF MACHINE LEARNING

Machine Learning problems are abounding. They make up core or difficult parts of the software you use on the web or on your desktop every day. Think of the “do you want to follow” suggestions on twitter and the speech understanding in Apple’s Siri.

- **Spam Detection:** Given email in an inbox, identify those email messages that are spam and those that are not. Having a model of this problem would allow a program to leave non-spam emails in the inbox and move spam emails to a spam folder. We should all be familiar with this example.
- **Credit Card Fraud Detection:** Given credit card transactions for a customer in a month, identify those transactions that were made by the customer and those that were not. A program with a model of this decision could refund those transactions that were fraudulent.
- **Digit Recognition:** Given a zip codes hand written on envelopes, identify the digit for each hand written character. A model of this problem would allow a computer program to read and understand handwritten zip codes and sort envelopes by geographic region.
- **Speech Understanding:** Given an utterance from a user, identify the specific request made by the user. A model of this problem would allow a program to understand and make an attempt to fulfill that request. The iPhone with Siri has this capability.
- **Face Detection:** Given a digital photo album of many hundreds of digital photographs, identify those photos that include a given person. A model of this decision process would allow a program to organize photos by person. Some cameras and software like iPhoto has this capability.
- **Product Recommendation:** Given a purchase history for a customer and a large inventory of

products, identify those products in which that customer will be interested and likely to purchase. A model of this decision process would allow a program to make recommendations to a customer and motivate product purchases. Amazon has this capability. Also think of Face book, Google Plus and Face book that recommend users to connect with you after you sign-up.

- **Medical Diagnosis:** Given the symptoms exhibited in a patient and a database of anonymized patient records, predict whether the patient is likely to have an illness. A model of this decision problem could be used by a program to provide decision support to medical professionals.
- **Stock Trading:** Given the current and past price movements for a stock, determine whether the stock should be bought, held or sold. A model of this decision problem could provide decision support to financial analysts.
- **Customer Segmentation:** Given the pattern of behavior by a user during a trial period and the past behaviors' of all users, identify those users that will convert to the paid version of the product and those that will not. A model of this decision problem would allow a program to trigger customer interventions to persuade the customer to covert early or better engage in the trial.
- **Shape Detection:** Given a user hand drawing a shape on a touch screen and a database of known shapes, determine which shape the user was trying to draw. A model of this decision would allow a program to show the platonic version of that shape the user drew to make crisp diagrams. [5]

V. TOP 10 PROBLEMS IN MACHINE LEARNING

- **Churn Prediction:** Churn prediction is one of the most popular use cases for people who want to leverage machine learning. It has a large business value and benefit attached to itself especially in industries like the telecom and banking. Several challenges such as the skewed nature of the data set available and the ability to decide which models to use are going to be under a lot of debate.
- **Sentiment Analysis:** A lot of decisions these days are being taken on the opinion of others. We buy a product more because it has received a positive opinion and we visit a hotel most

likely because it got the best rating online. Sentiment analysis has its own challenges such as how granular can the sentiment be determined, how subjective is the sentiment and so on, and hence sentiment analysis will be a good place to attack machine learning.

- **Truth and Veracity:** There is a lot said online these days and it is hard to determine what is true and what is fake. We have bots smart enough to publish content like human beings and there are social aspects attached to the ratings of various entities online. Machine learning will be leveraged as a big challenge to determine the veracity/truth of information online.
- **Recommendations:** There is such a myriad of choices available online that it is becoming even more difficult to choose a book, restaurant or even a simple product. The ability to do smart recommendations based on the user's context and not just based on the preferences of the crowd is going to be a great challenge because it is a great deal to understand the user's context.
- **Online Advertisement:** There is a lot of work and many start ups around the space of intelligent online advertisements, but to be able to push the right advertisement at the right time in the right way to the user needs a lot of understanding of the when to target a particular customer. Machine learning exhibits a great challenge in this space in my opinion for determining the user's behavior online to push the correct advertisement instantly when the user really needs it.
- **Aggregation:** Plenty of news is being generated around us from various different places about a variety of topics. Yet we all have a constant thirst to consume all the news relevant to us as much as possible. How are we going to aggregate news according to the user's preference? Does his taste vary with time? How do we learn this variation? All this is going to be a challenge for machine learning and it involves a great deal of making sense of news and articles.
- **Scalability:** Data is constantly expanding in variety, velocity and volumes. Can the traditional machine learning algorithms that were developed a decade back be applied to big data? They will all undergo some kind of refurbishment to be able to operate on data at large scale. Can SVMs train faster? Can it be made parallel? This is going to be a good problem to focus with the rise of big data.

VI. TOP 10 MACHINE LEARNING ALGORITHMS

- Content Discovery/Search: There are millions of people around the world on various social networks and within enterprise. How can you discover people who share similar interests as yours and what parameters are you going to consider measuring this similarity? How do we measure similarity and can we quantify this? This is a nice problem for machine learning where we will face the challenge of trying to find the needle in a haystack.
- Intelligent Learning: For example, it is still difficult to identify a behavior in a video sequence and there has been a lot of research around this space. In my opinion, with the state of art learning algorithms, one of the top problems is to be able machines to be able to see, hear and recognize like the human brain does. This means a good problem would be to leverage machine learning algorithms to use different modes of learning to achieve a particular task, be it recognition or anything similar.

Machine Learning for Medicine: This is probably the most interesting machine learning problem for 2013 and coming future. There are so many diseases that need our attention and a lot of human time spent in researching for their cure by analyzing a lot of symptoms. Yet, two patients with similar health problems receive different kinds of treatment and eventually different extents of cure. Can we use machine learning to understand how a patient is feeling at a particular moment and appropriate recommend the right treatment for him? This will change how we are going to live and will help doctors discover a lot of new medical methodologies.

