ISBN : 978-81-946375-7-8

Deccan Education Society's
NAVINCHANDRA MEHTA INSTITUTE OF TECHNOLOGY AND DEVELOPMENT [NMITD]
(Approved by AICTE, DTE, and Affiliated to University of Mumbai)
NAAC ACCREDITATED "B++" Grade

Organized by

Deccan Education Society's
Navinchandra Mehta Institute of Technology and Development
Dadar (W), Mumbai

Proceedings
of
NCMAT'20

Proceedings
of
NCMAT'20

EMPYREAL PUBLISHING HOUSE
India | UAE | Nigeria | Uzbekistan | Montenegro

Proceedings
of
NCMAT'20

Editors

Dr. Samadhan Khamkar
Director
DES NMITD, Mumbai

Dr. Rasika Mallya
Associate Director
DES NMITD, Mumbai

First Impression: 2020

Proceedings of NCMAT'20

ISBN : 978-81-946375-7-8

Rs. 650/- ($18)

Published by:
Empyreal Publishing House

Editorial Note

The size of the India economy is currently around INR 190 lakh crore (GDP at current market prices, 2018-19) which comes out to be $2.8 trillion, converting it in dollar terms taking average exchange rate of the financial year. This means the Indian economy needs to almost double its size in next five years to realize the target of reaching the $5 trillion mark and get global recognition as the third largest economy in the world by 2025.

The Government of India has also set an ambitious target of increasing the contribution of manufacturing output to 25 per cent of Gross Domestic Product (GDP) by 2025, from 16 per cent currently. India's manufacturing sector has the potential to touch US$ 1 trillion by 2025. There is potential for the sector to account for 25-30 per cent of the country's GDP and create up to 90 million domestic jobs by 2025.

Further, India is ready to take a quantum leap globally, due to service sector. Service economy is now considered the golden key to India's future. Services sector contributes mora than 60% to India's economy and 28% to the total employment. Attracting highest FDI inflows, services sector is on a growth trajectory driven by digital efforts of the government and highly skilled and low cost manpower.

Current advances in technology present newer avenues for growth. It is interesting to note that in the next 5 years, global application technology market is expected to be worth US$86 billion and the Indian application technology market is also expected to grow at 12% CAGR and touch US$4 billion. The Indian Information Technology sector is visualizing their growth through emerging technologies like 5G, Block Chain. New developments like 4.0 and 5G technologies, block chain faces global headwinds like Brexit but developments around Globalisation 4.0 and 5G technology will bring growth for the industry this year.

The Big Data Analytics market in India is currently valued at $2 Billion and is expected to grow at a CAGR of 26 percent reaching approximately $16 Billion by 2025, making India's share approximately 32 percent in the overall global market. The IoT market shows a similar growth trajectory and is fore casted to reach $15 Billion by 2020

The objective of this special issue of our journal is to provide a platform for deliberations and constructive dialogue on strategies, policies and issues pertinent in India's growth story. This will be achieved through multi-disciplinary views published as papers and deliberated in the proceedings from both academics and practicing executives from industry.

Table of Contents

 ISBN: 978-81-946375-7-8

Recommendation System for E-commerce using Data-Mining Techniques

Drashti Burman and Diksha Sihotra
MCA, Deccan Education Society's Navinchandra Mehta Institute of Technology and Development, University of Mumbai

Abstract
Nowadays, the internet has grown up as a familiar medium to investigate new data. Several organizations and companies getting on the web to present their products or services across the world. E-commerce is a sort of business or saleable exchange that involves the exchange of statistics over the internet or web. In this circumstance, an immense amount of information is acquired and dumped into web services. Recommender systems are changing from novelties utilized by a couple of E-commerce sites, to serious business tools that are reshaping the universe of E-commerce. The greater part of the most significant commerce websites is already using recommender systems to help their customers discover products to get. Recommender systems are changing from novelties utilized by a couple of E-commerce sites, to serious business tools that are re-shaping the universe of E-commerce. The greater part of the most significant commerce websites is already using recommender systems to help their customers discover products to get. By using a data mining algorithm system that will trace customer shopping behaviour learn his/her up-to-date preferences adaptively based on association rule mining using apriori algorithms extracts the varied useful information from the massive data sets. So implementing the above data processing technique in JAVA and data sets are dynamically generated while the transaction is processing and extracting various patterns.

Keywords: E-commerce, Recommended System, Data Mining, Association Rule Mining

Introduction

There are various types of items or materials that are acquired or utilized in everyday lives. The products have product specifications, the customer contrasted the items and comparable characteristic on the Internet, read the criticism from unknown clients, and afterward settling on choice about the item to purchase. Subsequently, the Web is changing from a basic data space to a progressively complete market space introducing a wide assortment of business administrations, going from electronic Web stores to barters, online booking and different administrations [1]. In this specific circumstance, clients may explore among thousands Websites to explore desired sources and make their purchases. However, to perform the varied tasks typically involved in e-Commerce transactions, a customer has often to spend an outsized amount of your time on the online [1][2] . On the opposite hand, electronic suppliers have similar problems for proposing their products to customers within the best suited way, taking under consideration their preferences, habits, etc[3].

The largest E-commerce sites offer a large number of items available to be purchased. Choosing among numerous options is challenging for consumers. Recommender frameworks systems have emerged in response to the present problem[4]. A recommender framework for an E-business site prescribes items that are probably going to meet her requirements[4][5]. Recommender frameworks are utilized by E-commerce sites to suggest products to their customers. The products can be recommended dependent on the highest overall sellers on a site, given the socioeconomics of the client, or dependent on an examination of the past purchasing conduct of the client as a prediction for future buying behaviour[5]. Extensively, these strategies

are a part of personalization on a site, since they assist the site adjust to every client. Recommender frameworks automate personalization on the Web, enabling individual personalization for every customer[6]. It additionally help in suggesting, for example, like useful products on e-commerce websites, videos on YouTube, friends' recommendations on Facebook, book recommendations on Amazon, news recommendations on online news websites, and therefore the list goes on[6].

Often e-Commerce systems act also as recommender systems that generate some recommendations computed by using different possible approaches as:

1. Content Based: Recommending to a customer those products that appear most like those he/she already accessed within the past[8].
2. Collaborative Filtering: Searching similarities among customers and consequently suggesting to a customer some products also considered by similar customers within the past[8].
3. Hybrid: Using both content-based and collaborative filtering techniques to get recommendations. Generally, these systems use a profile of the customer to represent his/her interests and preferences, and lots of them propose the utilization of software specialists to construct such a customer's profile[8].

To overcome the disadvantages of collaboration filtering, the recommender system based on data mining is proposed in the paper. It uses a variety of data mining techniques such as web usage mining, association rule mining etc. Based on these, the system could trace the customer's shopping behaviour and learn his/her preferences. Therefore, the paper is harmonize as follows. Section 2 provides a summary of literature related to recommended system, Section 3 provides research design of the algorithm with its implementation. And the Section 4 provide conclusion and future scope of the paper[9].

Literature Review

A several works were done to improve e-commerce efficiency through web use mining. Web based mining utilizing basic association rule proposed by Mei-Ling Shyu has drawbacks include generation of irrelevant rules, generating too many rules contradict the prediction of results, resulting reduction in accuracy[10]. Summary of several data mining approach for improving e-commerce's working efficiency. Collaborative filtering utilizing k nearest neighbour is proposed in [11], has drawback of time complexity in dynamically finding k nearest neighbour. Moreover, personalization through association rule mining is proposed by Mobasher [11] utilize multiple support and confidence. [12] In, recommendations are made for e-customers through k-mean clustering and Apriori algorithm for association rule mining. It utilize cosine distance measures. The cluster tightness isn't acceptable as distance isn't updated dynamically when new object enters into clusters. In the proposed work homogeneity of clusters determining highly similar user groups through simple computations. Khalil gave Markov model for web access forecast in which for analysis of every test session, all training data are considered, some of them can be less relevant to test session which affects prediction accuracy [13].Also, Markov model utilise only strict consecutive and sequential page access for matching session during prediction. It may lose some of the loosely connected but interesting sessions.

Research Design

Data mining is characterised as the mining of the unknown information from the accessible past information and utilizing of it for estimating the upcoming activities. At present data mining has been executed in several industries such as retail, manufacturing, health care, banks and telecommunication [14]. From last few decade data mining has played the significant roles in

the sales of the merchants and also offer to expand the benefits by introducing statistical and computational techniques [14]. Data mining helps define the patterns by utilizing those statistical methods such that the retailer can differentiate about the buying percent of every item, so that the retailer can evacuate the undesirable products. On the Bases of the patterns mined the sellers can have a structured arrangement of the items in the store with the end goal that the purchaser will appear consideration in purchasing the items for whenever. Consequently, the hit count rate of the website will add the profit for the organization by encouraging some announcements on the website [14].

Other technique in data mining is classification of the existing data. Classification of data is done by making use of some fixed attributes such that we can decrease the data to be processed for every transaction. The profits in recognizing interest of the customer based on his/her attributes taken from his/her previous transactions from the site [10]. Majorly the classification is utilize for the estimation in mining of the given sales data. The labelled data and classifying the data will generate some patterns about that data taken so that the up-coming events can be recognized and processed accordingly. Association rule mining is used to find the frequently items sold and the relationship between those items, the outlier items that have to be excluded from the stores. The association define as the direct relation between two attributes and their dependencies [15]. It is a computational method where we can find the parameters such as support and confidence. These two will decides the type of relationship between the items and the customers. Not just a high confidence and support is useful but also a low support and confidence will determine an unusual behaviour of transactions which can predict the fraud from the customers. Apriori algorithm is utilized for mining the item sets which occur frequently by some Boolean expressions on the data set given. Apriori algorithm is an iterative search algorithm which finds the strong relations by finding the maximum confidence and support [15].

Java platform is utilized for the developing web pages dynamically hence whole data deposited can be stored by using MY SQL and HTML pages and are designed to provide more efficient inter-active web pages .Therefore statistical analysis of the data is done by applying the data mining technique. Most of the data miners use JAVA platform for knowledge mining techniques for faster results [16]. JAVA is available in several forms and also to design JSP pages can combine different formats of statistical methods into a single program. In JAVA the algorithms can be written in servlet programs so that the users could easily code and extract the required data. Additionally JAVA consists of some in built functions and libraries so that the default cases were avoided from the user to write large programs. The operations can done on the active memory of the computer. To gather the statically information and strong web mining we need to prefer some built in algorithms of JAVA. The approach in this paper is executed by using JAVA platform by loading the data dynamically CVS file into the JAVA. By applying association mining based on ontology by using JAVA built in packages for obtaining certain patterns[16]. The statistical data which is taken form the result of the algorithms specifies the retail traders to perform the future process. Java provides mathematical formulas required that are in built in R for pre-processing.[17].

Apriori algorithm

It is utilised for finding the frequent itemset by using the Boolean values. The main aim of the apriori algorithm is that the subset of a frequent itemset must be frequent. Let's consider an item set comprises of a list of items and their transaction id's [17]. So the minimum support and confidence are specified to avoid unwanted items and improve the efficiency of frequent itemsets based on the support and confidence computational results [17]. This procedure is applied recursively until it found the association list with same support and confidence and

arrange all the items in the same place or give out the recommendation of outstanding products on buying one product from the associated item list [17].

Step 1: Consider the data set with their transaction ids.[17] Step 2: Specify the minimum support and confidence. [17]

Step 3: Find the frequent items by using confidence and support. Exclude the items which don't have the minimum support and confidence.[17]

Step 4: Next iterate the similar procedure is done by taking the combinations of the selected items.[17]

Step 5: Repeat this procedure until there is no set with the association.[17] Step 6: Find the resultant set which is associated.[17]

Step7: Organize them in the same place.[17]

Ontology based association mining

Ontology is a representation of official knowledge. It gives a consistent representation of language and method that help the individuals to detect difficulties and handling with affairs, deliver public languages and determine different levels of formal implications of terms and relationships between terms.[17][18] It is systematized by taxonomy and contains the particular type of model of the native language of the ontology which provides a public and reliable understanding of the area. It defeats the semantic content of the communication and mismatch problem.[18] Ontology structure is categories into six levels:

1) Recognize the reason and goal of the ontology application: The area of study ontology can be determine based on the domain of ontology.[18]
2) Ontology Investigation: Determine the relationship between a wide range of Ontology and terms.[18]
3) Ontology Representation: It can select the reasonable method of ontology which allows the System necessity.[18]
4) Ontology testing: It tests the quality, reliability, stability and scalable. [18]
5) Ontology Structure: It will check the ontology permitting to the above principles, to encounter the necessities, where it can store the files, or else it will switch to the next process.[18]
6) Ontology representation: To represent Ontology many other languages are in existence.[18]

Applications of Ontology-based Web Mining

1) Web search improve
2) Better browse capability
3) Personalizing Web data access.

The principle of the recommender system is to get the customers' preferences using web mining techniques, to provide decision support for their web shopping.[17][18] Figure 1 gives an outline of the personalized recommender procedure of the framework. We select some customers as the target customers for providing recommender services, considering the efficiency and accuracy of the system running and maintenance. The recommender process comprises of three phases as shown in figure 1. After required data cleansing and transformed in the usable form for the system, the preferences of target customers are mined first in phase

1. In this phase, customer's previous shopping behaviour is trace effectively it is important as it can use to make preference analysis. In phase 2, distinct association rule based on ontology are mined from the customer purchase database, combining it we can use for discovering product associations between products.

In phase 3, by using match algorithm with uses Apriori algorithm to generate frequent match to customer preferences and product associations discovered in the previous two phases, so the recommended products list consists of the products with the highest scores(confidence and support), are returned to a given target customer.[17][18]

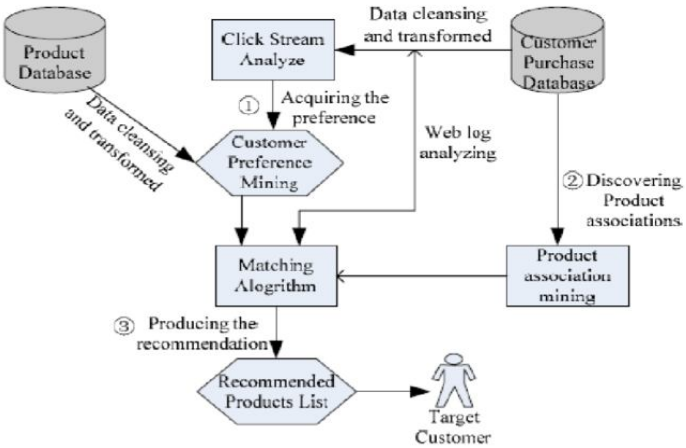


Figure 1 Overview of Recommended System.[1]

Implementation

There are different platforms to implement various data mining techniques for analysis of Big Data are Hadoop, R, weka and python etc. Numerous organizations use Hadoop for better data analysis [17]. By utilizing the Hadoop for any E-commerce business we can take the input data, process the data, manage the data, analysing the information and storing the required statistics for the fast organization growth. For predictions from the given data, we use weka for real-time applications [17]. Weka utilizes a graphical user interface and the command-line interface. Therefore it becomes faster and easier for an analyst to process in weka. Python is a programming language which is easy to learn and it is a flexible language that can be utilized for analysis and manipulation of data [17].

CSV file data

Figure 2 CSV file data.

 ISBN: 978-81-946375-7-8

Calculation of Confidence and support for each transaction

```
All transcastions loaded into memory.

[Nuts, ROYAL], 18.181818181818183%
[Onions, lipstick], 18.181818181818183%
[Onions, Grey], 27.272727272727727%
[Salt], 27.272727272727727%
[Salt, ROYAL], 18.181818181818183%
[Onions], 63.636363636363636%
[ARAMIS], 18.181818181818183%
[Salt, Carrots], 18.181818181818183%
[Nuts], 36.363636363636637%
[Onions, Nuts], 18.181818181818183%
[Milk-Orange], 18.181818181818183%
[Grey], 27.272727272727727%
[Milk-Orange, Onions], 18.181818181818183%
[Carrots], 27.272727272727727%
[ROYAL], 36.363636363636637%
[Onions, ROYAL], 27.272727272727727%
[Onions, Nuts, ROYAL], 18.181818181818183%
[BeetRoot], 18.181818181818183%
[lipstick], 27.272727272727727%

==High-confidence association rules( min_conf=30.0% )
[ROYAL] => [Salt] (conf:50.0%)
[Carrots] => [Salt] (conf:66.666666666666667%)
[lipstick] => [Onions] (conf:66.666666666666667%)
[Grey] => [Onions] (conf:100.0%)
[Nuts] => [Onions] (conf:50.0%)
[Milk-Orange] => [Onions] (conf:100.0%)
[ROYAL] => [Onions] (conf:74.999999999999999%)
[Nuts, ROYAL] => [Onions] (conf:100.0%)
```

Figure 3 Calculation of Confidence and support for each transaction.

Preference analysis based on Association Rule

```
All transcastions loaded into memory.

==High-confidence Associations
[Salt], 21.428571428571427%
[Onions], 50.0%
[lipstick], 21.428571428571427%
[BeetRoot], 14.285714285714285%
[Onions, Nuts], 14.285714285714285%
[Nuts], 28.571428571428571%
[Salt, Carrots], 14.285714285714285%
[Carrots], 21.428571428571427%

==High-confidence association rules
Medium Rules: [Carrots] => [Salt] (conf:66.66666666666666%)
Medium Rules: [Salt] => [Carrots] (conf:66.66666666666666%)
```

Figure 4 Preference analysis based on Association Rule.

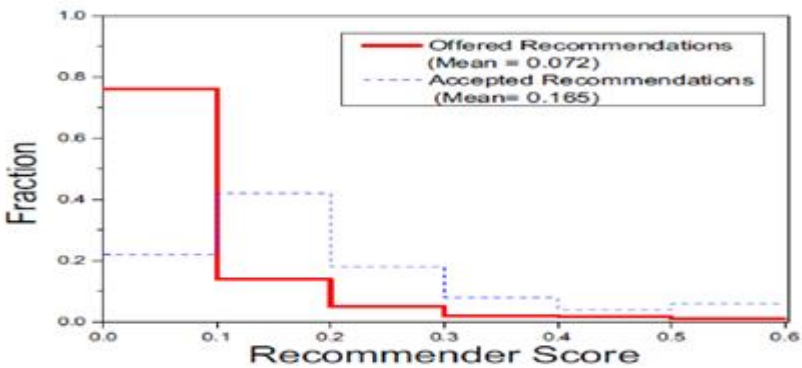


Figure 5 Distribution of score offer and accepted recommendation.

One major issue for assessing the recommender quality is the extent to which recommendations with higher recommender scores (confidence and support) are accepted preferentially over recommendations with lower scores (confidence and support). By addressing this issue by comparing the distribution of scores by finding the mean for accepted recommendations with the analogous distribution and offered recommendations. The results are shown in Figure 2. The scores depends on accepted recommendations are based on 120 products accepted from 50 distinct recommendation lists. The distribution of the offered recommendations is taken from about 300 recommendations made to the customers who accepted at least one recommendation during the primer phase of system running. Figure 2 shows that the scores (confidence and support) of the accepted recommendations are higher than the scores (confidence and support) of a large number of offered recommendations. For instance, 76% of the products placed onto the recommendations lists have scores below 0.1, yet just 22% of the accepted recommendations fall in this lower span. The mean of the offered recommendations are 0.072, while the mean scores for the accepted recommendations are 0.165. The difference between the two means is 0.093, falls well inside the 95% confidence interval (0.089, 0.106).

Conclusion and Future Work

Recommender systems are a strong new technology for extracting additional value for a business from its customer databases. These frameworks help customers find products they need to buy from a business. Recommender frameworks benefit customers by enabling them to find products they like. Conversely, they assist the business by generating more sales. Recommender systems are rapidly becoming a crucial tool in E-commerce on the online. Recommender framework are being stressed by the large volume of customer data in existing corporate databases, and will be stressed even more by the increasing volume of customer data available on the Web. New technologies are needed which will dramatically improve the scalability of recommender systems. This paper discuss about development of a personalized model by using data mining techniques such as association rule mining based on ontology in JAVA platform to predict the customers behaviour, purchasing habits, based on different parameters such as age, status, occupation, gender, brand etc., Recommendation frameworks can easily identify which age group persons buys what kind of product, which brand is rating more in market and can do analysis of various items' sales and pricing limits for various customers based on their age and also identifies which item sets are frequently purchased by the customers. Hence by providing the rough scenario of predicting the customer's needs and

can provide the personalization when new user login/sign-up on the e-commerce sites. So the ecommerce sites can increase their sales on various products and which can provide huge profits to the E-commerce organization. In future, we can combine our methodology by clustering algorithms as well as prediction generation schemes to improve the prediction quality.

References

1. Sarwar, Badrul M., et al. "Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering." Proceedings of the fifth international conference on computer and information technology. Vol. 1. 2002.

2. Schafer, J. Ben, Joseph Konstan, and John Riedl. "Recommender systems in ecommerce." Proceedings of the 1st ACM conference on Electronic commerce. ACM, 1999.

3. Rosaci, Domenico, and Giuseppe ML Sarné. "A multi-agent recommender system for supporting device adaptivity in e-Commerce." Journal of Intelligent Information Systems 38.2 (2012): 393-418.

4. IA, Hassan. "EVALUATION OF ACCURACY BETWEEN ITEM-BASED AND

MATRIX FACTORIZATION RECOMMENDER SYSTEM." Annals. Computer Science Series 17.1 (2019).

5. Billsus, Daniel, and Michael J. Pazzani. "Learning Collaborative Information Filters." Icml. Vol. 98. 1998.
6. Kazanidis, Ioannis, et al. "Proposed framework for data mining in e-learning: The case of open e-class." IADIS AC (2). 2009.
7. Sarwar, Badrul M., et al. "Recommender systems for large-scale e-commerce: Scalable neighborhood formation using clustering." Proceedings of the fifth international conference on computer and information technology. Vol. 1. 2002.
8. Breese, John S., David Heckerman, and Carl Kadie. "Empirical analysis of predictive algorithms for collaborative filtering." Proceedings of the Fourteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., 1998.
9. Goldberg, David, et al. "Using collaborative filtering to weave an information tapestry." Communications of the ACM 35.12 (1992): 61-71.
10. Babu, B. Sekhar, P. Lakshmi Prasanna, and P. Vidyullatha. "Personalized web search on e-commerce using ontology based association mining." International Journal of Engineering & Technology 7.1.1 (2018): 286-289.
11. Mobasher, Bamshad, et al. "Effective personalization based on association rule discovery from web usage data." Proceedings of the 3rd international workshop on Web information and data management. ACM, 2001.
12. Khribi, Mohamed Koutheaïr, Mohamed Jemni, and Olfa Nasraoui. "Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval." 2008 Eighth IEEE International Conference on Advanced Learning Technologies. IEEE, 2008.
13. Khalil, Faten, Jiuyong Li, and Hua Wang. "Integrating recommendation models for improved web page prediction accuracy." Proceedings of the thirty-first Australasian conference on Computer science-Volume 74. Australian Computer Society, Inc., 2008.
14. Book ”Data science and big data analytics” by Wiley.
15. Khattak, Asad Masood, et al. "Analyzing association rule mining and clustering on sales day data with xlminer and weka." International Journal of Database Theory and Application 3.1 (2010): 13-22.
16. Matloff, Norman. "The art of R programming." http://heather. cs. ucdavis. edu/~ matloff/132/NSPpart. pdf>. Acesso em 1.03 (2009): 2018.
17. Jovic, Alan, Karla Brkic, and Nikola Bogunovic. "An overview of free software tools for general data mining." 2014 37th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, 2014.
18. Venkatachari, Kavitha, and Issac Davanbu Chandrasekaran. "Market Basket Analysis Using FP Growth and Apriori Algorithm: A Case Study of Mumbai Retail Store." BVIMSR’s Journal of Management Research 8.1 (2016): 56.

 ISBN: 978-81-946375-7-8

IOT with Fog Computing

Abhishek Ambokar
Student, MCA, DES Navinchandra Mehta Institute Technology and Development, Dadar (W) / University of Mumbai

Archana Mestry
Student, MCA, DES Navinchandra Mehta Institute Technology and Development, Dadar (W) / University of Mumbai

Dr. Rasika Mallya[3]
Associate Professor, MCA department, DES Navinchandra Mehta Institute Technology and Development, Dadar (W) / University of Mumbai

Abstract
With the huge, demand to the Internet of the thing (IoT) Application. The centralized cloud computing faces many challenges such as high latency, low capacity and network failure. To address these challenges hence introduced fog computing which bring cloud closer to IoT devices. Fog computing stores the data locally in IoT devices instead of directly stored to cloud, Also fog computing provide services with faster response and greater quality. Therefore fog computing can be considered as a best choice which provides efficient and secure services to many IoT users.

In this paper, prove that fog computing is appropriate platform for number of critical IoT application namely connected car, smart traffic light, smart homes and healthcare.

Keywords: IoT, Cloud computing, fog computing, Fog Node, real time

I. Introduction

Fog computing is sometimes it is termed as edge computing it is growing it application with many other technology .The fog is distributed architecture, where storage control of device and communication are delivered to the user in a seam less way. IoT provides a best way to implement and utilize fog computing Architecture .Internet of Things is ever growing and fast capturing all the sectors of the industries .The Iot integrate hardware, computing machine, software and any kind of living and non-living objects which helps to communicate and share data. Fog computing basically process data on the locally on network nodes, which use to be done on the centralized cloud servers. It not only support the cloud paradigm it also address the unfit and unsupported application by cloud computing paradigm. As the cloud computing is not the replacement its only which pushes relevant data to the cloud. The devices which support or install on the network for fog computing are termed as fog nodes. Any device with processing power, storage and connectivity is fog node. As a nontrivial extension of cloud computing, it is inevitable that some issues will continue to persist, especially security and privacy issues [1]. Fog nodes are installing and deployed on network by different standard and on standard fog service provides as the security may be compromised it should be addressed in some way.

II. Literature Review

2.1 IoT Device

Using large number of sensors and actuators it will support for collection, Analysis, and Control the data. The IoT devices are efficient data handler like to minimize raw data and maximize actionable information. It will support connectivity and communication between sensors and cloud. It provides security, minimal latency and fault tolerance.

 ISBN: 978-81-946375-7-8

It consists of three tier architecture fig. [2.1].

The first layer is the perception layer which is built of sensors, actuators and edge device which will be interacting with the environment.

The second layer is the network layer which will be used for coordination with application layer also it will connect and discover the devices in a network.

Third layer is the application layer which consists clouds system or service. It will process the data or store the data.

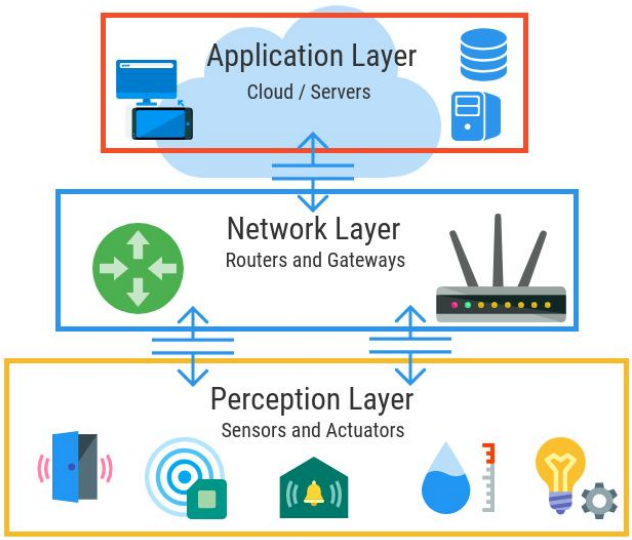


Fig. 2.1 Three tier Architecture cloud computing [2]

1.2 Cloud Computing:

The collision of IoT and clouds known as the CoT (clouds of things).

Fig. 2.2 Cloud computing [3]

The CoT simplifies the flow of data gathering processes and provides quick, low cost installation and integration for complex data processing and deployment [4, 5].Because of the IoT device which collects a large number of data it is difficult to store therefore, for the analyzing and storing data cloud computing is used. Using CoT it will help to cost effective and efficient. CoT having lots of issues like performing the task is the time sensitive or the internet is poor. “This is the scenario for vehicle to vehicle common avoiding collision or accidents cannot tolerate the latency cost by centralizing cloud approach.”[6]

Therefore, to handle the problems of CoT the Cisco suggests new technology “fog computing”.

1.3 Fog Computing

Fog computing is a paradigm with limited capabilities such as computing, storing and networking services in a distributed manner between different end devices and classic cloud computing. Fog computing was originally coined by Cisco. Fog computing provides many services to IoT users like data processing and storing. Fog computing stores the data locally instead of sending it to the cloud. The purpose of fog computing in an IoT is to improve performance, efficiency and fewer amounts of data is sent to cloud for processing, analysis and storage. Fog computing, act as a middle layer between cloud and IoT devices.

Fog computing has the ability to provide real-time and efficient communication between different IoT application.

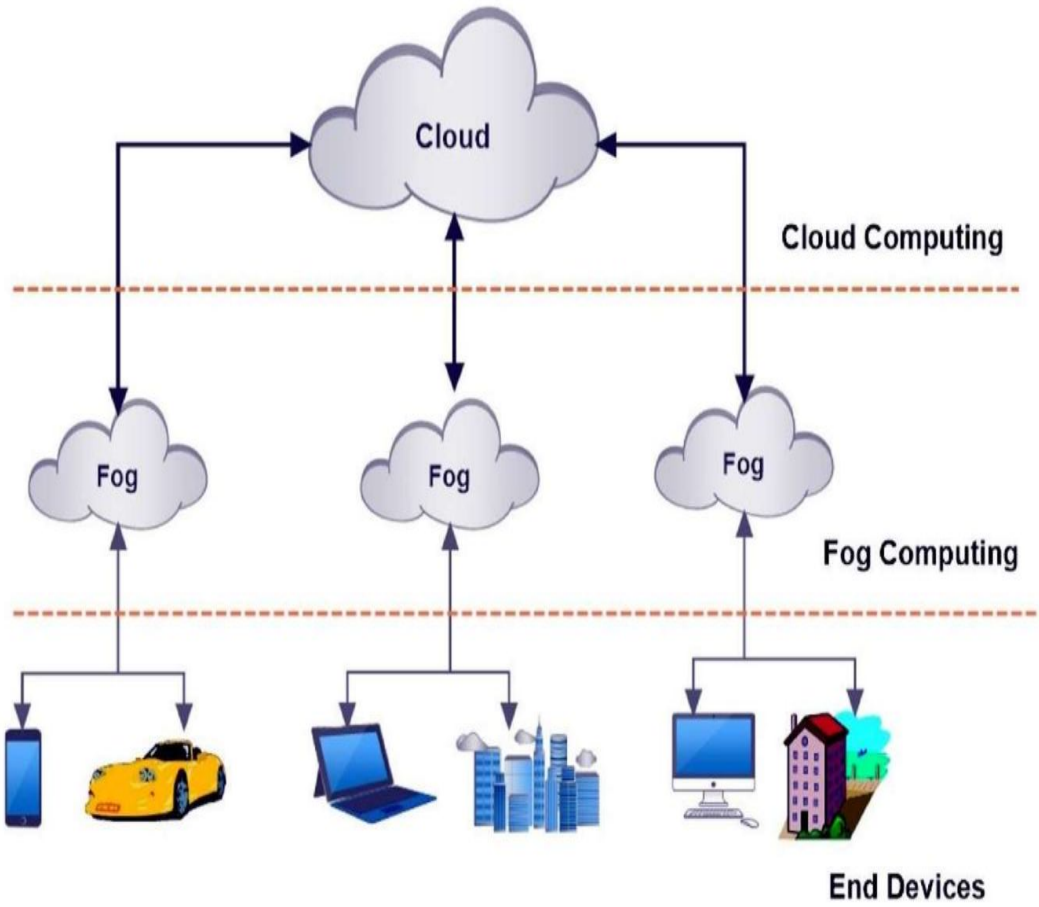


Fig. 2.3 fog computing [7]

 ISBN: 978-81-946375-7-8

1.4 Fog computing architecture

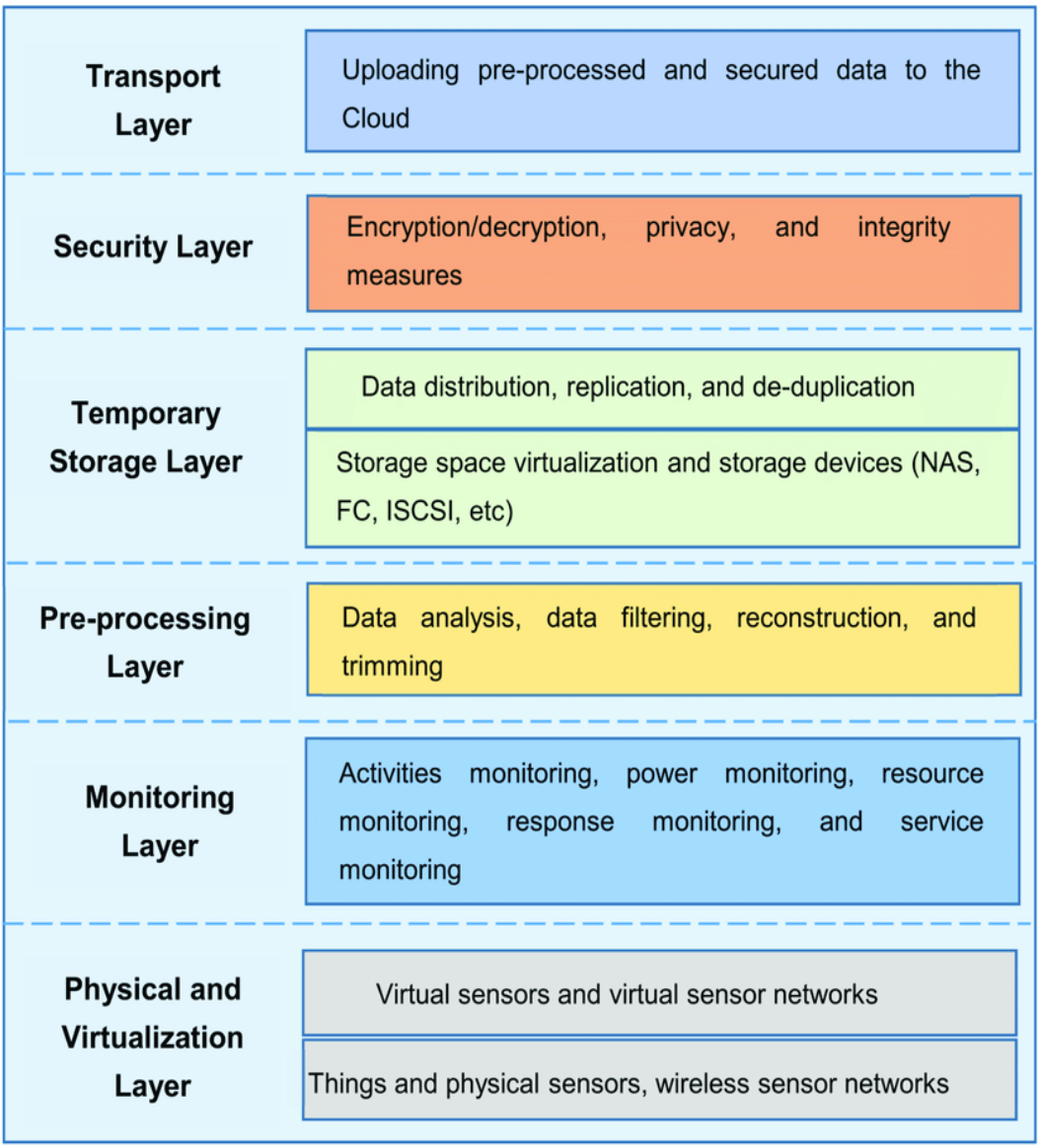


Fig. 2.4 Fog computing six layer architecture [8]

As shown in Fig. [2.4] architecture of layers is as follows:

The physical and virtualization layer involve many different types of nodes and virtual sensors network. All these nodes are managed and maintained depending on their type and demands. All these sensors are distributed geographically to sense the environmental things and send these all collected data to upper layer via gateway for further process.