In addition there are other problems such as:

- Statistical Predicate Invention
- Generalizing Across Domains
- Learning Many Levels of Structure
- Deep Combination of Learning and Inference
- Learning to Map between Representations
- Structured Prediction with Intractable Inference
- Reinforcement Learning with Structured Time
- Expanding SRL to Statistical Relational AI
- Learning to Debug Programs [3]

- Kernel Density Estimation and Non-parametric Bayes Classifier
- K-Means
- Kernel Principal Components Analysis
- Linear Regression
- Neighbors' (Nearest, Farthest, Range, k, Classification)
- Non-Negative Matrix Factorization
- Support Vector Machines
- Dimensionality Reduction
- Fast Singular Value Decomposition
- Decision Tree
- Bootstrapped SVM [6] [7]

VII. CONCLUSION

The Future of Human Assisted Machine Learning:

In today's data-driven analytical world, there's simply too much data for any human to crunch the numbers on. As a result, we're seeing machine learning playing a significant role in the development of ever-more sophisticated analytics models through Amazon, IBM's Watson, and Microsoft's Azure, and more.

Algorithms can help these systems to learn to draw better insights by improving their data collection and analysis over time. But what about things that still require a human touch?

In trying to analyze online conversation, a machine would have hard time picking up on sarcasm. This is something that still requires some sort of human intervention (for now, at least). Similarly, the voice-fed questions and instructions that are said to Siri, Cortana and Google Now will invariably lead to confusion on the part of the machines, with responses like, "I didn't understand that."

Face book announced that it is launching M, a virtual assistant that will add a distinctly human touch to voice commands through its Messenger app. Think of it like a combination of Siri and a concierge, as M will be a hybrid of artificial intelligence and human interaction.

While you can't completely scale human interaction, is the future inextricably tied to more automation? Maybe it's a hybrid of both.

Machine learning, in the near future, will be extensively used in the fields of Genetics, Robotics, Finance, Linguistics and Computer Systems. [8]

REFERENCES

- [1] Stanford University department of Computer Science. (www.cs229.stanford.edu)
- [2] Target tech publication. (<http://techtargget.com>)
- [3] “Structured Machine Learning: Ten Problems for the Next Ten Years” a paper by Pedro Domingos.
- [4] (<https://homes.cs.washington.edu/~pedrod/papers/ilp07.pdf>)
- [5] Quora. (<https://www.quora.com/topic/machine-learning>)
- [6] Oracle Advanced Analytics Data Mining Algorithms and Functions.
- [7] <http://www.oracle.com/technetwork/database/enterprise-edition/odm-techniques-algorithms-097163.html>)
- [8] Analytics 1305 and Ants 11. (<http://jtonedm.com>)
- [9] “The Master Algorithm” a book by Pedro Domingos.
“A Few Useful Things to Know about Machine Learning” a paper by Pedro Domingos.
<https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>

A Comparative Study: IoT Tools Raspberry Pi and Arduino use in Healthcare

Guided by: Ms.Sonam Pareek

Pankaj Singh
TIMSCDR

Ramratan Singh
TIMSCDR

AjeetKumar Yadav
TIMSCDR

Abstract: - The objective of this research paper is to make aware how internet of things (IoT) is transforming healthcare and the role of IT in healthcare more than others stand to be revolutionized from the sensor-driven, connected climate created by the IoT—not the least of which is health care. Let's take a moment to , to take advantages of IoT, it is necessary that medical enterprises and community should trust the IoT systems in terms of performance, patient security, privacy, reliability and return on investment, which are open challenges of current IoT systems. We also focus on how IoT tools like Arduino and Raspberry Pi beneficial to health care sector. A comparative study of IoT tools Arduino and Raspberry.

I. INTRODUCTION

Through the Internet of things, anything in the healthcare system can be identified tracked and monitored on demand anytime any where Internet of things is considered as remarkable revolution after the blooming of Internet with ICT based industry. Internet of things has three basic components, namely RFID systems, middleware systems and Internet systems Savant. RFID system is one of the major components of IOT and it enables data to be transmitted by a portable device, called “a tag”, which is read by an RFID reader and processed according to the needs of a particular application The Healthcare industry remains among the fastest to adopt the Internet of Things. The reason for this trend is that integrating IoT features into medical devices greatly improves the quality and effectiveness of service, bringing especially high value for the elderly, patients with chronic conditions, and those requiring constant supervision. . However, to take advantages of IoT, it is essential that medical enterprises and community should trust the IoT systems in terms of performance, security, privacy, reliability and return on investment, which are open challenges of current IoT systems. For heightening of healthcare system;

tracking, tracing and monitoring of patients and medical objects are more essential. But due to the inadequate healthcare situation, medical environment, medical technologies and the unique requirements of some healthcare applications, the obtainable tools cannot meet them accurately. Connected medical devices and applications are already part of the Internet of Medical Things. How much they will integrate into patient care and other aspects of healthcare remains to be seen. This guide section dives into how IoT can be used to collect share patients' biometric data and monitor them after they've been discharged from a hospital. Keep reading to see what other possibilities could be realized through healthcare IoT.

II. BENEFITS OF IOT TREND IN HEALTHCARE

- Collecting, sharing and analyzing health data received in real time
- Providing better treatment based on more specific data
- Reaching a wider range of healthcare providers
- Protecting patient privacy and the sensitive health data inside these devices.
- Additionally, if these devices interface with medical billing records, then patients risk losing both medical and financial information
- Health-monitoring products provide real-time feedback about nutrition, fitness, pulse, blood pressure, and other vital signs.
- Using the available data, and aided by decision-support systems that also have access to a large corpus of observation data for other individuals, the doctor can make
- a much better prognosis for your health and recommend treatment, early intervention, and life-style choices that are particularly effective in improving the quality of your health.

- Connect any wearable or portable device to the cloud, pull and analyze collected patient data in real time
- Monitor patients at home using live video and audio streaming
- Monitor vital health indicators collected by portable devices such as smart phones and tablets
 - Set intelligent emergency notifications sent to a physician or family.
 - Charts and diagram visualization based on data collected from health monitoring devices

Arduino and Raspberry Pi are the most popular boards among the students and health care professionals. Experienced and professionals know the utility and differences between the two. But beginners and students often get confused between them, like which board to use for their project or which board is easy to learn or why should they use Arduino over Pi and vice versa. So here I am covering mostly all the aspects which make them easy to take the decision over the choice of Arduino vs. Raspberry Pi.

III. INTRODUCTION OF RASPBERRY PI

Raspberry Pi is a fully functioned computer, a system-on-chip (SoC) device. In other words Raspberry Pi is a full-fledged credit card sized computer with 512 MB RAM and 700 MHz micro processor. It can run a full Linux based operating system and has hardware support for SPI, I2C and Serial [1].