At the monitoring layer all node and network elements are monitored which node is performing which task at what time what will it required. The performance of all application is monitored.

Finally in transport layer, the data from pre-processed layer are uploaded to cloud. For efficient power utilization only the collected data is uploaded to cloud.

1.5 Without fog computing and with fog computing:

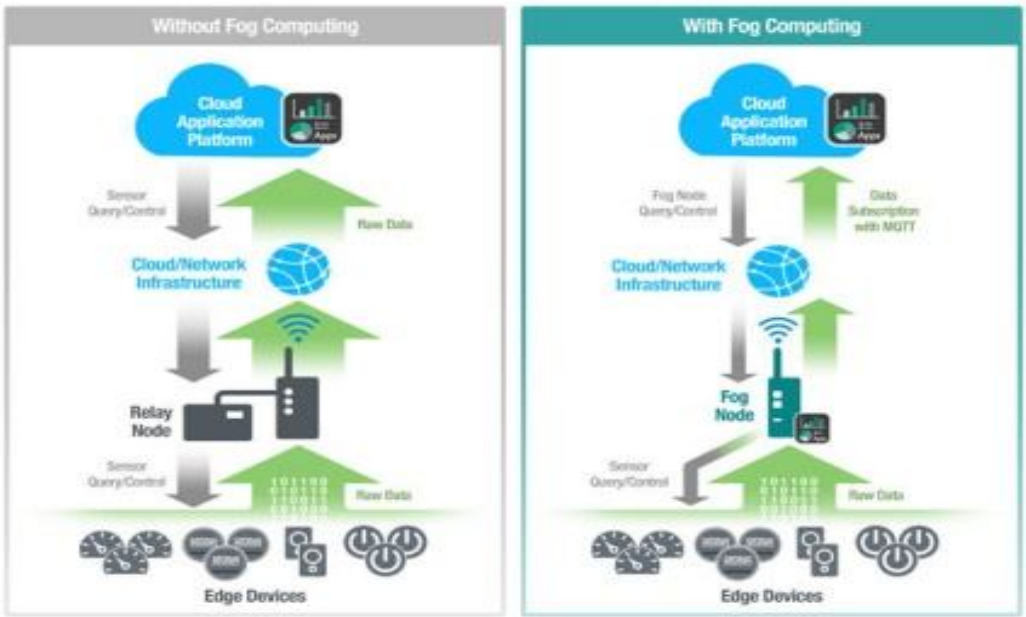


Fig 2.5 difference between with fog and without fog [9]

1.5.1 Connected cars

Many automatic cars are available now a day, this car has in build IOT devices, which can interact with other cars through internet. IOT devices have limited storage and processing power and store the data directly to the cloud. Traditionally centralize cloud have many issues as it take maximum processing time because it store data to the single location and network failure. Connected car required real-time response for better working of car and to avoid an accident.

2.5.2 Smart traffic light

Smart traffic light is use to control the traffic. IOT sensors are used in traffic signals to change the signal according to vehicles queue on the road. IOT device are work with cloud computing which store the data at single place. To retrieve the data from single location take large time but smart traffic light need real-time response to avoid traffic jam.

2.5.3 Smart home

In our day to day life, use many sensors and devices which are connected in our home like lightning, heater, air-conditioner (AC) etc. However these devices came from different platform and to make them work together is difficult and in addition some task need huge amount of computation and storage.

2.5.4 Health care monitoring System:

Before the IoT devices or computing, interaction between patient and doctor the person has to visit to the clinic there were no other way for communication. It is not possible for the doctor to monitor the health condition of the patient and provide proper treatment to them as early as possible. Nowadays the communication is more efficient, easy and accessible it will increase the satisfaction of the patient However, these are possible because of IoT devices and computing models. Because of remote health monitor system the hospital stay and readmission will be less. For better understanding healthcare is divided in two parts Concept of healthcare and Application of healthcare

III. Application of fog computing with IOT

3.1 Connected car:

Now a days there are many automated cars are available, which did not need any person to direct as well as automatic steering and "hand free" operations are performed which are also known as Driverless Cars.

Now it is possible that all cars can communicate with nearby cars and the internet. A connected car is a vehicle with internet access, which interact with each other's and the devices outside the cars, houses and surrounding infrastructure using sensors.

Connected car is solution of vehicle as it can communicate with smart city, it can get real-time information about road conditions, accidents, traffic, and current weather. This data help to change the route, avoiding traffic jams and possible accidents, also the vehicle can gather the information about the objects on the road around it, speed breaker and other vehicles. It can also suggest you the effective routes to wherever you want to go, avoiding traffic jams, busy streets or road that are under constructions which save your travel time [10].

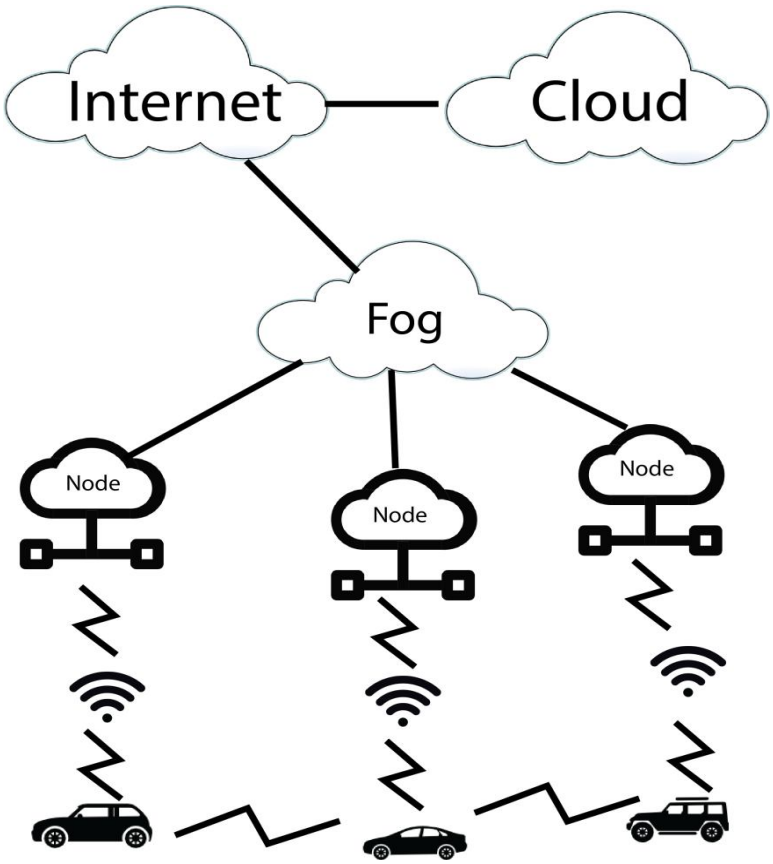


Fig. 3.1 Connected cars using fog computing

3.2 Smart traffic light:

As city population grows, transportation become critical, Smart traffic systems are way to handle this traffic conditions.

Smart traffic light interact with many sensors, it senses the presence of walkers, bikes and cyclists and measure the distance as well as speed of nearby vehicles. It also interacts with nearby lights to co-ordinate the green traffic.

Cameras and queue detector detect the real-time traffic conditions on busy roads. After a specific period of time control system calculate whether it is necessary to adjust the traffic light and then adjust it to reduce the number of queue of vehicles to improve the efficiency of traffic and reduce the pollution.

Sensors are mounted under roads as a part of smart traffic system which detects the location and speed of the vehicle and the sensors can send the emergency alerts when accident has taken place also it notify the drivers of upcoming hazards [11].

Smart traffic light can be consider as a fog node which synchronized with each other smart lights to send warning signals to nearby vehicles.

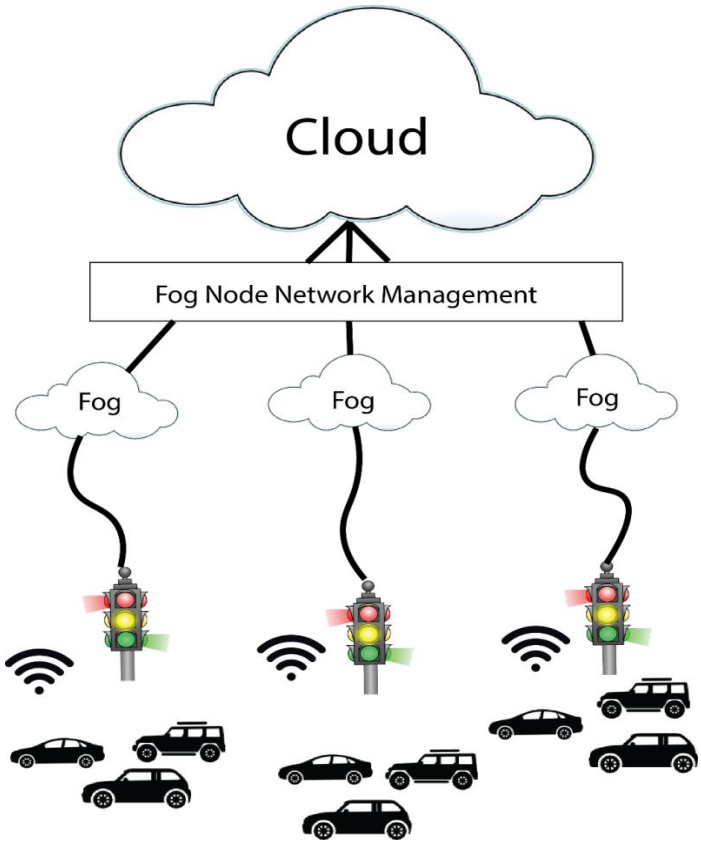


Fig. 3.2 Smart traffic light using fog computing

3.3 Smart Home

There is constant increase in the IoT sectors. The homes are buzzed with the sensors and devices. In market there are lots of options for one sensor regarding price and quality of the sensor. Every vendor has its unique API built for own devices use always it has some extra or different implementation. Some process requires huge amount of processing and storage. E.g. video analytics are infeasible it required more amounts of graphics processing which is not possible for embedded devices. So, to solve these kinds of problems using fog computing which help to integrate everything into a single platform of home application. It mainly focus of the fog is the classic resource.

1. Unified Interface:

Fog target to provide uniform platform for integration of devices from different vendors.

2. Flexible resources:

The switching between the required computation power and storage as per application utilization. There are rapid ups and downs in the resources fog computing better handle it. Sometime in security system there might be constant open and close in the security locks. So it better to provide it computation on the fog node rather than the cloud.

3. Low latency response

For security system data processed at cloud sometime due some network delay has great flaws in it. The thief might run away till alarm blows. So it better to have low latency and provide faster real time response.

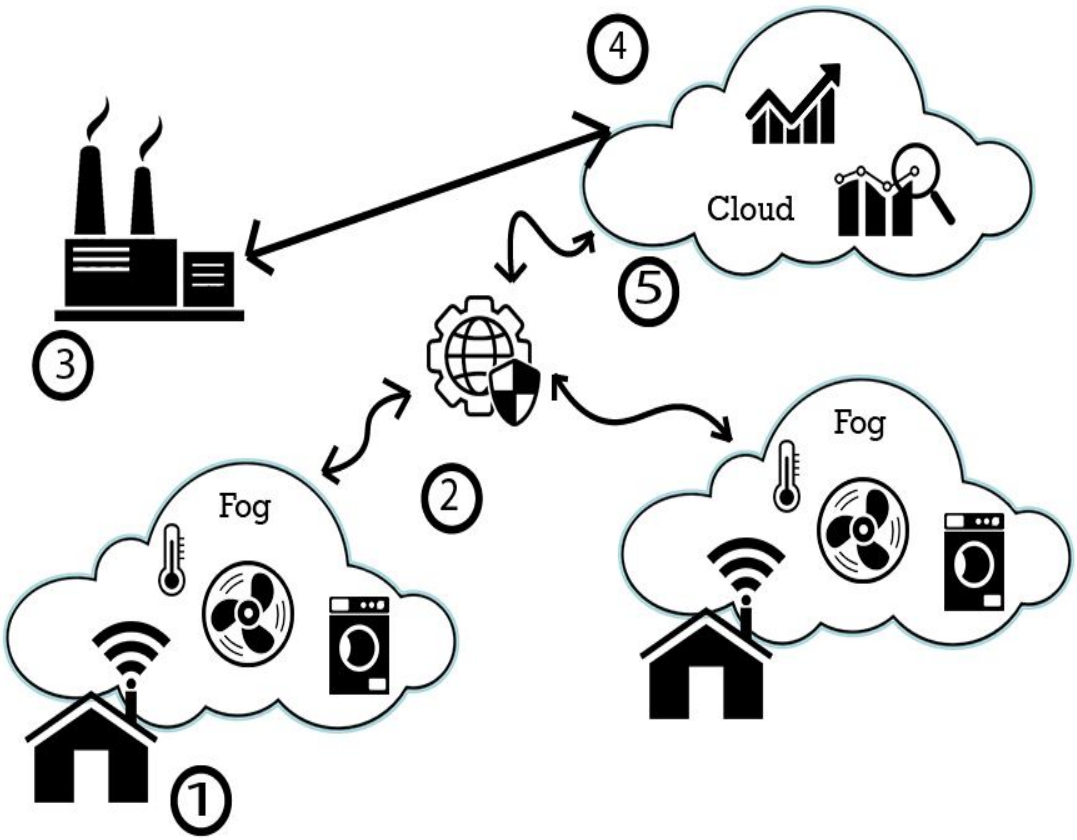


Fig.3.3 smart home using fog computing

3.4 Health Care monitoring system:

3.4.1 Healthcare in fog computing:

Cloud computing possess the scalability, mobility, security benefits. Cloud computing is one of the importance pillar of the IoT health care system. With cloud computing it can share data among health professionals, doctors, patient with minimum risk of loss of data. But cloud computing does not support the real time feedback like remote monitoring. Therefore, to support highly scalable computing the fog computing technology has been used. Cisco said that using fog computing any one can exceed the computing power and storage of the cloud to the network edge [12]. It will help to reduce the network latency and bandwidth usage. In an easiest way state that bring the cloud closest to the user. In 2017, a smart healthcare gateway for fog computing was introduced [13]. Fog computing is at mediator level.

 ISBN: 978-81-946375-7-8

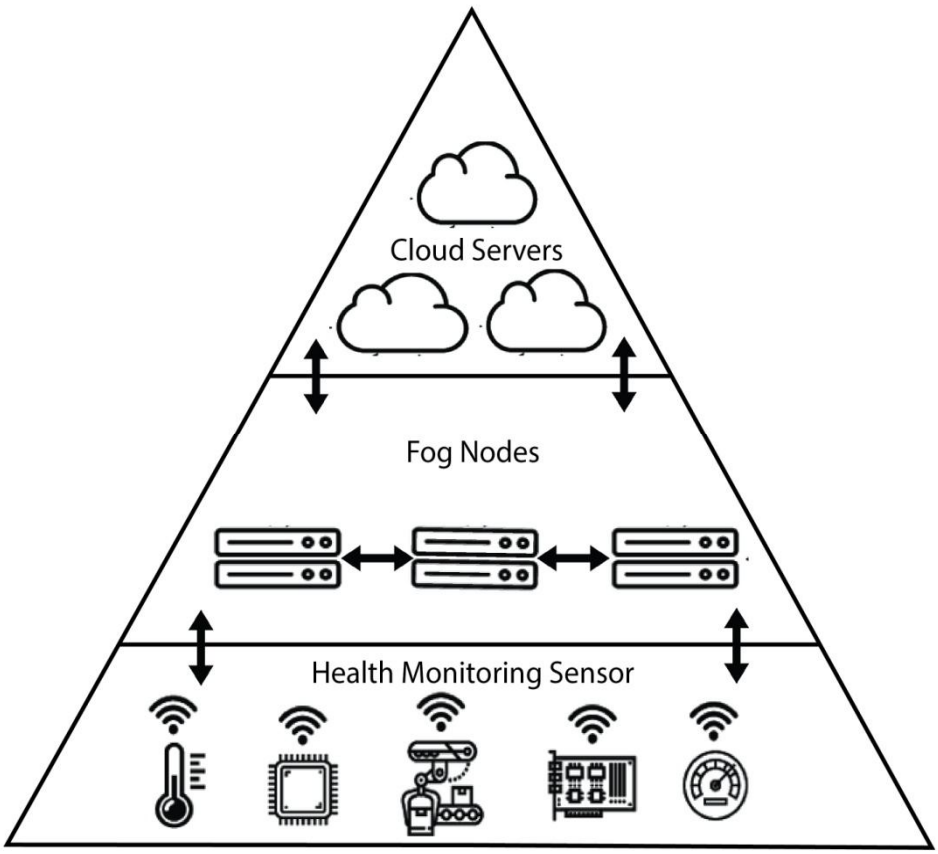


Fig.3.4 Health care monitoring system using fog computing

The raw information has been received by the sensors and that information is feed to cloud.

3.4.2 Application of healthcare system:

3.4.2.1 ECG Monitoring:

There is an activity which gives information about HR and Rhythm. It provides instant early prediction of heart performance, high BP problem or other heart diseases. ECG is a mandatory test for the patient who is having high BP diseases. To measure the ECG patient has visit the clinic and check the status of heart rate through ECG Machine. But nowadays a wearable gadget performs operations in real time. Now days have seen lots of emerging smart watches or HR monitor have features of the ECG in the gadget itself. The device has a sensors to get readings in real time and process data in fog computing and send a relevant data to cloud computing.

IV. Conclusion

In few years, the Internet of Things attracts the attentions of many peoples. It is becoming an integral part of our lives. It has the ability to connect almost everything to everything else in our day-to-day life. Internet of Things devices are dynamic in nature, which have limited storage and processing power. Traditional centralize clouds have many issues, such as high latency and network failure. To solve this issues, fog computing is developed as an extension of cloud, which bring cloud closer to the Internet of Things devices for real-time response.

In this paper shown that the integration of fog computing and Internet of Things devices bring many benefits to different Internet of Things applications.

References

1. R. Roman, J. Lopez, and M. Manbo, "Mobile Edge Computing, Fog et al.: A Survey and Analysis of Security, Threats and Challenges," Futur. Gener. Comp. Syst., vol. 78, pp. 680–698, 2018
2. https://www.netburner.com/learn/architectural-frameworks-in-the-iot-civilization/
3. https://www.researchgate.net/publication/311065854_Secure_Integration_of_Internet-of-Things_and_Cloud_Computing
4. Atlam, H.F.; Alenezi, A.; Alharthi, A.; Walters, R.; Wills, G. Integration of cloud computing with internet of things: challenges and open issues. In Proceedings of the 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Exeter, UK, 21–23 June 2017; pp. 670–675.
5. Ai, Y.; Peng, M.; Zhang, K. Edge cloud computing technologies for internet of things: A primer. Digit. Commun. Netw. 2017, in press. [CrossRef]
6. Wen, Z.; Yang, R.; Garraghan, P.; Lin, T.; Xu, J.; Rovatsos, M. Fog orchestration for internet of things services. IEEE Internet Comput. 2017, 21, 16–24. [CrossRef]
7. https://www.mdpi.com/2504-2289/2/2/10
8. https://www.researchgate.net/figure/Layered-architecture-of-fog-computing_fig2_324280213
9. https://www.moxa.com/en/articles/does-your-iot-application-need-fog-computing
10. https://www.iotforall.com/connected-car
11. https://moility.here.com/smart-traffic-systems-101- components-benefits-and-big-data-connection
12. Solutions, C.F.C. Unleash the Power of the Internet of Things; Cisco Systems, Inc.: San Jose, CA, USA, 2015.
13. Negash, B.; Gia, T.N.; Anzanpour, A.; Azimi, I.; Jiang, M.; Westerlund, T.; Rahmani, A.M.; Liljeberg, P.;Tenhunen, H. Leveraging fog computing for healthcare iot. In Fog Computing in the Internet of Things;Springer: Berlin/Heidelberg, Germany, 2018; pp. 145–169

 ISBN: 978-81-946375-7-8

Implementation of Spark-Scala Vs. Scala

Amruta Jadhav
Student, Master of Computer Application, Navinchandra Mehta Institute of Tenchnology and Development, Dadar(W)/University of Mumbai, Mumbai

Hemangi Gadade
Student, Master of Computer Application, Navinchandra Mehta Institute of Tenchnology and Development, Dadar(W)/University of Mumbai, Mumbai

Mr. Pritam Warke
Assistant Professor, MCA Department, Navinchandra Mehta Institute of Tenchnology and Development, Dadar(W)/University of Mumbai, Mumbai

ABSTRACT
Developers will get Object oriented conception terribly simply. By mistreatment, Scala language an ideal balance is maintain between productivity and performance. Syntax for Scala language is a smaller amount advanced as compare to CPP and Java thus it is smart for brand spanking new developers to start out mistreatment Scala.

Scala is such a robust language that has its own potential and this will be seen from it increasing demand within the market. Scala is combination of purposeful programming and object minded programming. Scala is such a purposeful and compiled language that is quicker than R and Python.

We see currently a day, there is an increasing demand of Scala among developers. The explanation behind this is often IT Company and very different organizations have an interest in those developers, which will masters of strong programing language, which will be utilized in information analysis, and programming in Apache Spark. Scala is one amongst the most effective possibility for this. Apache Spark is written in Scala because it is additional scalable on JVM. Scala helps to dig deep into the Spark's ASCII text file that aids developers to simply access and implement new options of Spark.

Keywords: Spark, Scala, RDD

I. INTRODUCTION

The foremost difficult thing for massive information builders nowadays is selecting a programing language for giant records applications.

Python and R programming are the languages of preference among statistics scientists for building machine-learning models at the same time as Java remains the go-to programing language for growing Hadoop programs.

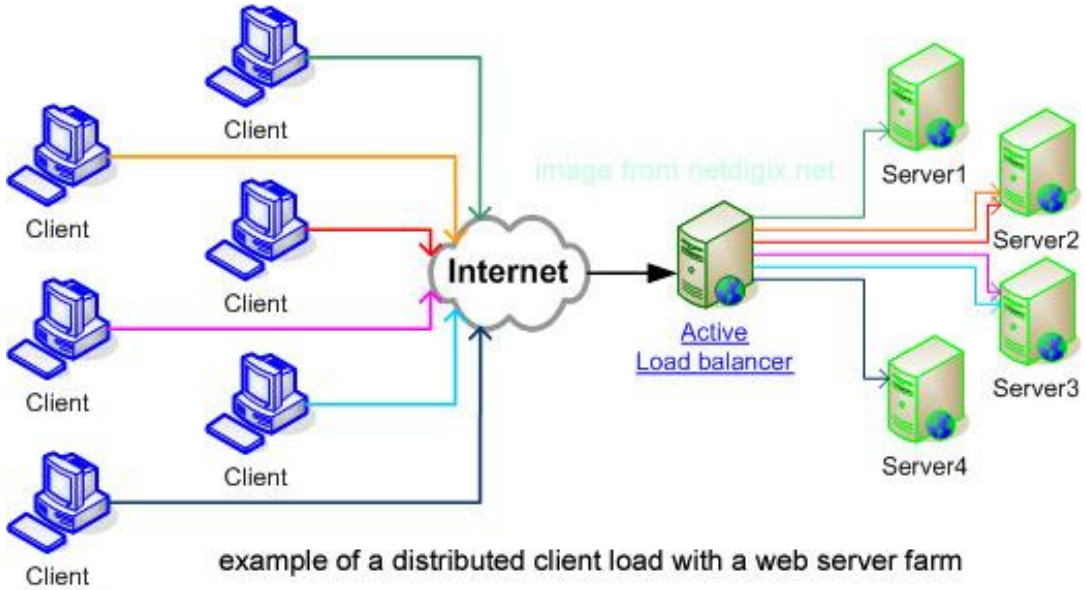
With the appearance of varied huge statistics frameworks like Apache Kafka and Apache Spark-Scala programing language has received prominence amid massive statistics developers. To make information analytics speedy Scala is used for Apache Spark Programming.

70% of spark users makes use of Scala Programming language as the pleasant programming language in line with the recent survey.

II. BACKGROUND

Computer clusters may want to also be configured for various purposes ranging from general-reason enterprise needs like web-service guide, to computation-intensive clinical calculations.

Note that the attributes described beneath aren't unique and a "compute cluster" also can use a excessive-availability method, etc.



- **LOAD BALANCING**

"Load-balancing" clusters are configurations throughout which cluster-nodes proportion computational workload to supply higher overall performance. For instance, a web server cluster may assign exceptional queries to exceptional nodes, consequently the general reaction time are going to be optimized.[4] However, approaches to load balancing may also considerably range among packages, e.g. A high- overall performance cluster used for scientific computations would balance load with special algorithms from a web-server cluster, which could just use an clean round-robin method by assigning each new request to a unique node. "Computer clusters" are used for computation-extensive functions, in place of coping with IO-orientated operations like web provider or databases. [5]

- **CLUSTER MANAGEMENT**

One of the challenges within the use of a laptop cluster is that the cost of administrating it, which may additionally once in a while be as excessive because the cost of administrating N independent machines, if the cluster has N nodes. In a few cases, this affords a plus to shared reminiscence architectures with decrease administration costs. This has also made digital machines popular, thanks to the benefit of administration.

- **TASK SCHEDULING**

When an oversized multi-consumer cluster need to get entry to very huge amounts of understanding, assignment scheduling turns into a challenge. The MapReduce approach was cautioned through Google in 2004 and other algorithms like Hadoop are applied. However, as long as at some stage in a complicated utility environment the overall performance of every job relies upon on the traits of the unique cluster, mapping duties onto CPU cores and GPU gadgets are very huge challenges. This is usually a neighborhood of ongoing studies and algorithms that integrate and enlarge MapReduce and Hadoop are proposed and studied.

- **NODE FAILURE MANAGEMENT**

When a node in the course of a cluster fails, techniques like "fencing" can also be employed to live the rest of the system operational. Fencing is that the system of keeping apart a node or

shielding shared sources whilst a node appears to be malfunctioning. The two training of fencing strategies are; one disables a node itself, and therefore the opposite disallows access to sources like shared disks. The STONITH technique stands for "Shoot the alternative Node in the Head", it method that the suspected node is disabled. As an example, strength fencing uses a power controller to reveal off an inoperable node. [7]

1. Spark

Apache Spark started as a search project at UC Berkeley within the AMP Lab, which focuses on big data analytics.

Our goal was to style a programming model that supports a way wider class of applications than MapReduce, while maintaining its automatic fault tolerance. Especially, MapReduce is inefficient for multi-pass applications that need low-latency data sharing across multiple parallel operations.

Traditional MapReduce and DAG engines are suboptimal for these applications because they're supported acyclic data flow: an application has got to run as a series of distinct jobs, data gets read from stable storage (e.g. distributed file system) and writes it back to stable storage. They incur significant cost loading the info on each step and writing it back to replicated storage.

They incur substantial price loading the information on each step and writing it again to replicate garage. Spark affords an abstraction referred to as resilient disbursed datasets (RDDs) to guide these programs very efficiently. RDDs are regularly saved in memory without requiring replication between queries. Instead, the lost facts are rebuilding on failure the use of lineage: each RDD remembers how it has been built from different datasets (via adjustments like map, join or group through) to simply rebuild itself. RDDs permit Spark to outperform models which can be present through up to 100x in multi-skip analytics. We showed that RDDs could help a good form of iterative algorithms, also as interactive data processing and an especially green SQL engine (Shark). [1]

- **Spark cluster computing architecture**

Although Spark has similarities to Hadoop, it represents a substitute cluster-computing framework with useful variations. First, Spark become designed for a particular kind of workload in cluster computing namely, those who reuse a running set of information throughout parallel operations (such as gadget gaining knowledge of algorithms). To optimize for those sorts of workloads, the idea of in-memory cluster computing is delivered by using Spark, where datasets are often cached in memory to cut back their latency of get right of entry to. Spark additionally introduces critical idea of an abstraction referred to as resilient dispensed datasets (RDDs). An RDD may be a examine-only collection of items allotted across a set of nodes. These collections are resilient, because they will be rebuilt if a number of the dataset is lost. The method of rebuilding some of the information set is based on a fault-tolerance mechanism that keeps lineage (or facts that permits the portion of the dataset to be re-created supported the approach from which the information became derived).

An RDD is represented as a Scala item and can be made out of a record; as a parallelized slice (spread throughout nodes); as an alternate of every other RDD; and eventually via changing the endurance of an current RDD, like requesting that it's cached in memory. Applications in Spark are referred to as drivers, and those drivers enforce the operations carried out either on one node or in parallel across a set of nodes. Spark helps a single-node cluster or a multi-node cluster same as Hadoop.

Spark is predicated at the Mesos cluster manager for other multi-node operation, An green platform for resource sharing and isolation for allotted programs are provided via Mesos (see Figure). This setup lets in Spark to coexist with Hadoop at some stage in a single shared pool of nodes.

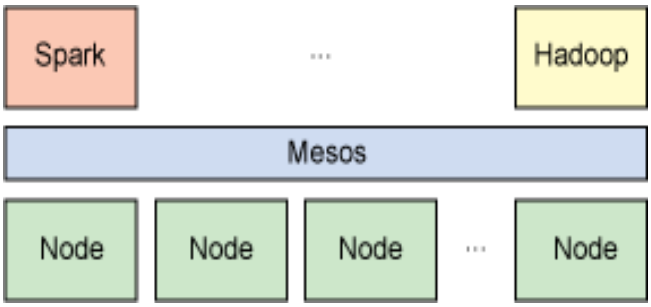


Figure. Spark relies on the Mesos cluster manager for resource sharing and isolation.

Spark is predicated on the Mesos cluster manager for resource sharing and isolation. Although Hadoop captures the most interest for allotted facts analytics, there are many options which give a few interesting blessings to the acknowledged standard Hadoop platform. Spark may be a scalable information analytics platform that includes primitives for in-memory computing and for this reason physical games a few performance blessings on all over Hadoop's cluster storage method. Spark is carried out in and exploits the Scala language, which provides novel surroundings for processing. Get to understand the Spark method for cluster computing and its differences from Hadoop. [2]

- **Features of Spark**

❖ Swift Processing

Using Apache Spark, we achieve a excessive processing velocity of about 100x quicker in memory and 10x faster on the disk. That is frequently made possible by means of reducing the quantity of examine-write to disk.

❖ Dynamic in Nature

Spark gives 80 excessive-degree operators as we can without problems develop parallel software.

❖ In-Memory Computation in Spark

With in-reminiscence processing, we will boom the processing speed. Here the information is being cached so we'd like no longer fetch records from the disk every time as a result the time is saved. Spark has DAG execution engine, which facilitates in-memory computation and acyclic information flow leading to excessive pace.

❖ Reusability

The opportunity to reuse the Spark code for batch-processing, be a part of circulate against historical statistics or run ad-hoc queries on flow state.

❖ Fault Tolerance in Spark

Apache Spark offers fault tolerance via Spark abstraction-RDD. Spark RDDs are designed to deal with the failure of any worker node in the cluster. Thus, it guarantees that the loss of knowledge reduces to zero. Learn alternative methods to make RDD in Apache Spark.

❖ Real-Time Stream Processing

Spark functions a provision for real-time flow processing. Earlier the problem with Hadoop MapReduce changed into that it is able to manage and manner information, which is already

present, but now not the real-time statistics. But with Spark Streaming we are able to remedy this problem.

❖ Lazy Evaluation in Apache Spark

All the differences we make in Spark RDD are Lazy in nature, that is it does now not supply the result proper away as a substitute a brand new RDD is formed from the prevailing one. Thus, this will increase the efficiency of the machine. Follow this manual to learn greater approximately Spark Lazy Evaluation in excellent detail.

❖ Support Multiple Languages

In Spark, there's Support for more than one languages like Java, R, Scala, Python. Thus, it affords dynamicity and overcomes the limitation of Hadoop that it can build programs handiest in Java.

❖ Active, Progressive and Expanding Spark Community

Developers from over 50 businesses were involved in making of Apache Spark. This undertaking was initiated inside the year 2009 and is still expanding and now there are approximately 250 builders who contributed to its expansion. It is the most critical task of Apache Community.

❖ Support for Sophisticated Analysis

Spark comes with devoted gear for streaming information, interactive / declarative queries, system studying which add-on to map and reduce.

❖ Integrated with Hadoop

Spark can run independently and also on Hadoop YARN Cluster Manager and as a consequence it could read current Hadoop facts. Thus, Spark is flexible.

❖ Spark GraphX

Spark has GraphX, which may be a component for graph and graph-parallel computation. It simplifies the graph analytics obligations through the collection of graph set of rules and builders.

❖ Cost Efficient

Apache Spark is price powerful answer for massive records hassle as in Hadoop great deal of garage and therefore the massive statistics centre is required throughout replication.[6]

2. Scala

Scala (Scalable Language) is an open language created by Professor Martin Oder sky, the founding father of Type safe, which promotes and provides commercial support for Scala programing language. It's a multi-paradigm programing language and supports functional also as object oriented paradigms. From the functional programming perspective, each function in Scala may be a value and from the thing oriented aspect - each value in Scala is an object.

Scala may be a JVM based statistically typed language that's safe and expressive. With its extensions which will be easily integrated into the language-Scala is taken into account because the language of option to achieve extensibility.