A. Benefits of Raspberry Pi

- Entire Linux software stack is available, which runs on a Linux operating system specially designed for it, named Rasbian. Rasbian is the official OS for Raspberry Pi, where other third party OSes likes Firefox OS.
- It is like a mini computer, it has memory, processor, USB ports, audio output, and graphic driver for HDMI output and as it runs on Linux, most of the Linux software applications can be installed on it.
- It is very easy to connect to interne
- It has several models and revisions like Raspberry Pi, Raspberry Pi 2, and Raspberry Pi Model B+ etc.
- Can be programmed using variety of programming languages.

- Regarding devices like Raspberry Pi and Beagle Board, they are designed to function on a much higher level. With already integrated hardware that takes care of things like Ethernet, video and audio processing, large quantities of RAM and an almost unlimited amount of storage space, they are really mini-computers.
- Raspberry Pi is best used when you need a full-fledged computer: driving a more complicated robot, performing multiple tasks, doing intense calculations (as for Bit coin or encryption).

IV. INTRODUCTION OF ARDUINO

Arduino is a collection of three things. A hardware prototype platform, Arduino language and IDE & libraries. Arduino is best used for simple repetitive tasks: opening and closing a garage door, reading the outside temperature and reporting it to Twitter, driving a simple robot .The board is based on 8-bit AVR microcontroller. It has built-in hardware support for SPI, I2C and Serial. If you were to design a control panel for a microwave, you could use an AVR like Arduino. It could drive the LCD, interface with the buttons, and sense the conditions inside the appliance. You could have also used a device like Raspberry Pi or Beagle Board, but it would be overkill for the task.

A. Benefits of Arduino

- Arduino is best used for simple repetitive tasks: opening and closing a garage door, reading the outside temperature and reporting it to Twitter, driving a simple robot.
- Very easy to get started
- It doesn't need any OS and software applications to run, we just need to write few lines of code to make it use.
- There are many Arduino boards like Arduino UNO, Arduino PRO, Arduino MEGA, Arduino DUE etc.
- Very easy to extend it and has tons of user contributed shields and libraries. Shields are available to do pretty much anything.
- Can be used to for real-time applications.
- Not much programming knowledge needed to do basic stuff.

V. SIMILARITY BETWEEN RASPBERRY PI AND ARDUINO

Although they are quite different but there are some similarities in terms of their inception. They both are invented in European countries;

like Raspberry Pi is developed by Eben Upton in UK and Arduino is developed by Massimo Banzi in Italy. Both the inventors are teachers and they develop these hardware platforms as a design learning tool for their students. Raspberry pi was first introduced in year 2012 while Arduino in 2005. [1]

➤ If your project requires both software and hardware equally, then you can also use both Arduino and Raspberry Pi together in your project and get the best of both worlds.

1.Strength: Raspberry Pi runs on a OS so it must be properly shut down before turning OFF the power, otherwise OS & applications may get corrupt and Pi can be damaged. While Arduino is

Feature	Raspberry Pi	Arduino
Processor Speed	700 MHz	16 MHz
Programming Language	No limit	Arduno, C/C++
Real-time Hardware	No real-time	In real-time
Analog to Digital Convertor	No	Yes
Hardware Design	Closed source	Open source
Internet Connection	Very easy	Not easy, but doable

just a plug and play device which can be turned ON and OFF at any point of time, without any

risk of damage. It can start running the code again on resuming the power.

2.Price: Obviously Arduino is cheaper than Raspberry Pi, Arduino costs around \$10-20 depending on the version, while price of Raspberry is around \$35-40

3.Power consumption: Pi is a powerful hardware, it needs continuous 5v power supply and it is difficult to run it on Batteries, while Arduino needs less power can easily be powered using a battery pack.

VI. CONCLUSIONS

IT is used everywhere in each and every field. Health sector is also utilizing IT to improve the services and quality. To use IT on large scale some actions are needed such as government funding to provide IoT Kits those used to in healthcare must be increased, infrastructure should be improved, standard and protocols used should be improved, staff must be aware of new technology and government should conduct the hands on training program on IoT for their staff in partnership with Industries sector.

REFERENCES

- [1] <http://makezine.com/2015/12/04/admittedly-simplistic-guide-raspberry-pi-vs-arduino>
- [2] <http://circuitdigest.com>
- [3] <http://hardwarefun.com/tutorials/difference-between-raspberry-i-and-arduino>. adafruit.com

Virtualization In Cloud Computing

Guided by: Mr. Pankaj Mudholkar

Prakash singh
TIMSCDR

Anurag Singh
TIMSCDR

Ritesh Singh
TIMSCDR

Abstract: Cloud Computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources, (for example networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. But it has many security issues they are Lack of trust, loss of control and multitenency. In cloud computing at IaaS Infrastructure is provided so Virtual machines are running in this level and service has been provided. SO there is one security threat is VM Migration. So for preventing this VM Migration problem one protocol is used so through that protocol this problem becomes less. At this level different VMs of different users are running on the same physical server. So another problem at this level is risk of co-resident attack, in which user of one VM can get information from another VM. So for reducing this co-resident Attack different policies are there which perform the best when servers are properly configured but if not then there is a risk of co-resident attack so for reduce the risk of this new policy is there in which Previously selected server First is used that is for allocating the server for user. But still there is problem so that another policy in which one agent is there which checks the user and authenticate them and if there is any malicious VM then it discard it so that possibility of co-resident attack is becoming less and efficiency will be improve and secure services are being provided. So in this paper survey on virtual machine threats mainly VM Migration and Co-resident Attack is discussed.