Scala programing languages are regularly determined in use at some of the simplest tech businesses like LinkedIn, Twitter, and Four-square. Scala's overall performance has ignited interest amongst several monetary institutions to apply it for derivative pricing in EDF Trading. The most crucial names in the digital financial system are making an investment in Scala programming for large processing are - Kafka created through LinkedIn and Scalding created

through Twitter. With effective monoids, combinatory, pattern-matching capabilities, provision to make DSL's and extra -Scala as a device for giant processing on Apache Spark is truly a certainty.[3]

For example, at Twitter, the social networking provider, Robey Pointer moved their middle message queue from Ruby to Scala. this alteration was driven by the companies were given to reliably scale their operation to satisfy fast growing Tweet rates, already accomplishing 5000 per minute in the course of the Obama Inauguration. Robeys thinking at the back of the Twitter Kestrel project is explained in the developers stay journal. His concise 1500 lines of Scala code are often seen as he has generously made them to be had as an open source project.[7]

- **Features of Scala**

❖ **Scala is object-oriented**

Scala can be a pure item-orientated language in the sense that every fee is an object. Types and conduct of gadgets are defined with the aid of training and traits. Classes are prolonged by using sub classing and a versatile blend in-primarily based composition mechanism as a easy replacement for a couple of inheritance.

❖ **Scala is functional**

Scala may be additionally a functional language inside the experience that every feature is a fee. Scala offers a lightweight syntax for outlining anonymous capabilities, it supports higher-order capabilities, it lets in features to be nested, and supports currying. Scala's case lessons and its built-in assist for sample matching model algebraic types utilized in many purposeful programming languages.

Furthermore, Scala's belief of pattern matching naturally extends to the processing of XML records with the assistance of proper- ignoring series patterns. Throughout this context, collection comprehensions are useful for formulating queries. These capabilities make Scala ideal for developing programs like web services.

❖ **Scala is statically typed**

Scala is provided with an expressive type system that enforces statically that abstractions are utilized in a stable and coherent manner.

Especially, the sort system helps:

- generic training,
- variance annotations,
- higher and lower type bounds,
- internal classes and abstract kinds as object members,
- compound kinds,
- explicitly typed self-references,
- views,
- polymorphic strategies

A local type inference mechanism takes care that the consumer isn't required to annotate this system with redundant type information. Together, these functions provide a robust basis for the safe reuse of programming abstractions and for the type-safe extension of software

❖ **Scala is extensible**

In practice, the event of domain-specific packages often calls for domain-specific language extensions. Scala presents a novel aggregate of language mechanisms that make it smooth to smoothly upload new language assemble in type of libraries:

- any method can also be used as an infix or postfix operator, and
- Closures are built automatically counting on the predicted type (target typing)

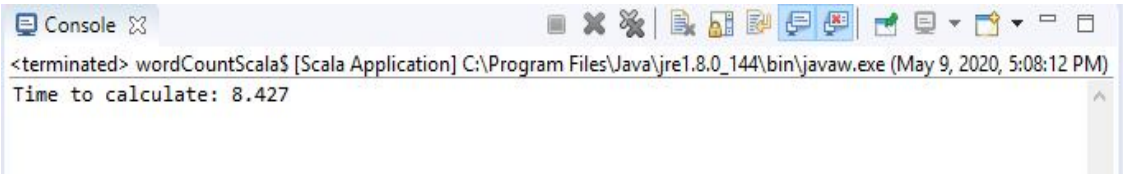
A joint use of both features enables the definition of latest statements without extending the syntax and without the usage of macro-like meta-programming facilities.[7]

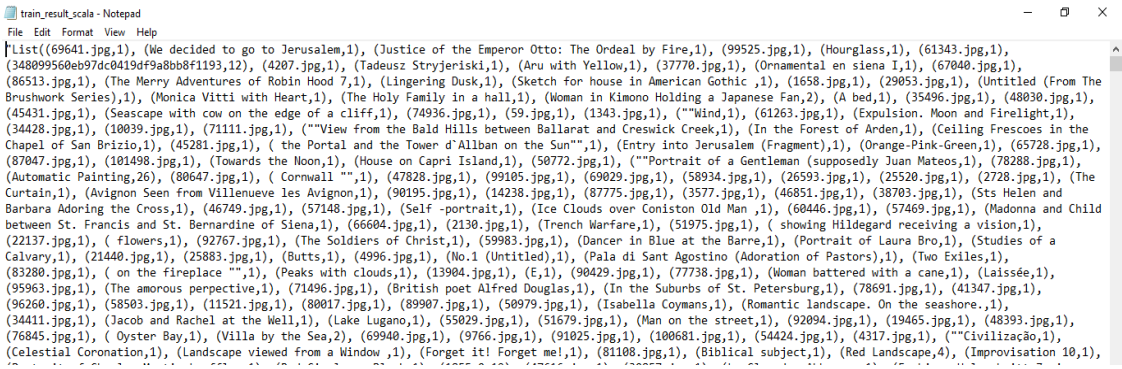
III.RESEARCH DESIGN/WORK:

Scala Code:

```
package com
import scala.io.Source
import collection.JavaConverters._
import collection.mutable._
import java.io.{BufferedWriter, FileWriter}
import scala.collection.JavaConversions._
import scala.collection.mutable.ListBuffer
import scala.util.Random
import au.com.bytecode.opencsv.CSVWriter
object wordCountScala {
def main(args: Array[String]) = {
val startTime = System.currentTimeMillis()
val filedata = Source.fromFile("C:\\manali\\train_info.csv").getLines().toList
val words = filedata.flatMap(line => line.split(","))
val keyData = words.map(word  =>  (word,1))
val groupeData = keyData.groupBy(_._1)
val result = groupeData.mapValues(list =>{list.map(_._2).sum}).toList.toString()
val outputFile = new BufferedWriter(new FileWriter("C:\\manali\\research_paper\\train_result_scala"))
val csvWriter = new CSVWriter(outputFile)
csvWriter.writeNext(result)
val endTime = System.currentTimeMillis()
println("Time to calculate: " + (endTime - startTime)/1000.0)
outputFile.close()
}
}
```
[8]

Output:



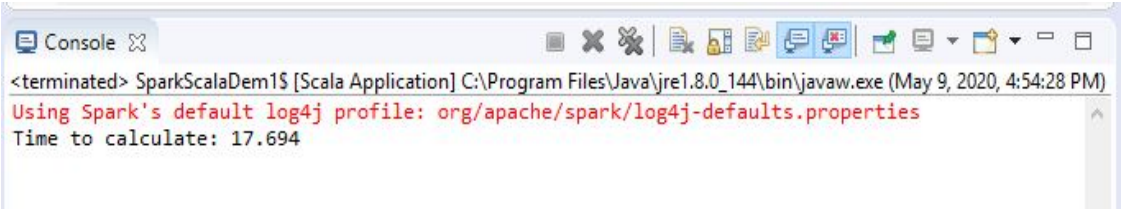


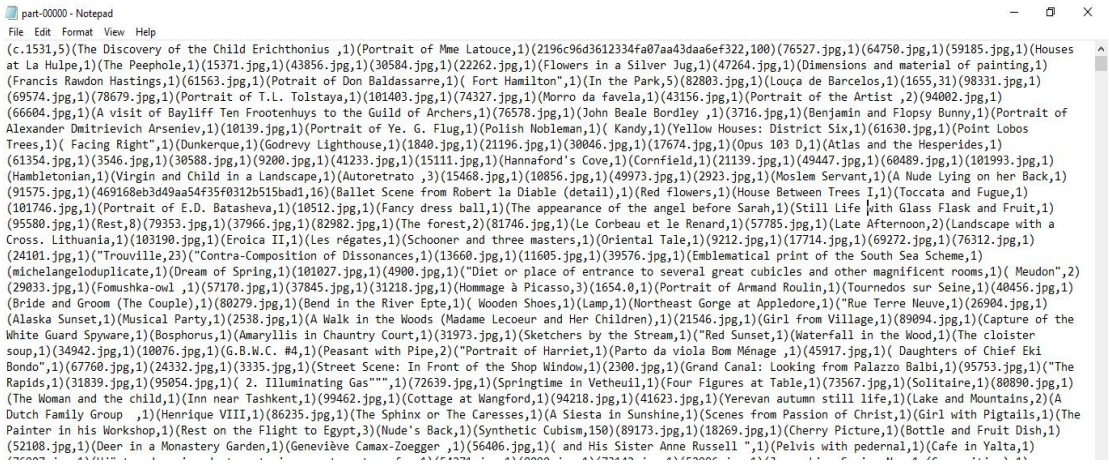
IV. Spark-Scala Code:

```
package com
import org.apache.spark.sql.SparkSession
import org.apache.log4j._
object SparkScalaDem1 {
def main(args: Array[String]): Unit = {
Logger.getLogger("org").setLevel(Level.ERROR)
val spark = SparkSession.builder()
.master("local")
.appName("SparkOperations")
.getOrCreate()
val startTime = System.currentTimeMillis()
val csvData = spark.sparkContext.textFile("C:\\manali\\train_info.csv")
val wordcnt = csvData.flatMap(line => line.split(",")).map(word => (word,1)).reduceByKey(_+_)
wordcnt.saveAsTextFile("C:\\manali\\research_paper\\train_result")
val endTime = System.currentTimeMillis()
println("Time to calculate: " + (endTime - startTime)/1000.0)
}
}
```
[7]

OutPut

 ISBN: 978-81-946375-7-8





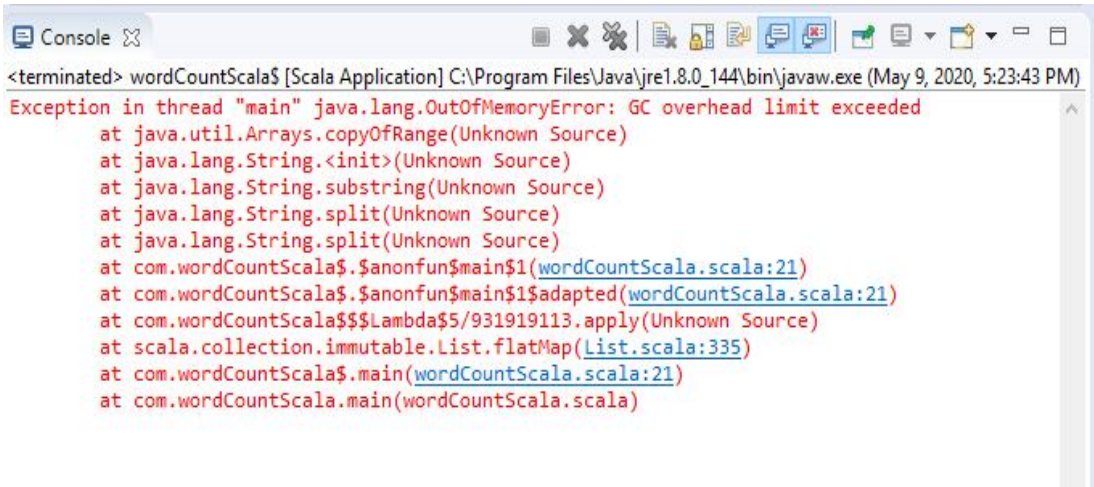
V. IMPLEMENTATION

Scala code

```
package com
import scala.io.Source
import collection.JavaConverters._
import collection.mutable._
import java.io.{BufferedWriter, FileWriter}
import scala.collection.JavaConversions._
import scala.collection.mutable.ListBuffer
import scala.util.Random
import au.com.bytecode.opencsv.CSVWriter
object wordCountScala {
def main(args: Array[String]) = {
val startTime = System.currentTimeMillis()
val filedata = Source.fromFile("C:\\manali\\all_data_info.csv").getLines().toList
val words = filedata.flatMap(line => line.split(","))
val keyData = words.map(word  =>  (word,1))
val groupeData = keyData.groupBy(_._1)
val result = groupeData.mapValues(list =>{list.map(_._2).sum}).toList.toString()
```

```
val outputFile = new BufferedWriter(new FileWriter("C:\\manali\\research_paper\\train_result_scala"))
val csvWriter = new CSVWriter(outputFile)
csvWriter.writeNext(result)
val endTime = System.currentTimeMillis()
println("Time to calculate: " + (endTime - startTime)/1000.0)
outputFile.close()
}
}
```
[8]

Output

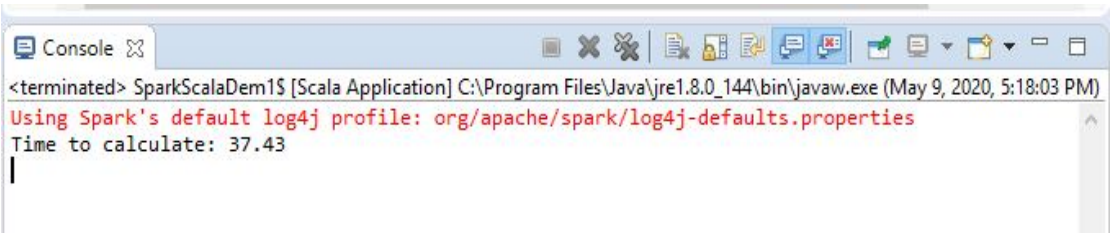


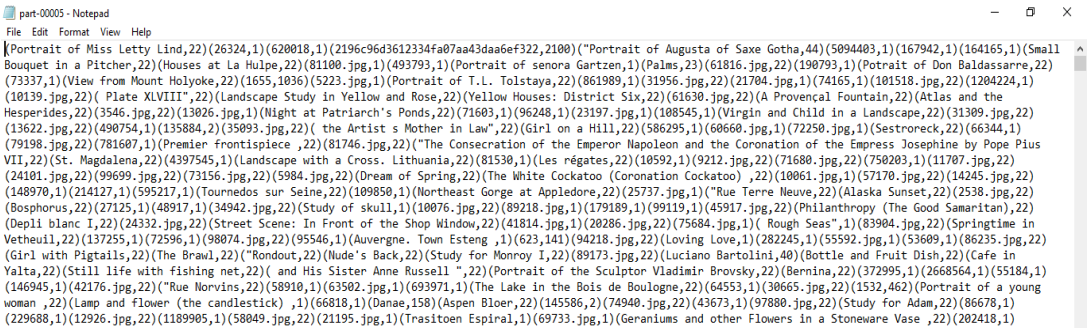
Spark-Scala code

```
package com
import org.apache.spark.sql.SparkSession
import org.apache.log4j._
object SparkScalaDem1 {
def main(args: Array[String]): Unit = {
Logger.getLogger("org").setLevel(Level.ERROR)
val spark = SparkSession.builder()
.master("local")
.appName("SparkOperations")
.getOrCreate()
val startTime = System.currentTimeMillis()
val csvData = spark.sparkContext.textFile("C:\\manali\\all_data_info.csv")
```

```
val wordcnt = csvData.flatMap(line => line.split(",")).map(word => (word,1)).reduceByKey(_+_)
wordcnt.saveAsTextFile("C:\\manali\\research_paper\\train_result")
val endTime = System.currentTimeMillis()
println("Time to calculate: " + (endTime - startTime)/1000.0)
}
}
```
[7]

OutPut





VI. CONCLUSION

Spark is Open Source distributed computing platform.

For large data scale processing speed always matter. So Spark is faster than Scala for large data processing.

Using single command, multiple operations can be perform.

Spark not only support Map & Reduce. Also support Machine learning, graph algorithms, SQL queries, etc.

Spark is Multilingual means it supports many programming languages like python, scala, java, etc.

In Spark, there is issue with small size files.

VII. REFERNCES

1. https://spark.apache.org/research.html
2. www.ibm.com/developerworks/library/os-spark/
3. https://www.dezyre.com/article/why-learn-scala-programming-for-apache-spark/198
4. https://people.eecs.berkeley.edu/~franklin/Papers/hotcloud.pdf
5. https://data-flair.training/blogs/apache-spark-features/

 ISBN: 978-81-946375-7-8

6. http://www.ijsrp.org/research-paper-1112/ijsrp-p1181.pdf
7. https://spark.apache.org/examples.html
8. https://scala-lang.org/files/archive/spec/2.13/

 ISBN: 978-81-946375-7-8

Multi-factor Authentication for Privileged and Non-Privileged Users

Anuja Patil
MCA, DES'S NMITD, University of Mumbai

Abstract
Today, passwords are the most widespread form of user authentication. In recent years, there has been a disturbing trend in the number and types of cyber breaches worldwide. These violations showed that everyone is vulnerable, including the most sophisticated Information Technology companies, the largest and most looked up to financial institutions. Due to the use of only one factor authentication, that is the use of only user id and password, privileged user accounts may be used to compromise the system confidentiality, integrity and availability and act as a starting point for attacking other critical resources in the environment. To overcome this cyber threat exploit, organizations are turning to Multifactor Authentication. Many of these violations involve the use of compromised privileged credentials. Only a small number of cyber security professionals believe that user name and password-based security remains an adequate form of protection. This is one of the main reasons many organizations are turning to multi-factor authentication (MFA) or two-factor authentication to not only reduce the risk of stolen passwords but also to verify the user, which is one of the main concepts. By adding a second

authentication factor to security policies, attackers are unable to access critical n network systems and devices or gain privileged access without a smartphone (e.g., something you have) or the fingerprint (e.g., something you are) required to complete the process of authentication. This work sheds light on the cyber threat exploits in terms of privileged users and how the use of different types of multi-factor authentication will help overcome these breaches and shortcomings.

Keywords Multifactor Authentication, MFA, Two-factor Authentication, 2FA, Privileged Access

1. Introduction

Two-factor Authentication (2FA) is one member of a group of closely related technologies which seek to strengthen user authentication by requiring a method of verifying the user's identity in addition to their password. Typically, the additional factor will either be something the user has in his or her ownership (e.g. a one-time use code provided via security token, telephone, text, or email), or a biometric feature (e.g. facial recognition, fingerprint or voice recognition). 2FA is a form of multifactor authentication (MFA), while two-step verification (2SV) does not technically require a second "factor", but requires a passcode to be sent via a separate communication strip. The study described in this document encompasses many online services offering variants of these technologies, some offering several**.** For example, Facebook supports the use of an SMS code, a software token or a hardware token. Since this document covers a variety of suppliers and implementations, the term 2FA will generally be used to refer to such technology, unless it refers to the implementation of a specific supplier by name (for example, the two-step verification from Google). Historically, 2FA was primarily accessible only to government and other large companies. Recently, however, it has become widely available to home end users, often at no additional monetary cost.

Table 1 shows a variety of popular websites which offer users Multifactor authentication, covering key categories of online activity such as social media; e-commerce and banking; and email, mobile and cloud services

Examples of sites offering Multifactor Authentication		
Social Media	**Email/Mobile/Cloud**	**E-Commerce/Banking**
Facebook	Gmail/Google Drive	Citi
Twitter	Apple(iCloud)	eBay
Instagram	Microsoft(Hotmail, OneDrive)	
LinkedIn	Dropbox	
Google Plus	Amazon Drive	

Table 1:shows a variety of popular websites which offers users Multifactor Authentication

Multi-factor authentication is an essential security tool required for today's increasingly complex, hybrid IT environment, especially with new attack surfaces that include the cloud. It adds an additional layer of security around your confidential data, validating the user before granting access during login or privilege elevation, which is an essential concept in achieving least privilege and a Zero Standing Privileges posture — a best practice prescribed by Gartner.

Passwords are the most common form of user authentication on the web today [1]. Although many password replacements schemes have been proposed, none of them compete with the deploy-ability and usability of passwords [2]. Recently, huge service providers, including, Facebook, Google have integrated an optional 2FA layer as part of their authentication processes to defend against account compromise. Two-factor authentication requires users to present two of the following types of authentication factors: 1. Something they know (traditionally a password) 2. Something they have (like a phone or a hardware token) 3. Something they are (referring to biometrics, such as a fingerprint) Several 2FA methods are used, methods such as SMS, TOTP (time-based one-time password) and hardware code generators (such as RSA SecureID) requires the user to enter a one-time passcode in addition to their password. These passcodes are either sent to the user via a separate channel, or generated on the fly by the user's device. In commercial and government environments, smart cards are a second commonly used feature, requiring the user to insert an identification badge in a card reader connected to his computer. Banking systems, frequently use variants of hardware code generators and card readers in their 2FA implementations. Bigwigs such as Google, Dropbox and GitHub have deployed hardware USB tokens (aka security keys), such as YubiKey, internally [3]. Two-factor authentication provides a solid defence against account compromise. The number of recent password database leaks [4] highlights the risk of account compromise. Since users tend to re-use the same username and password on multiple sites [5], password leaks from one site can result in a chain reaction of account vulnerabilities when attackers access other accounts with the same credentials [6]. Even if an attacker steals or guesses a user's password, the attacker must compromise the user's phone or steal a physical token to gain access to the account. Thus, it is much more difficult for a remote attacker to compromise an account protected by a second authentication factor.

2. Literature Review

Asif and Amin [7] have proposed to generate OTPs in such a way, in order to secure the system, the generated OTP must be hard to guess, retrieve, or trace by hackers. Therefore, it is very important to develop a secure OTP generation algorithm. Several factors can be used by the OTP algorithm to generate a password that is difficult to guess. Users seem to be willing to use simple factors such as their mobile number and a PIN for services such as authorizing mobile micro payments, so a Secured Cryptographic algorithm is thus proposed.

The username, password, date of birth of user is taken from the user and then concatenated with the current date, time and the time stamp for which the one-time password is valid. This concatenated string is then given as input to Secured Hash Algorithm (SHA1) Algorithm. SHA-1 algorithm returns its message digest which is 20 bytes value. These 20 bytes are reduced to 5 bytes by XORing a group of 4 bytes , i.e. byte no. 1, 4, 8, 12 are XORed ; 2, 5, 9, 13; 3, 6, 10, 14; 4, 7, 11, 15; 5, 8, 12, 16; 17, 18, 19, 20 are XORed. Then from this 5-byte value, every byte is right shifted with 4 digits and then is converted to hexadecimal. Finally, by converting the ASCII values to a character string, it is displayed as a onetime password to the user.

Jessica Colnago, Maggie Oates [8] of University of California studied people who adopted 2FA at CMU found it annoying, but fairly easy to use, and believed it made their accounts more secure. The likelihood that a user would subsequently adopt 2FA for other accounts was related to their opinions about 2FA ease-of-use and perceived value. While they found some evidence that people who were required to adopt 2FA had more negative perceptions than those who adopted voluntarily, the differences were smaller than expected. They also explored the relationship between usage patterns and perceived usability

3. Research Design

Here we are proposing to deploy MFA everywhere. Applying multifactor-authentication only for certain systems, administrators, and privileged applications only exposes organizations to attacks and exploits. Take advantage of enhanced (contextual) or adaptive (risk-based) authentication with MFA that balances user security and convenience. Offer a choice of authentication methods for maximum flexibility and a better user experience. A variety of authentication methods help IT meet the needs of different user populations. Go for standards-based MFA solutions because standards allow MFA to work well with your existing IT environment and help prevent vendor lockout. Integrate MFA with least privilege access to further strengthen protection against compromised passwords. · Continually reassess the 2FA to determine if the deployment is still meeting the evolving needs of the organization. Apply role-based access to cloud environments with policy-driven elevation of privilege combined with session auditing and monitoring

A key benefit of these approaches is the improvement of the user experience. Rather than being constantly asked for the 2FA, the user is invited to provide an additional factor only when necessary. For example, a user connecting from the corporate network to a managed device can be granted access with only one factor, a password. But, a user connecting from an unknown device on an unmanaged network may be asked to provide additional authentication.

Today, organizations have a wide range of authentication methods, including:

Hardware tokens: small hardware devices that a user carries to authorize access. They come in different forms, including smart cards, key chains or USB devices. The hardware device generates a one-time password (OTP) that the user enters when prompted.

Single factor cryptographic devices: FIDO U2F and its successor without password, FIDO2, are authentication standards managed by the FIDO Alliance. They are designed to be open, secure, private and easy to use. Considered as next-generation two-factor authentication, the benefits of FIDO U2F and FIDO2 include enhanced security because public key cryptography protects against phishing, session hijacking, and malware attacks, as well as the ease of use and high confidentiality.

Software tokens: these are software tokens or applications that generate a one-time password (OTP). These are usually mobile apps installed on a smartphone and can take advantage of push notifications to improve user convenience. The widespread adoption of mobile devices has

made software tokens a popular option. Software tokens have two main advantages over hardware tokens. First, users are less likely to lose or forget their phones than a single-use hardware token. Second, software tokens are easier and cheaper to distribute to users.

SMS/Text message: An **OTP** can be sent to a cell phone via SMS. Once received, the user enters it in the login screen.

Telephone call: with this authentication method, a user receives a telephone call to a registered telephone number (landline or mobile number). The user then provides the correct response to the voice prompt to complete the authentication.

Email: a user receives an email with a link to verify the request. Clicking on the link settles the authentication process.

Security questions: Instead of tokens, users give answers to security questions. These questions can be predefined or the user can define his own questions.

Biometric: These methods include fingerprints, retina scanners, facial recognition, etc. Many of the smartphones today support biometrics such as Touch ID on iPhones and Fingerprint for Samsung Galaxy devices. The FIDO2 standard includes the ability to use biometric data on the device such as Microsoft Hello or Apple FaceID.

A wide range of authentication methods gives users flexibility and choice improving user experience

In case of privileged accounts, the privileged account must be asked for a second factor of authentication by default when the privileged user enters his credentials, he must be receiving an OTP or if he has a software token, he must enter it.

The already existing applications like Google Authenticator and Authy can be integrated with such accounts

Authentication verification process by studying patterns of the user if the user behaviour does not match the pattern it should ask for elevated level of authentication

In case if the pattern does not match of privileged user, the user can be verified by sending a link to his email account, by clicking on that link the super users will be notified and only after they verify the user, the user will be able to access his account.

4. Implementation

The proposed model should by default ask for the second authentication factor when a privileged user logs in to the account, choice of authentication should be offered to maintain flexibility.

Also, in case of non privileged users, it should offer adaptive approach and should study the user behaviour pattern if the pattern does not match with the behaviour it should ask for the second authentication factor.

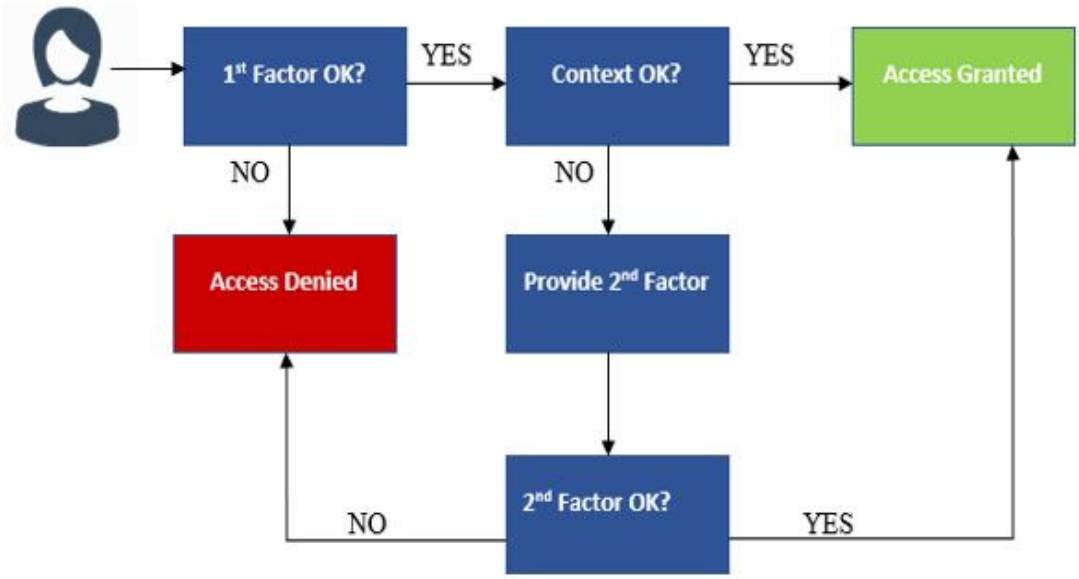


Figure 1: Represents the flow proposed

5. Conclusion

The purpose of this research paper is studying the user pattern be it a privileged user or a non-privileged user and accordingly being asked for the second authentication factor. Studying this pattern will help organization to easily recognize the privileged as well as non privileged and organization can then take necessary actions against those users who frequently deviate or show risky pattern.

References

1. A Usability Study of Five Two-Factor Authentication Methods Ken Reese, Trevor Smith, Jonathan Dutson, Jonathan Armknecht, Jacob Cameron, Kent Seamons Brigham Young University
2. Joseph Bonneau, Cormac Herley, Paul C Van Oorschot, and Frank Stajano. The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes. In 2012 IEEE Symposium on Security and Privacy (SP), pages 553–567. IEEE, 2012.
3. Juan Lang, Alexei Czeskis, Dirk Balfanz, Marius Schilder, and Sampath Srinivas. Security Keys: Practical Cryptographic Second Factors for the Modern Web. In International Conference on Financial Cryptography and Data Security (FC), pages 422–440. Springer, 2016.
4. Data Breach Investigations Report, 2017.
5. Anupam Das, Joseph Bonneau, Matthew Caesar, Nikita Borisov, and XiaoFeng Wang. The Tangled Web of Password Reuse. In Network and Distributed System Security (NDSS), volume 14, pages 23–26, 2014.
6. Blake Ives, Kenneth R Walsh, and Helmut Schneider. The Domino Effect of Password Reuse. Communications of the ACM, 47(4):75–78, 2004.
7. IJCSMC, Vol. 6, Issue. 7, July 2017, pg.5 – 8 TWO FACTOR AUTHENTICATION

8. “It’s not actually that horrible”: Exploring Adoption of Two-Factor Authentication at a University Jessica Colnago, Summer Devlin*, Maggie Oates, Chelse Swoopes, Lujo Bauer, Lorrie Cranor, Nicolas Christin Carnegie Mellon University, *University of California, Berkeley

9. Joseph Bonneau and Sören Preibusch. The Password Thicket: Technical and Market Failures in Human Authentication on the Web. In The Ninth Workshop on the Economics of Information Security (WEIS), 2010.

 ISBN: 978-81-946375-7-8

Review on Securing Web Services using CAPTCHAs

Charandeep Singh Arora and Paras Doshi
Master of Computer Application, Navinchandra Mehta Institute of Technology and Development, Mumbai University

Abstract
CAPTHCAs are the nets first line of arguments by person for whom law process is against automated account work of art and support wrongly use. Google's reCaptcha, one of the most having general approval captcha systems, is currently used by hundreds of thousands of places in the world of open internet to keep safe against made automatic attackers by testing whether a user is truly human. The end, purpose of their new system is 2-directions; to make seem unimportant the hard work for within the law users, while having need of tasks that are harder to knowledge processing machines. reCAPTCHA is driven by an increased danger observations system that revaluates requests and selects the trouble of the captcha that will be came back. users may be needed to push key to in a checkbox or get answer to a questioning by making out images or sound with like what is in.

Keywords: CAPTHCA (Completely Automated Public Turing Test to Tell Computers and Humans Apart), reCaptcha, GUID (Global Unique Identifier)

1. Introduction

The internet contributes in many aspects to do with man living such as making connections, education, and online trading, business like activities and so on. Some internet services have online registration and use services such as Swiggy, Zomato, Ticket booking etc. However, many programs have been have undergone growth by low computer experts which automatically completes the registration with fake credentials given which can cause trade goods against the law congestion, making seem unimportant the operation of the system and in some examples, even causing it to fail, particularly where a place in the net has a very great number of data ..1997 Andrei Broder et al. developed a mechanism to distinguish between human users and computer programs and also in the same year, the Altavista website used this method to block bot programs from entering by displaying a distorted English word to the user and asking the user to copy it [1]. 1996 Moni Noar suggested the use of an Automated Turing Test to distinguish between human users and bots. [2].

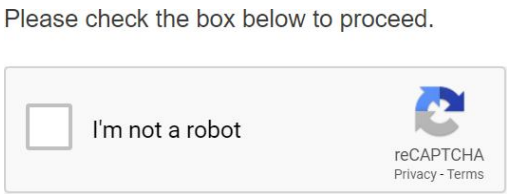


Fig 1. Captcha Verification Question. [8]

In recent years, many types of captchas have been got greater, stronger, more complete. Some are based on Optical Character Recognition (OCR) such as text captcha. In view of the fact that others are based on Non-Optical Character Recognition (Non-OCR) which uses distorted sound, such as voice and video. However, the user who are visually impaired are incapable of these seeing captchas pointing the work of art of sound CAPTCHAs of a certain sort sound

CAPTCHAs form of different speakers saying words or any numbers 0 to 9 at as if by chance spaced intervals at a random fixed in level key or go quickly, often with a way of saying words and distortion/noise. To get answer to the captcha a user must correctly make out the any numbers 0 to 9 or words talked in the sound thing to grip. Attacks have been put examples on view of on these sound captchas with changing degrees of a good outcome in the past. This is usually done by training nearby machine-learning models to make out the talked words, a high-resource and time-taking way in. in addition, although persons making observations have had a look for using on-line talk being seen supports, including Sphinx or Google speech being seen, these services have not been accurate 1 enough to take part in competition with off-line services or get answer to the captcha safely.

Fig 2. Captcha Verification Complete. [9]

1. Literature Survey

Thomas V.A and Kaur K offered marker light on computer viewer CAPTCHA Implementing captcha using instrument for pointing marker light on computer viewer, with the growing use of the net and its supports,a greatly sized number of organizations are making use of it to give and make attempts news given of the people using those supports. This has made the chances of attacks on such services by putting a stop to them sending number times another requests to the computers making ready these services programmatically. (Completely made automatic Public Turing test to say to knowledge processing machines and of men separate) provides a way to note as being different users into humans and bots 2, or knowledge processing machine lists of knowledge processing machine orders. This paper includes a new expert way that puts to use image from thing generally done instrument for pointing cursors and outdoes some most having general approval captcha techniques such as teaching book - based captchas and earlier image - based captchas

Gossweiler R, Kamvar M, and Baluja S proposed CAPTCHA Based on Image Orientation; Present a new CAPTCHA which is based on identifying an image's upright orientation. This task requires analysis of the often-complex contents of an image, a task which humans usually perform well and machines generally do not. Given a large repository of images, such as those from a web search result, use a suite of automated orientation detectors to prune those images that can be automatically set upright easily. Then apply a social feedback mechanism to verify that the remaining images have a human- recognizable upright orientation. This CAPTCHA lends itself to rapid implementation and has an almost limitless supply of images, conducted extensive experiments to measure the viability of this technique.

3. reCaptcha Background

The reCaptcha system is dependent on an increased risk analysis engine. As the user acts between, along with reCaptcha (clicking buttons and typing), the system comes to a decision about a level of doubt about that user. Today, many users will discover that they simply need to check the checkbox and be made certain of without needing to get answer to a captcha. This takes place when the reCaptcha is fairly with self-belief that the user is human and not a made automatic attacker (this is named as

"noCaptcha reCaptcha") If the system is uncertain if the user is a human (but is not very sure), It will give birth to a middle question to the user (an simple, not hard image or hard question or a short sound cord of numbers to copy into another form). This often takes place when a user does

not yet have a long enough history of effect on one another with Google. however, as the reCaptcha system becomes increasingly having feeling that something is wrong, it gives harder questions: 10 any numbers 0 to 9 in the sound question, or pointing the user to get answer to number times another questions. In 2011 the Google used reCAPTCHA with CAPTCHA technology to digitize the archive of New York Times and Google Books [3] By Default, 1, a user with no past history with Google services will be automatically given the hardest questioning. It is these most hard questions that noCaptcha attempts to get answer to. In the backend.