Keywords— Artificial Neural Networks, ANN, Mobile Location Prediction, Cellular Network

I. INTRODUCTION

Cloud Computing is an on demand service model for IT provision based on Virtualization and distributed computing technologies [1]. In the current era, it is a wide field. It provides different services and platforms. It provides multi-tenancy, Massive scalability, elasticity, self-provisioning of resources. In cloud resources are shared so it is easy to use them whenever we want. It provides different service delivery models. They are Software As A Service (SaaS), Platform As A Service (PaaS), and Infrastructure As A Service (IaaS). SaaS rents software on a subscription basis. User can access the service through authorized

device. In short SaaS is a software distribution model in which Application resides on cloud service Provider (CSP) and are available for its clients via a web browser (e.g. Google docs) [2]. PaaS offers development environment to application developers. So it refers to the delivery of operating systems, associated tools, toolkits, building blocks over the internet. A Client deploys his application on the cloud service provider without installing any tool or platform on their local machines. In IaaS, CSP outsource the processing power, Storage, network and all other computing resources in the form of Virtual Machines. Hypervisor is provided at this level. Cloud computing provide different deployment models. They are Public cloud, Private cloud and Hybrid cloud. Public clouds are hosted, operated and managed by third party vendor. Their security and day to day management is also done by vendors. So it is available for all the consumers. While private clouds are restricted to any firm or organization. In which networks, infrastructures and data Centre's are owned by the organization. Hybrid cloud is a combination of both in which sensitive applications are provided by private cloud while non-sensitive applicators are provided by public cloud.

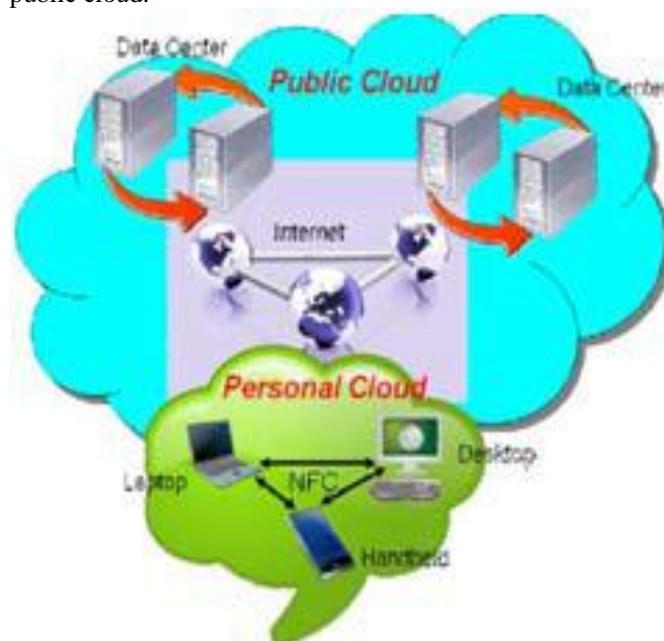


Fig 1: Cloud Usage Enterprise Environment

II. BACKGROUND

As Cloud infrastructure consist of large scale virtualized resources, traditional security mechanisms are not enough. There are some security issues that need to be considered. Main security issues are loss of control, lack of trust and multi-tenancy. Apart from that at different level like network, host and application layer Security should be considered. In SaaS customer has very less control over resources, therefore CSP is responsible for the required security mechanism. Whereas the PaaS offers greater custom control as compare to SaaS so CSP and customer are responsible for security mechanism. While IaaS offers greater custom control over security as compare to SaaS and PaaS. The PaaS and SaaS models are dependent on IaaS so any breach in IaaS model will affect the security of PaaS and SaaS.

Virtualization is a key at IaaS Model. So at IaaS Virtualization is provided. Different instances running on the same physical machine are isolated from each other is a major task of virtualization. Therefore this dynamic nature makes it difficult to achieve and maintain consistency security. The hypervisor manages and allocates the physical resources among the VMs. [5]



Fig 2: Virtualization

Diversity, VM Denial of Service and VM Mobility. [4]

- VM hooping (co-resident Attack): an attacker on one VM gains access to another VMs. The attacker can monitor any VM's resource usage, delete stored data and modify its configuration.
- VM Mobility/Migration: content can be moved or copied from one host to another.
- VM Diversity: Security management is done.
- VM Denial of Service: when multiple VMs share resources at a single time then VM denial of service is also one issue.

In virtualization Full Virtualization and Para Virtualization are two kinds of virtualization

provided in cloud computing. For that Virtual Machine Monitor (VMM) is used which abstracts the physical resources used by the multiple virtual machine. [6] VMM provides a virtual processor. VMs are running at a one time so VM Lunch, VM authentication, VM Migration, VM license all things should need to be consider. In cloud computing IaaS user can not verify the provider promised cloud platform integrity so it is a security risk. To prevent this issue one VM Lunch is introduced which allows the cloud user to securely bind the VM to a trusted computer. [7] Main issue is VM migration. VM live migration is done still there are some security issues that need to be consider.[11,12]

III. VM MIGRATION

VM Migration is a process in which running VM is migrated from one platform to another. Virtualization can provide significant benefits by enabling virtual machine migration to balance load across data center. It enables highly responsive and robust provisioning in data center. It provides hardware/system maintenance, transparent mobility, work load balancing, consolidated management and high availability services. It is a administration tool which deals with situations like performing platform hardware maintenance without disrupting provisioned services, optimization of workload with provider resource pool. Xen and VMware have implemented "live" migration of VMs that involves short downtimes ranging from tens of millisecond to a second [8]. The major benefit is, it avoid hotspots, still it has some issue. That is detecting workload hotspot and initiating a migration lacks the ability to respond to sudden workload changes. And in memory state should be transferred consistently with integrated consideration of resources for application and physical servers. In Xen an attacker can gain control over guest OS or VMM due to vulnerabilities in migration module and similarly VMware exposes the sensitive information of guest OS during the VM migration[3]. So Live VM migration without security features becomes single point of failure for cloud environment.

Co-resident Attack

In cloud computing environments, in order to maximize the utilization rate of hardware platforms, it is common practice that the virtual machines (VMs) of different users run on the same physical server (i.e., these VMs are co-resident), and are logically isolated from each other. However, malicious users can circumvent the logical isolation, and obtain sensitive information from co-resident VMs[12]. If cloud providers cannot ensure data confidentiality and hence lose the basic trust

from users, the future of cloud computing will be jeopardized.