Fig 3 reCaptcha. [10]

1.1 Types of CAPTCHAs 3.1.a) CAPTCHA Based on Text.

Text based is the simplest to implement where a sequence of letters and digits present to the user with adding some modifications to the letters and digits such as sound or rotation of character in defined angle with 3D letters, these modifications added to prevent bot programs from reading the actual letters or digit. There are many methods on text-based CAPTCHAs.

EZ-Gimpy: This is simple version of Gimpy captcha used in chat rooms like Hangout and also used for authenticating user while registration.

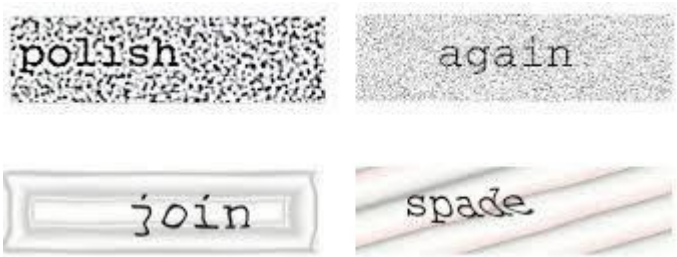


Fig 4. Example of EZ- Gimpy Captcha [11]

Scatter Type Method: The methods involve the segmentation of letters. These are modified by cutting each word into displaced pieces. The existing methods for the recognition of letters cannot be separated easily, because it is divided numerous small pieces. And letters are chosen randomly

Fig 5. Example of Scatter Type Method [12]

3.1.b) CAPTCHA Based on Images.

The captcha system presented a user with a set of images all connected with the same thing or idea of a quality common to a group of images. In recent years Google used image-based captcha like identifying objects, Streets Marks, trees etc to improve its driverless car dataset.[4][5]. The user was needed to move into the thing or idea to which all the images to e.g.

the road-map of work might present pictures of round map of earth, Volleyball, moving body moving round sun and baseball with young user to correctly get together all these pictures with the word ball. In general image-based captchas present a seeing good example or idea that the user needs to make out and act as in agreement. Different image- based captcha design uses different designs or ideas of a quality common to a group which are simple, not hard to be took in by the users and hard for the bot programs to simulate. In addition to several others, this part of captcha methods covers: Microsoft Asirra, image producing for internet checking to make certain (mind picturing), Captchas Mosaic-based to do with human effecting on one another fact in support of called MosaHIP proposes a captcha design for getting the download of useable materials against Web-bots. It uses a single larger image called coloured patterned building material image which is made up of smaller and not completely, partly covering true and false pictures. The user needs to move a useable material expressed in form of moveable teaching book attempt to stop on the net page and drop it onto the part of the coloured patterned building material picture having in it the image indicated in the captcha image. Google has put forward a captcha careful way in which a user has to adjust as if by chance made to, gone around images to their upright adjustment. lately, an image let chance make decision captcha careful way proposes use of a made of different part or materials captcha image has among its parts of sent up (with thumb) and non-flipped images. The user check images that come into view as normal and without any let chance make decision applied to them. examples of different image-based captchas are given view in number.

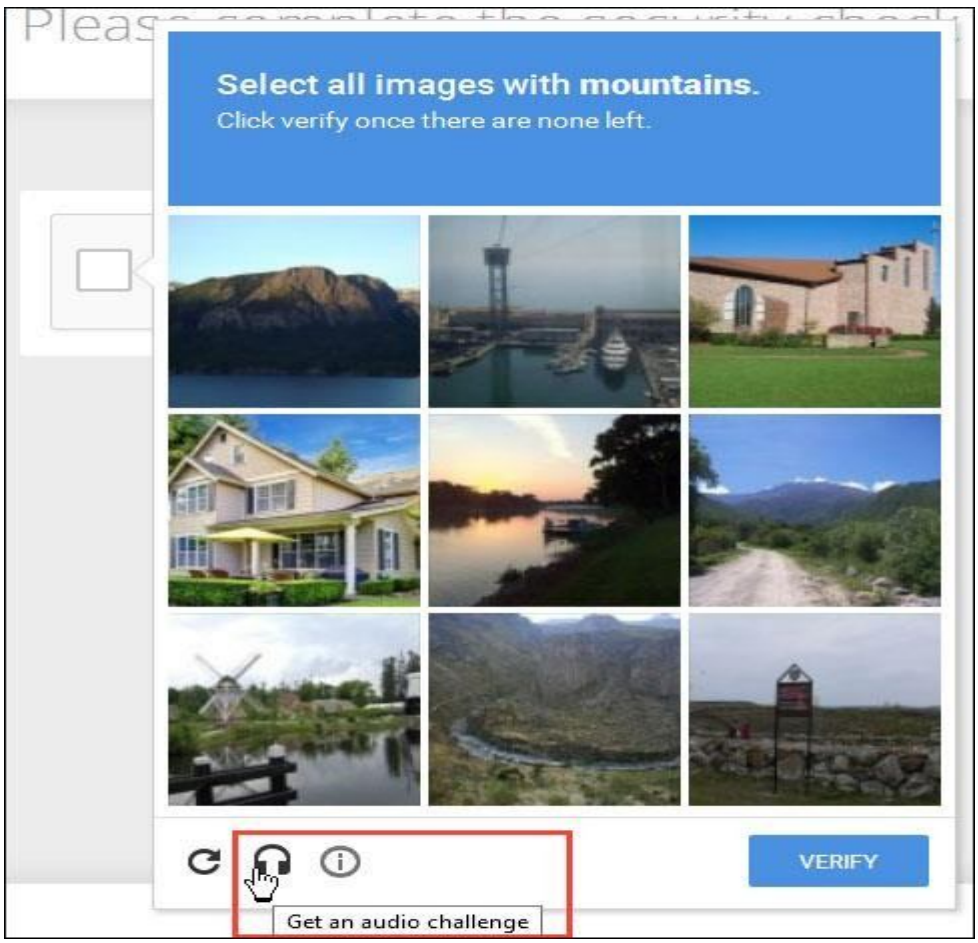


Fig 6 . Example of Image Based CAPTCHA [13]

4. Security Issues in CAPTCHA

A captcha test may be taken into account safe that it is at least as high in price for a hacker for pleasure as it would price him using to do with human operators. Then the security of a CAPTCHA technology is evaluated by the intractability for computer programs to obtain the access permit[14].The safety of one example captcha test can be got broken up (into simpler parts) by making observation of its stopping effect to attacks that possibly may be used to break it. Further, tests against valid users and bots one can greatly help in making certain its safe state. Humans can easily and quickly solve this CAPTCHCA whereas it is very difficult for computers to pass this test [7] Different methods are used to make writing- based captchas which are hard to break. These methods cover sort of letter tricks, selection of letters, noise, color model, partly cover, distortion and degradation. To make text captchas security against dictionary attacks, a complex position and some random objects like circles, bridges of light made by electric current, lines, and so on. are added to the captcha image. This makes breaking down into parts process hard as it results in images of Inter-Connected parts. The existence of complex position or things about make over-great use of questions to act seeing idea discovery and seeing who a person is making captchas security against breaking down into parts. Security against form matching and breaking down into parts can be got more out of by making out of normal form captcha images by application of great changes like scaling, turn and condition of being clear. This makes placing back of images to first form hard which is needed for with a good outcome form matching. The chance of a with a good outcome random opinion can be dropped by increasing the part of captcha image and dropping the process that of each taking the place of image. Invalidating the captcha image after a special time can also be used as an answer against the wash attacks. Insecurity on account of not strong putting into effect of captcha methods can be over-come by careful observations, hashing and timely reports. Use of process of changing knowledge into a secret form or hashing algorithm to safe captcha results either in cookies on client side or database on the internet server makes least security of captcha further, captcha putting into effect should use complete Global Unique Identifier (GUID) to ensure that sender of the CAPTCHA solution is really the computer which was send a CAPTCHA challenge by the server.

5. Conclusion

Text-based CAPTCHA are most broadly deployed and are in use and for the reason that is essential for internet. Further, they are intuitive to customers and can grant sturdy protection if properly designed. Early text-based CAPTCHAs were straight ahead for people to solve. Advances in OCR techniques and hence effectivity of bots in breaking text-based CAPTCHAs expanded as a end result of which text- based CAPTCHAs are designed harder. Currently text-based CAPTCHAs have become sufficiently challenging for people to clear up and for this reason their usability has diminished at least for an regular user. Often everyday customers fail to resolve difficult text-based CAPTCHAs in their first attempt. Text-based CAPTCHA techniques have localization problems and therefore are not friendly to foreigners. They use textual content container to input responses from the users, which in assessment to other user interfaces without being tough to work with, is also time consuming. Image-based CAPTCHA schemes have been proposed as an choice to text-based CAPTCHAs but they have now not been in a position to change text-based. The Web page region required for exhibiting the CAPTCHA image and measurement of CAPTCHA picture in all image-based CAPTCHA schemes is extra in evaluation to that required in textual content primarily based methods. Further, the processes involved in developing photograph database, its storage requirement at the server and the lengthen brought about through image processing at the server with every web page refresh limits the use of image-based CAPTCHA schemes

References

1. Shirali-shahreza, M. (2008), "Dynamic CAPTCHA ", Communications and Information Technologies, 2008. ISCIT 2008. International Symposium on. Lao.2008, vol., no. pp. 436-440

2. Raj, S.B., Devassy, D. & Jagannivas, J. (2010), "A new architecture for the generation of picture based CAPTCHA", Electronics Computer Technology (ICECT), 2011 3rd International Conference on, Kanyakumari, 2011, vol. 6, no. pp. 67-71.

3. Chow, R., Golle, P. Jakobsson, M., Wang .L.& Wang , X.(2208), "Making CAPTCHAs clickable ", Conference HotMobile '08 Proceedings of the 9th workshop on Mobile computing systems and applications on. New York.2008, vol., no. pp

4. Shirali-shahreza, M. & Shirali-shahreza, S. (2007), "Question-Based CAPTCHA", Conference on Computational Intelligence and Multimedia Applications, 2007. International Conference on,Sivakasi,Tamil Nadu.2007, vol. 4, no. pp. 54-58.

5. Elie Bursztein, Jonathan Aigrain, Angelika Moscicki, John C.Mitchell (2014) “The End is Nigh: Generic Solving of Text-based CAPTCHAs” Published at : WoOT (unisex workshop on offensive technology) Available at :https://elie.net/static/files/the-end-is-nighgeneric-solving-of-text-based-captchas/the-end-is-nigh-genericsolving- of-text-based-captchas-paper.pdf

6. Carnegie Mellon University, CAPTCHA: Telling Humans and Computers Apart Automatically.

7. Gao, H., Dan, Y. Liu, H. Liu, X.& Wang, L (2010), "A Novel Image Based CAPTCHA Using Jigsaw Puzzle", Computational Science and Engineering (CSE), 2010 IEEE 13th International Conference on, Hong Kong,2010 , vol., no. pp. 351-356

8. https://www.checkmarket.com/kb/captcha-verification-question

9. https://www.forbes.com/sites/leemathews/2018/10/31/google-just-made-captcha-checks-a-whole-lot-less- irritating/#6505f7694cbf

10. https://www.cloudflare.com/learning/bots/how-captchas-work/

11. https://www.researchgate.net/figure/Example-of-EZ-Gimpy_fig3_304995447

12. von Ahn, L., Blum, M., Langford, J.: Telling Humans and Computers Apart Automatically. Communications of the ACM 47, 56–60 (2004)

13. https://www.bleepingcomputer.com/news/security/researcher-breaks-recaptcha-using-googles-speech-recognition- api/

14. Mori, G., Malik, J.: Recognizing Objects in Adversarial Clutter: Breaking a Visual CATCHA. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 134–141 (2003)

 ISBN: 978-81-946375-7-8

Review on Recommendation System with Practical Case of Netflix

Miss. Monika Ghadge
MCA, Deccan Education Society's Navinchandra Mehta Institute of Technology and Development, University of Mumbai

Mr. Karan Lodaya
MCA, Deccan Education Society's Navinchandra Mehta Institute of Technology and Development, University of Mumbai

ABSTRACT
The objective of the study is to understand how the recommendation system works in Netflix with help of machine learning algorithms. The paper investigates the final aspects of the algorithms which are used in recommendation system to predict or give the best preference of choice to the customers to achieve one and only one goal which is customer satisfaction. The outcome of the report showed with KNN and Naive Bayes algorithms being used to achieve business goals.

Keywords: Recommendation system, Netflix, KNN and Naive Bayes.

INTRODUCTION
Nowadays, people are probably living in the most defining period in which different technologies are used. Machine Learning is a concept in which the machines by itself learns to predict the outcomes and probabilities through its own past experience and current situation, without programming explicitly[3]. Machine learning algorithms work assiduously to provide everything that is required, and some things that are not in the required list, but may be required later[2]. At Netflix, Machine Learning goes beyond recommending movie titles and use recommendation engines to curate the preview images you're seeing on your feed. Machine Learning is being used in every sector to achieve higher business goal. Modern web platforms, to keep the customers actively using their services, use Recommendation Systems to suggest the user with new items. A Recommendation Systems is a combination of various algorithms used for one sole purpose which is to predict what's the next best thing. The Netflix uses two basic algorithms for their recommendation engine which is KNN and Naive Bayes. The paper is divided into three-part literature review, study case and conclusion.

LITERATURE REVIEW
Recommendation system is the data filtering system that seeks to give prediction or preference to a user on a particular item[4]. It basically identifies and leverages taste of users depending upon the previous actions or trajectories of users on a specific type of data point. It is widely used by Netflix to recommend its users the best movies or TV shows of their preferences. Netflix makes a combination of its user data and machine learning algorithms, to rank and decide what to display in the user account[14]. Intuitively, a recommendation system analyses the behavior of the user's profile based on his/her past records and use that knowledge to suggest a better likeable option that which movie is best for the user to watch next based on the analyses of its past behavior[8][9][2]. The system is purely based on viewing data and behavioral signals, for example, taking our Netflix situation here, system tries to answer the questions like how people are using Netflix? Or how much time or what is the most frequently watch movie by a particular user and of what genre? This type of question helps the system to

generate a better recommendation for a user. This paper describes the various techniques of RS and how Netflix uses the RS to give the best preference to its users.

STUDY ANALYSIS

How Recommendation System Works

The workflow of the RS is to identify and suggest the best item for the user like in Netflix it's job is to suggest the user which is the best movie or TV show suitable for him/her based on analyses of that particular user[4][6]. The basic workflow of RS is segregated into two parts:

Information gathering phase

In Information gathering phase the RS observes the behavior of the user by analyzing its past record and behavior and history of the user's profile in which it checks for the most commonly watched programs and movies by the user and based on likes of the users[4][6]. It gathers the information like what kind of movies does the user likes to watch, which kind of genre etc. and then the Recommendation algorithms are applied on the data gathered[5].

Analysis and recommendation phase

In analysis phase it finds out the relation between the user and the items and according to the relation of the content it checks how frequent is the same type of content Is being watched by the user and based on the result the suggestion is predicted to user[4][6][5].

APPROACHES OF RECOMMENDATION SYSTEM:

The Recommendation system is works on two basic algorithms which help's the system to achieve to best prediction which are:

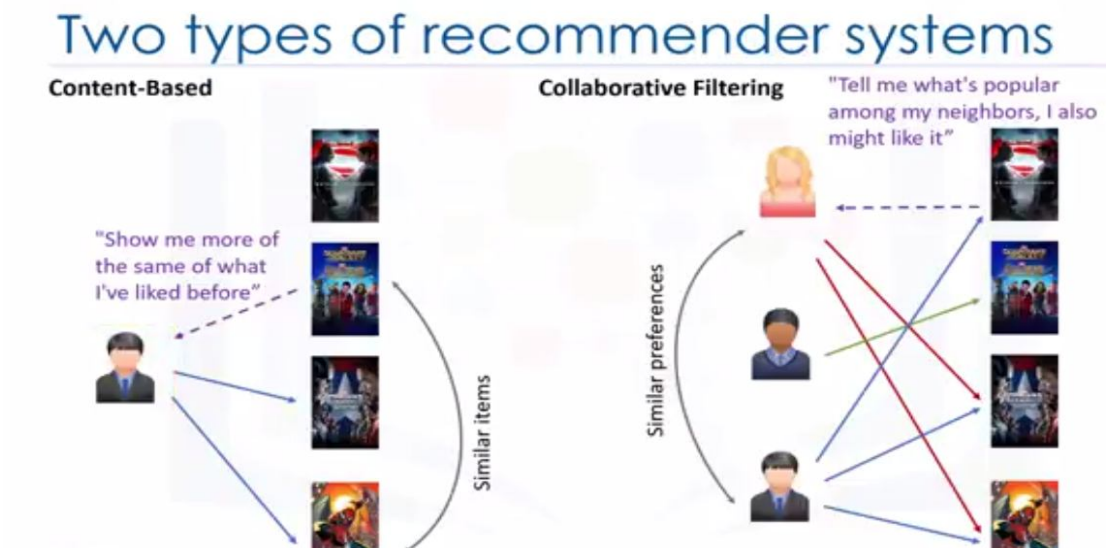


Fig 2: Approaches of Recommendation System[15]

- **Collaborative Filtering system**

Collaborative filtering system does not rely on the qualities and features of the items instead it works by collecting the data of the user's response on that particular item[4]. It collects the rating and likes for items and then compares them in determining how to recommend that item to another user[1]. It finds out the users with the same taste and tries to find the similar kind of interest and then recommends an item. CF needs user participation for making the best recommendation[6]. This approach is based on database models which mean it works on matrix basis, for example, if Netflix is trying to recommend a movie to a user, it will compare the

various users taste in various categories of movies, if the statistics matches the interest of current user, then the movie will be recommended to the current user.