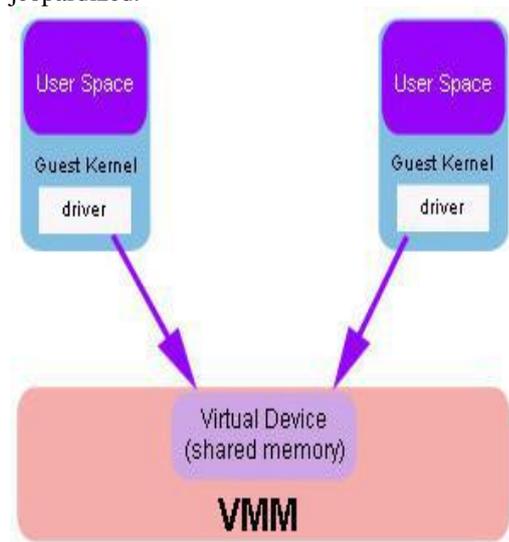


Fig 3: VM Migration

Although, in principle, programs running on co-resident VMs should not be able to influence each other, there are a variety of ways this can occur in practice. For example, the cache utilization rate has a major influence on the execution time of cache read operations. Therefore, the attacker is able to estimate the victim's cache usage by performing extensive cache read operations and comparing the execution time on a co-resident VM [10]. With similar approaches, attackers can also infer other private statistics, such as the traffic rate of a website. In addition, co-resident VMs share the instruction cache and other hardware resources. This can also be exploited by malicious users to extract private information, such as cryptographic keys[13], although it requires overcoming several major challenges. One way to encounter this kind of threat is to fundamentally eliminate the side channels between VMs[14]. However, the proposed methods require substantial changes to be made to existing commercial platforms, and hence are impractical and not suitable for immediate deployment.

IV. RELATED WORK

A. VM Migration

The importance of secure VM migration solution for IaaS clouds has earlier been identified. In that platforms within the CSP network could be malicious and untrusted which could result in security threats during VM migration. To overcome this problem Trusted Cloud Computing Platforms (TCCP) which are register with a Trusted Third Party (TTP) called External Trust Entity (ETE) are introduced which define trustworthiness of the platform during VM Launch

and Migration [3]. In Migration following requirements are analyzed

- 1) Security: That is the user VM must be protected against any unauthorized migration on vulnerable platforms, either unintentionally or intentionally so that the user can trust the complete cloud service provider's infrastructure.
- 2) Transparency: In this main goal is to keep VM migration transparent from the user without compromising on the trustworthiness of the destination cloud platform.
- 3) Scalability: There are many cloud deployment models and each can have different VM migration requirements so the secure migration solution should scale well to high number of simultaneous VM migration.
- 4) Scheduling Flexibility: For the IaaS provider to implement efficient load balancing, it must be possible for the provider to introduce scheduling mechanisms within the provide network.

To overcome these requirements in VM migration one protocol is introduced that is Trust Token which certifies the trust worthiness of a platform up to a certain level.

V. CO-RESIDENT ATTACK

Now, for co-resident attack which is also known as side channel attack or VM Hopping. For preventing side channel attack different technique are being used. In VM side channel attack requires two things they are placement and extraction. Placement refers to the adversary or attacker arranging to place their malicious VM on the same physical machine. Extraction means after successfully placement of the malicious VM to the targeted VM extract the confidential information, file and documents on the targeted VM. So for that it might be accomplished by the combination of firewall and random encryption decryption. That is for preventing placement the virtual firewall appliance in the backend of the cloud computing and for preventing extraction random encryption decryption is used [15]. But problem with this things are sometimes someone gets key of encryption/decryption somehow. So this technique is used but it is not so efficient.

Now another thing is for preventing inter-VM Traffic. Virtualization is the fundamental of the cloud computing, security and availability are critical for cloud environments because their massive amount of resources, simplifies several attacks to the cloud services. So for inter-VM traffic one new approach that gives an identity to particular traffic, this identity is about the where and who sends the request. So for that one approach which propose a frame called frame tag

that holds the proper credentials which are the tenant and the application that sends the IP packet, providing data origin authentication and integrity and also proposing an agent which is able to generate, capture and analyze particular frame and respond to it by automated acceptance or rejection and security mechanisms in order to ensure the security and integrity of the frame tag [16]. Through this inter-VM traffic is being avoided but still there is risk of co-resident attack. For allocating VMs different policies are defined they are: Least VM/Most VM policy. For every new VM request, the policy selects one server randomly from those that host the least/most number of VMs, and have enough resources left. Random policy. For every new VM request, the policy randomly selects one server from those with enough resources. In least VM Allocation policy there is less number of VMs so chance of co-resident attack is become less. But problem with this policy is someone thinks it's not reliable that's why there is less number of VMs.

VI. PLACE OF VIRTUALIZATION IN CLOUD COMPUTING

Virtualization is one of the major components of cloud computing that helps to emergence of cloud computing. To understand cloud computing it is important to understand the concepts like network virtualization or storage virtualization.

The main components of virtualization in the cloud are virtual machines, because all of the operating systems and applications are inside them [7]. They are like a container which isolated and separated from each other, even in the same physical host. Based on the basic role of cloud as-pay-you-go, the vendors give you that ability to access these provided virtual machines and in some cases they will make these virtual machines like an actual computer and you can purchase them for a limited time and take advantage of this ability without any worries about how they operate. That thing you are purchasing is the availability of these services.[13] Vendors in exchange promise to give you these services without any interrupt that make the availability in high level.

VII. NETWORK VIRTUALIZATION

In the old days the data centre technicians need to go through in the hard work of cabling, but nowadays by help of virtualization the cables are gone (of course not entirely). Instead of physical connection between two physical computers, we can virtually connect two of virtual computers together. By help if virtualization even cables converted to the virtual cables and it reduced the time which the data centre technicians spend on the cabling and maintaining the cables.

The cloud computing systems are essentially using the TCP/IP based protocols on communication [8]. It is like communication in LAN networks. Each computer has one specific, unique IP and they are communicating with each other.

Here also the same with little different, instead of computers we have virtual computers or virtual machines (VM's). The important thing here is the IP's we talking about are differs from the IP's we can configure on operating systems which we can set to our platform if we have that access (in case of IaaS service), statically or dynamically.[9] These IP's are available from pool of network which we have learned about it in network virtualization.

Network virtualization also continues the term of scalability in a cloud environment. The important point also here is a dynamic scalability of network resources. In cloud network communication, the physical devices will connect via the physical connection and they can communicate with their physical NIC's, but virtual devices use the virtual NIC's which is distributed within hosts. [13]

If two VMs want to communicate to each other through a network, there are two scenarios. The first scenario is communication between two VM on a same physical host. In this case all devices will be virtualized devices like virtual switches or virtual firewalls. As shown in the figure, the base communication will be through a hypervisor [5].

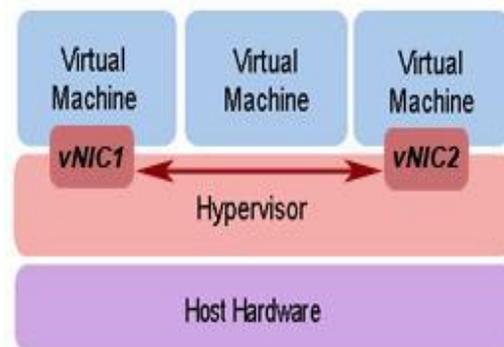


Fig 4: VMs communication on one a physical host.

In the second scenario, two VMs are on different physical host which communication is a combination of physical communication and virtual communication. As shown in the figure, the virtual NIC will be linked to the physical NIC of host one. After that the communication will be between physical NICs and at the last step, the packets will send to another virtual NIC [3].

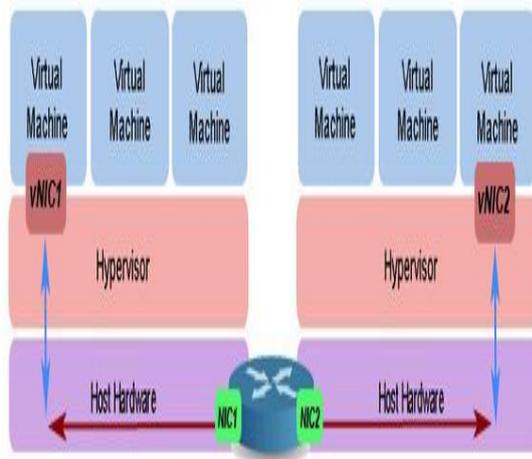


Fig 5: VMs communication on two physical hosts.

But, still there are some difficulties in this environment. One the challenges in the virtualized environment are about the IPs. As mentioned before, the IP of each VM comes from the pool of IPs. In a life cycle of a VM, the assigned IP will be selected from the pool of IPs and after the completion of the life; the assigned IP will be released [12].

Another challenge in such a network is communication between the VMs on the networks with IPv4 with other VMs on IPv6 networks. In some management software designers only considered only IPv6 or vice versa, but most of the heterogeneous data centres are must support both IP versions [9]. Thus, like all other types of networks, virtual network also and its management should consider such difficulties.

VIII. CONCLUSION

Virtualization is provided at IaaS level in cloud computing so it is one important thing. In cloud at a time so many requests are being handled so for providing it properly virtualization is necessary. But in virtualization there are some securities threats which are need to be considered. So in this paper we have discussed about that threats but in them there are main two threats they are VM \migration/mobility and VM Hooping/co-resident attack. So In this paper some techniques which are used for preventing these two problems are discussed. So that using them one can avoid the problem of VM Migration and Co-resident attack and CSP provide more efficient and secure services to its users. In this case by pushing the important concept of virtualization, we broke down the evolutionary chain between of technologies in the computer world and make the confusion. Without understanding the concept of virtualization, it is very difficult to realize the cloud computing concept.

REFERENCES

- [1] Zhang, Qi, Lu Cheng, and Raouf Boutaba. "Cloud computing: state-of-the-art and research challenges." *Journal of internet services and applications* 1.1 (2010): 7-18.
- [2] Bhadauria, Rohit, et al. "A survey on security issues in cloud computing." *arXiv preprint arXiv: 1109.5388* (2011).
- [3] Aslam, Mudassar, Christian Gehrman, and Mats Bjorkman. "Security and Trust Preserving VM Migrations in Public Clouds." *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*. IEEE, 2012.
- [4] Tsai, Hsin-Yi, et al. "Threat as a Service?: Virtualization's Impact on Cloud Security." *IT Professional* 14.1 (2012): 32-37.
- [5] Nurmi, Daniel, et al. "The eucalyptus open-source cloud-computing system." *Cluster Computing and the Grid, 2009. CCGRID'09. 9th IEEE/ACM International Symposium on*. IEEE, 2009.
- [6] Garfinkel, Tal, and Mendel Rosenblum. "A Virtual Machine Introspection Based Architecture for Intrusion Detection." *NDSS*. Vol. 3. 2003.
- [7] Aslam, Mudassar, et al. "Securely launching virtual machines on trustworthy platforms in a public cloud." (2012).
- [8] Clark, Christopher, et al. "Live migration of virtual machines." *Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation-Volume 2*. USENIX Association, 2005.
- [9] Ristenpart, T., Tromer, E., Shacham, H., and Savage, S.: "Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds," *Proc. 16th ACM Conference on Computer and Communications Security (CCS 2009)*, 2009, pp. 199-212
- [10] K. Hashizume, D. G. Rosado, E. Fernández-Medina, and E. B. Fernandez, "An analysis of security issues for cloud computing," *Journal of Internet Services and Applications* 2013.
- [11] P. Mell, T. Grance, "The NIST definition of cloud computing". NIST, Special Publication 800-145, Gaithersburg, MD
- [12] Zhang, Y., Juels, A., Reiter, M., and Ristenpart, T.: "Cross-VM Side Channels and Their Use to Extract Private Keys," *Proc. 2012 ACM Conference on Computer and Communications Security - CCS '12*, 2012, pp. 305-316
- [13] Jin, S., Ahn, J., Cha, S., and Huh, J.: "Architectural Support for Se-cure Virtualization under a Vulnerable Hypervisor," *Proc. 44th Annual IEEE/ACM International Symposium on Microarchitecture - MICRO '11*, 2011, pp. 272-283
- [14] *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249 – 8958, Volume-2, Issue-2, December 2012 "Security against Side Channel Attack in Cloud Computing" by Bhruhu Sevak
- [15] Benzidane, Karim, Sad Khoudali, and Abderrahim Sekkaki. "Secured architecture for inter-VM traffic in a Cloud environment." *Cloud Computing and Communications (LatinCloud), 2nd IEEE Latin*.
- [16] Han, Yi, et al. "Virtual machine allocation policies against co-resident attacks in cloud computing." *Communications (ICC), 2014 IEEE International Conference on*. IEEE, 2014.

Critical Study on E-Business Trends

Guided by: Mr. Pankaj Mudholkar

Rajkumar Pathak
TIMSCDR

Pankaj Rajbhar
TIMSCDR

Abstract: E-Commerce in recent times has exploded many folds due to the invention of new Information and Communication (ICT) technologies and applications. Business models of today are highly depending upon internet based transaction and trading platforms, use of Web portals and mobile applications. Proposition of new business models sans the inclusion of internet and their application has become almost impossible for business. Business data pertaining to E Commerce has become highly valuable in recent times. Social networking websites, sharing of product data and its reviews in blogs paves ways for promoting business in new methods. There is need to develop metrics to assess the strength of E Commerce penetration in Business and also evaluate the various E Commerce platforms. Use of E business tools for data analysis, prediction and decision making has become the latest order of the day. Changing trends in E business has a drastic impact on the organizations depending upon internet based applications for handling its business processes and day to day transactions.

I. INTRODUCTION

The impact of information and communication technology on varied areas of the Global economy and current trends in e business is an area of research that could lead to valuable business insights. The adaptation of business models to accommodate the customer's needs for providing an internet based business transaction platform is a reality and a compulsion for business as an order of the day. With so much necessity for business houses to include internet based marketing and applications to support one's business process and customer expectations, the area of E commerce and it's evaluation from a commercial perspective is still under researched.

The survey data from various sources of information is gathered and analyzed in the paper pertaining to internet based surveys, adoption of e commerce, e business trends and technology in order to assess e commerce success. Also the various terminologies and models on E commerce are clarified theoretically for better understanding for future researchers. Today retail organizations

are once again seeing a transformation in the way they do business.

Table 1 - E – Commerce Analytics Trends

Email Campaign Effectiveness	74%
Natural / SEO search rankings	70%
Paid search/SEM effectiveness	61%
Onsite search effectiveness	56%
Customized landing page effectiveness	55%
Social channel customer feedback	46%
A/B testing	40%
Cross-sell/up-sell merchandising effectiveness	34%
Cross-channel program effectiveness	25%
Loyalty program effectiveness	14%
Other	1%

The Research paper starts with an introduction to the topic on Ecommerce evaluation and trends. The objective of the study and research methodology is explained in the initial sections of the paper. The various perceptions on business models clarifies on understanding of business modelling and how Information technology cannot be separated from business modelling in growing globalized business. The term E Business is applied more in generic

sense for any application of information technology into business processes and hence it is explained and discussed from the view point of business modelling. Narrowing further into specific area of E commerce the research paper discusses on the E commerce metrics and elaborates on considerations of factors for evaluating E commerce platforms and applications. The top 10 trends in shaping the future of E commerce and the trends in E business evidenced in the later part of the paper serve as an eye opener for IT managers and e - tailers.

II. RESEARCH METHODOLOGY

The study is qualitative and descriptive in nature and most of the data is based on secondary sources of survey data. Such an approach is adopted in the study as the area of research is very broad and sources of data are also spread across multiple locations. To arrive at a conclusive idea of the larger picture on E- Business trends, analyzing the existing survey data would give a better result in finding the answers to the research question framed.

III. BUSINESS MODELS AND INFORMATION TECHNOLOGY

In today's fast paced business environment there are varied perceptions on Business model and its definitions.

According to Applegate (2001), Business model is a description of a complex business that enables study of its structure, the relationships among structural elements, and how it will respond to the real world. Business model (Petrovic et al., 2001, Auer & Follack, 2002) is a description of the logic of a —business system|| for creating value that lies behind the actual processes. And with growing use of Information technology for creating a competitive edge in business, even the definition of business model can include the idea of use of e commerce for Business modelling.

According to Osterwalder and Pigneur (2002), Business model is a description of the value a company offers to one or several segments of customers and the architecture of the firm and its network of partners for creating, marketing and delivering this value and relationship capital, in order to generate profitable and sustainable revenues streams. Shaffer, Smith and Linder (2005) uncovered twelve definitions published from 1998 to 2002, and they developed an affinity diagram to identify four major categories common to all or most definitions: strategic choices, creating value, capturing value and the value network.

A great deal of research has been directed towards classifying business models and grouping them into specific categories. The business models belonging to the same category usually share some common

characteristics, such as the same pricing policy or the same customer relationship model (Pateli and Giaglis, 2003). Hayes et. al (2005) by combining the work of Ticoll et al. (1998), Timmers (1999), and Kaplan and Sawhney (2000) proposed that electronic business models could be classified according to how they exhibit varying degrees of economic control, value chain integration, functional integration, business innovation and technical innovation. They developed the prerequisites framework for assisting decision makers assess the suitability of electronic business models during the intelligence phase of the decision making process and not in other phases. The basis of their prerequisites model is shown in Table 2. This framework hypothesizes that a particular business model is more likely to succeed in a particular industrial sector when the characteristics of the sector match the conditions required for the model.

A. 3 E – Commerce Metrics

A very useful representation of needs for e-commerce metrics as related to the three areas of the S shaped path for the diffusion of new technologies is presented. The idea is that there is a life cycle for research needs which follows the pattern of growth of e-commerce markets: at an initial stage there is need for information on the enabling factors and barriers to e-commerce; at a more mature stage one should look for the intensity of e-commerce use to enable policy makers to address imbalances; at a later stage one would be able to measure the impact of e-commerce on the economy and society. The three broad areas for indicators are:

1. E-commerce readiness - Included here are issues of preparing the technical, commercial and social infrastructures that are necessary to support e-commerce. It is essential for each country to be able to construct a statistical picture of the state of readiness of each infrastructure element to engage with e-commerce.
2. E-commerce intensity - These issues relate to the state of e-commerce usage, volume, value and nature of the transactions. The statistical requirement is to profile who is exploiting e-commerce possibilities and who is not, and to identify leading sectors and applications.
3. E-commerce impact - These issues relate to additionality (i.e. e-commerce goes beyond substitution effects and creates new value added) and multiplier effects. Statistics are needed to evaluate whether and to what extent e-commerce makes some kind of difference in terms of

efficiency and/or the creation of new sources of wealth.

B. Evaluating E – Commerce platforms

Selecting the right e-commerce application for the long term can be a difficult exercise. It's not easy to base a decision on both current requirements and a vague, undetermined set of future needs that have not yet even hit the planning stages. The difference between e-commerce application capabilities can spell the difference between an e-commerce site's success and failure. The white paper by Oracle (2011) offers 10 considerations to help guide the selection criteria for the next e-commerce platform—which should be the last e-commerce platform anyone ever needs to buy

C. E - Commerce trends

According to Michael Piastro (2010), the top 10 trends shaping the future of ecommerce are summarized In November 2011; Endeca (acquired by Oracle in February 2012) conducted the Trends for 2012 in Business to Consumer (B2C) Commerce survey to help businesses gain actionable insight into the evolving nature of e-commerce. When asked to identify the top three areas for investment in 2012, the respondents indicated that continued investment in the customer experience was the top priority, followed by mobile and commerce platforms

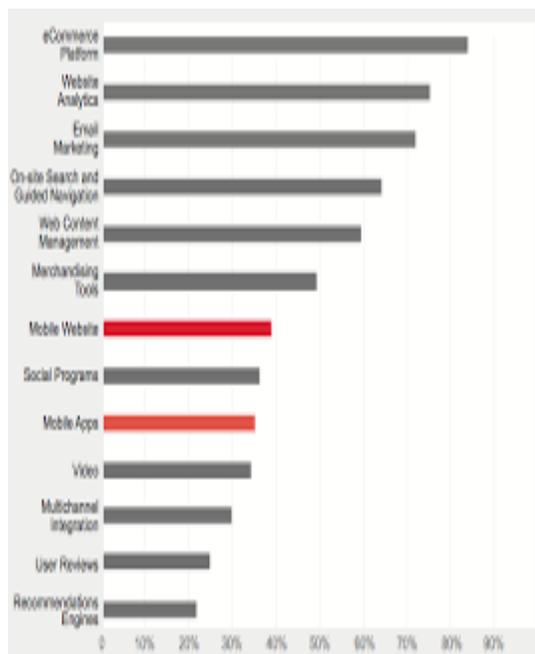


Fig.1 Trends for 2012 in Business to Consumer (B2C) Commerce survey – Oracle Survey, 2012.

Broken out from —other social programs, Face book takes the top spot for emerging areas to invest in, with 57 percent of the respondents indicating a planned investment in the social networking giant. Behind Face book, the mobile Web (47 percent), other social programs (such as Twitter or forums, with 43 percent), and mobile apps (43 percent) rank among the top emerging initiatives.

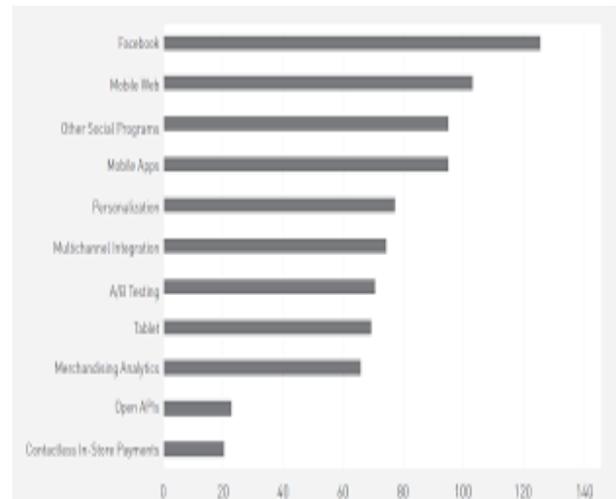


Fig:2 Impact of Face book, mobile Web and other social programs – Oracle Survey, 2012.

IV. E BUSINESS IN REAL TIME

As an alternative strategy adopted to reach the small businesses, some Information and Communication Technology (ICT) layers offer the software on demand through online access. While the road to Internet-based Software as a Service (SAAS) was a bumpy one in the early days, online services have become a credible – and often desirable – alternative to packaged software. SAAS layers are thriving in multiple market segments, from the Small and Medium Business (SMB) segment to the large enterprise sector, and in almost every application category.

As Internet-based computing becomes viable for an increasing array of individual and business requirements, customers can focus more on getting the functionality and outcomes they need from the application, and less on the underlying operating system, middleware and infrastructure requirements. The trend for the SAAS industry development is for growth, as top players such as Microsoft, Google, IBM and salesforce.com battle to build SAAS systems.

At the same time, new SAAS players and solutions will continue to emerge, and the likelihood of any one player dominating this landscape is very low.



ABOUT THE INSTITUTE

Thakur Institute of Management Studies, Career Development and Research was established in the year 2001 with a clear objective of providing quality technical education in tune with international standards and contemporary global requirements, offering 3 years postgraduate degree in Master of Computer Applications (MCA). The Institute is recognized by the AICTE norms and is affiliated to the University of Mumbai.

The Management's commitment to excellence is reflected in the marvelous infrastructure that is comparable to the finest institution of its type in the country. The sprawling campus with lawns, gardens, playgrounds, parking area, hostel accommodation and temple ensures a right academic ambience essential for a center of higher education.

At TIMSCDR, the importance of faculty is well understood which is reflected in qualified and experienced teaching staff. A closely monitored quality, assurance mechanism ensures proper coverage of syllabus within right time frame.

Application of modern technology in teaching-learning process and day to day governance of the Institute makes TIMSCDR unique. The organization supported by dedicated 16 Mbps broadband internet connectivity and also has WI-FI facility.

The Institute focuses on imparting knowledge to the students that persists even when they pass out and step into the corporate world. The syllabus has been given a new dimension through experienced faculty and state of the art infrastructure. The overall personality development through extra curricular activities like quiz, debates and seminars to name a few have been a hallmark of the Institute.



Thakur Educational Trust's (Regd.)
THAKUR INSTITUTE OF MANAGEMENT STUDIES, CAREER DEVELOPMENT & RESEARCH
(Approved by AICTE, Govt. of Maharashtra & Affiliated to University of Mumbai)

Thakur Educational Campus, Shyamnarayan Thakur Marg, Thakur Village, Kandivali (E), Mumbai - 400 101

• **Tel:** 6730 8301, 02, 28840484/91 • **Telefax:** 28852527

Email : timscdr@thakureducation.org • **Website :** www.timscdrmumbai.in • www.thakureducation.org