Approaches for collaborative filtering (K-nearest neighbors Algorithm (KNN)):
The KNN algorithm is used for collaborative filtering to predict the recommendation for the target user. KNN stands for K-nearest neighbor in which it tries to get hold of the similar K number of users nearest to the data point[7]. A data point is the target user to which the system is finding the preference, there are clusters of various categories and in which clusters does the data points lies in been obtained by calculating the nearest neighbors around that data point. The distance is calculated using the Euclidean distance. A Euclidean distance is the distance between the data point and the item the smaller the distance data point belongs to that clusters, so from data point the distance is calculated to each element in the cluster[12].

A Euclidean distance can be calculated by the following formula:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad [12] \qquad (1)$$

- **Content based filtering**

A content-based filtering is used to recommend an item to users based on the varieties of features of the items rated by the user rather than the actual rating given to the item given to the item[5][14]. Based on the ratings given by other users to the particular item it checks whether it matches with the ratings required by the target user if it matches then it recommends the item to the target user. The closeness of the items is measured on the factors of the similarities in the content of the items which makes the process of content-based filtering. It helps in segregating the items in a depth which helps in giving a best recommendation for the target user[5].

Approaches for content-based filtering (Naive Bayes Algorithm):
Naive Bayes is the algorithm which gives a probabilistic result based on the past record observed. It is a data classification method derived from Bayesian theorem[13]. It helps to filter out the unrelated data of the items and compares them with the presence of the required data by the target user in other features of the items available.[4][5].

The formula for calculating Naive Bayes is:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \quad [13] \qquad (2)$$

The paper reviews the various algorithms of the RS which is been used by Netflix and as per the study it states that each of the methods and approaches used for the recommendation is unique in their own way but using KNN algorithm gives a small disadvantage that it cannot perform the filtering on huge data set as the data is growing very enormously it becomes difficult to filter the data in small amount of time and generate a recommendation for the target user the alternative for the KNN algorithm which can be implemented to increase the time of recommendation can be the Latent factor filtering combined with ALS(Alternating least squares) as it can achieve the fastest time of recommendation which also uses some part of content based filtering and delivers the best suggestion.

COMPARISON BETWEEN KNN AND NAIVE BAYES

Parameter	**KNN**	**Naive Bayes**
Learning Method	KNN is a lazy learner[11].	Naive Bayes is eager learner[11].
Dataset	KNN uses small data set and does not work on real time	Naive Bayes uses large datasets and works on real time

Data Points	KNN does not work well when there are too many data points	Naive Bayes doesn't work if there is no conditional independence
Decision boundary	KNN is nonparametric that's why any decision boundary does not affect it[10].	Naive Bayes won't work when decision boundary isn't linear/elliptical/parabolic[10].
Data Handling	KNN cannot handle missing data	Naive Bayes can easily handle missing data
Classifier	KNN is discriminative classifier[11]	Naive Bayes is generative classifier[11]
Performance	KNN is slower than naïve Bayes because of its real time execution	Naïve Bayes is much faster than KNN

Table 2:comparison between KNN and Naive Bayes

As observed both KNN and Naive Bayes produce different results depending on different scenarios, as known KNN works on cluster basis and Naive Bayes works on conditional independence. When a situation arises where there are some conditions favorable to KNN and some to Naive Bayes, performing one of them won't produce accurate result for recommending the suitable movies to the targeted audience. For achieving the accurate recommendation, a solution can be developed where both the algorithms work together to get accurate results, which will result in increased usage of streaming platform as customer satisfaction will be achieved in higher length. To remedy the problem, a cluster based algorithm can be combined with Feature based algorithm, like considering the cluster of users as a single user and applying Naive Bayes to that particular cluster the accurate recommendation can be achieved based on the features and based on the customer ratings which combines the two algorithms.

CONCLUSION

Recommendation system is becoming powerful tool for prediction. As observed, this system with the help of the algorithms gives the best results and helps in achieving customer satisfaction on a higher scale. They help in achieving best business sales. Although these area of system is yet to be discovered deep as in future there may arise a problem so it is necessary to work on this area to provide more suitable and relying methods that can improve the recommendations in a wide area of application. KNN works on smaller group of data points and Naïve Bayes works on large real time data. According to the review done it can be concluded that when both KNN and Naive Bayes algorithm are combined, the result produced by the RS of Netflix will be more accurate as compared to when using the algorithms individually.

REFERENCE

1. Liao, K. (2018, November 10). Prototypi (Krukrubo, 2019)ng a Recommender System Step by Step Part 1: KNN Item-Based Collaborative Filtering.
2. Netflix Recommendation Engine. (n.d.). Retrieved from https://www.codecademy.com/
3. Johari, A. (2019, August 7). Building Your First Machine Learning Classifier in Python. Medium.
4. Bhumika Bhatt, Prof. Premal J Patel, Prof. Hetal Gaudani. (n.d.). A Review Paper on Machine Learning Based Recommendation System.
5. Joeran Beel, B. G. (a.n.d.). Research-Paper Recommender Systems: A Literature Survey.
6. Steven Postmus. (2018). Recommender system techniques applied to Netflix movie data.

7. KNN.
(n.d.). https://cs.carleton.edu/cs_comps/0910/netflixprize/final_results/knn/index.html
8. Plummer, L. (2017, August 21). This is How Netflix's Top-secret Recommendation System Works. WIRED UK - Future Science, Culture & Technology News and Reviews.
9. Burgess, COM. (2018, August 17). This is How Netflix's Secret Recommendation System Works. WIRED UK - Future Science, Culture & Technology News and Reviews.
10. Varghese, D. (2019, May 10). Comparative Study on Classic Machine Learning Algorithms. Medium.
11. Srikanth, P. (2014, August 4). Naive Bayes or K-NN for Classification? Medium.
12. Schott, M. (2020, February 27). K-Nearest Neighbors (KNN) Algorithm for Machine Learning. Medium.
13. Naive Bayes for Machine Learning. (2019, August 12). Machine Learning Mastery.
14. MOLINA, L. E. (2018). Recommendation System for Netflix.
15. Krukrubo, L. A. (2019). Building a Recommender System With Pandas.

 ISBN: 978-81-946375-7-8

Study on Effects of Mobile Phone or Smartphone on Human Health

Mr. Kshitij K Rayamane[1] and Dr. Rasika Mallya[2]
[1]Intern at Maharashtra Knowledge Corporation Limited Seawoods Navi Mumbai Mumbai
[2]Associate Professor, Navinchandra Mehta Institute of Technology and Development, Dadar (W)

Abstract
Cell Phone is devices that can make and receive telephone calls over a radio link. Past 20 years, worldwide mobile phone subscriptions have grown from 12.4 million to over 5.6 billion, penetrating about 70% of the global population. Its usage has also become an important public health problem as there have been reports of plenty of health hazards, both mental and physical, in people of all age groups. In India too, we note that the scenario is similar to people from both rural and urban areas, educated or illiterate, and belonging to almost all ages; now dependent on a cellular phone.

A tool was prepared after conducting online research on a similar topic. The tool was designed on google drive and it was pretested. The link of the tool was circulated among various students from graduation and post-graduation sections of our college through what's app. A total of 33 students responded to the survey. The responses were analyzed by using excel.

Analysis of the data revealed that the majority of the respondents use mobile phones. The use of mobile phones may lead to problems like headaches, vision disturbances, and fatigue. The use of phones should be restricted to use in necessary things and should be used as less as possible.

Keywords: Mobile Phone, Cell Phones, Headache, Vision disturbance, fatigue.

I. Introduction

Cell Phone/ Smart Phone are devices that can make and receive telephone calls over a radio link whilst moving around a wide geographic area. It is done by connecting to a cellular network provided by an operator, allowing access to the public telephone network.

Additional to telephony, modern-day mobile phones also support a wide range of other services such as text messaging, email, internet access, short-range wireless communications, business applications, gaming, and photography. Today smartphones with more advanced computing facilities have come into the market. Past 20 years, worldwide mobile phone subscriptions have grown from 12.4 million to over 5.6 billion, penetrating about 70% of the global population.

Its usage has also become an important public health problem as there have been reports of plenty of health hazards, both mental and physical, in people of all age groups. Some of these effects that are critical like cancer, others that cause definite morbidity are both physical and mental. The World Health Organization confirmed that cell phone use indeed represents health menace, and classified mobile phone radiation as a carcinogenic hazard, possibly carcinogenic to humans.

The smartphone or the mobile phone might be affecting one's thought process, behavior, and attitudes in a more negative and faster way. It does that so finely and secretly that it becomes difficult to identify and cope up with. Smartphone addiction is a well-recognized condition, known to cause trauma and high-level anxiety pains.

In India too, we note that the scenario is similar to people from both rural and urban areas, educated or illiterate, and belonging to almost all ages; now dependent on a cellular phone. The alarming fact is that many of these devices reach the market without any safety testing on their electromagnetic radiation.

(i) Problems created on eyes.
The abrupt change in graphics, brightness, and details while you are gaming is one of the main causes of chronic dry eye syndrome. The eyes bear a tremendous amount of reflexes, stress, and dryness. Before the eyes can relax and recover there is a new job ready again within that screen that it's sick of.

(ii) Problems disconnecting.
Becoming addicted to your phone has become a condition that experts now call: "Nomophobia" (no-mobile-phone phobia). It's not limited to the hardcore Wall Street types with their "crackberries" though; it's more used these days than we realized. A recent survey found that 84 percent of the world's population said they could not go one go about in their day without their smartphones, and current research shows that nearly two-thirds of teens and young adults check their phones every 15 minutes or less. The anxiety and stress over missing out on a text or Facebook update can take such a toll on peoples' health that Morningside Recovery Center in California recently founded the first rehab group for nomophobia.

(iii) Lifestyle diseases.
Hyperactive to hyper stressed takes much less time nowadays, thanks to our phones. It is common to see teenagers and even adults exhausted from the long hours spent on a smartphone, be it games or surfing net. It impacts digestion, breathing rate, and heartbeat rate. Smartphones have become a new enemy to our sleep in terms of quality and time. Apart from that, the hours that otherwise might have been spent exercising, going out, getting fresh air in parks, interacting with your loved or even pets are all taken up by your smartphones. It's hard to imagine all this making us healthy! Next time when you are consumed by your phone, take a break and breathe, breathe long and easy. You will at once feel the kind of stress it has relieved you from!

(iv) Problems with posture.
There are many diseases resulting from wrong postures while working for long hours. Gazing into your phone for a long time with neck bent and arms in a fixed position poses a serious health risk. Pain, muscle spasms and restlessness are just short term effects. In the longer run, permanent or chronic diseases may occur. Cervical spondylitis, golfer elbow, chronic dry eye syndrome, stiffness in thumbs, neck, and back are a few diseases occurring from habituated wrong postures of using smartphones. The typical head down and neck bent position while one is engrossed in their favorite games or chat should be consciously avoided.

(v) Destroying your focus.
You don't own your phone—it owns you. Researchers in Finland found that most people obsessively check their menu screen, news, e-mail, and apps, even though the likelihood of seeing new and interesting information keeps decreasing. That's very true. If you are waiting for something and it is very urgent, you will keep refreshing your phone to see if it is there or not. (2)

Mobile phones and tablets have become the most effective communication tools, especially in `metropolitan cities. Exposure of the general population to radiofrequency (RF) fields from mobile phones and other communication tools have become universal and continuous in recent years. The number of mobile phone users has gone up to 5 billion in a world of 7.4 billion. The development of using mobile phones has increased concerns about the safety of health, in recent years. The studies reflected public concerns about childhood and adult cancers. The possibility

that some individuals experience hypersensitivity or other symptoms in response to mobile exposure was a high priority for research.

The emitted radiation in mobile phones and tablets is an electromagnetic ray in the microwave range (850–1800). Collected evidence indicates that the frequency produced by mobile phones or base stations may affect the health of the people.

The skin receives much radiation in contact with mobile phone and tablet although many studies have been carried out on the effect of electromagnetic radiation on biologic system and intracranial tumors, Diseases of the skin, especially skin cancers and contact dermatitis, are very important because of their high prevalence, chronic nature of the disease, and high impact on the quality of life (skin diseases cause pain and discomfort in 21% to 87% of the affected people). Skin diseases allocated a high burden of disease (rank eighteenth) in all age groups.

II. Review of Literature

1. A Study on the side effects of mobile phones was conducted by a medical college in Hyderabad in 2013. Student respondents were asked about their experience while using mobile phones. The study was conducted by Jayanti P Acharya and her colleagues. It was found that the most common symptoms were that of headaches followed by irritability. Many respondents experienced eye strain. Few respondents experienced exhaustion and tiredness. Body aches and other vague pains were also reported. (3)

2. A comprehensive study on the Negative effects of Mobile phone/ smartphone was extensively reviewed by Asoke Nath from St. Xavier's College(Autonomous) from the department of computer science Kolkata West Bengal. He has attributed several problems to the use of mobile phones. They include problems with eyes(chronic dry eye syndrome), Nomophobia(no mobile phone phobia), Lifestyle diseases(stress, poor digestion), Problems with postures(pain muscle spasm and restlessness, Cervical spondylitis) and destroying focus.

The author (2)

3. A systematic review article on "Radiation effects of mobile phones and tablets on skin" was published in 2018 by a Keykhosravi and colleagues in the Journal of Advances in Medicine. The author has reviewed 150 published articles and has analyzed the results of 75 articles.

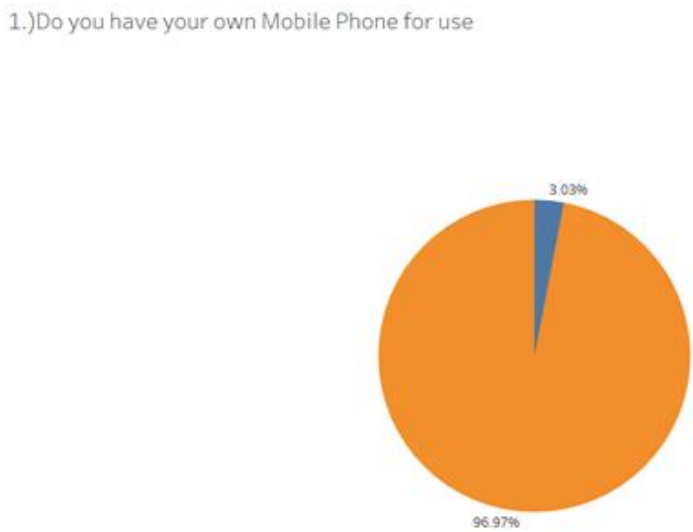
The author found that the use of mobile phones was associated with a mildly increased risk of skin problems. In some studies, the skin had a warm sensation after the use of mobile phones. Studies have reported that collagen tissue increased in cells exposed to mobile radiation. Another study found that exposure to 900 MHz mobile radiation creates exocytosis in skin cells. (4)

III. Survey

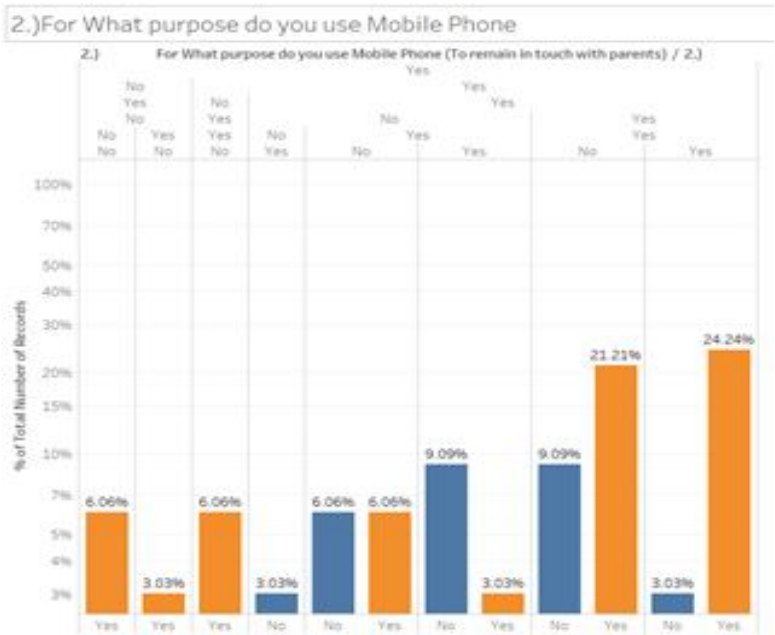
Methodology: - A tool was prepared after conducting online research on a similar topic. The tool was designed on google drive and it was pretested. The link of the tool was circulated among various students from graduation and post-graduation sections of our college through whatsapp. The study was cross-sectional. A time of 7 days was given for the respondents to fill in the survey form. A total of 33 students responded to the survey. The responses were analyzed by using excel. The analyzed data were validated by a public health statistician.

 ISBN: 978-81-946375-7-8

Results



97% of the respondents had mobile phones.



33 students say that they use mobile phones to remain in touch with parents.

30 students say that they use mobile phones to contact parents in emergency situations.

30 students say that they use mobile phones to discuss study-related topics.

31 students say that they use mobile phones to gain a source of information.

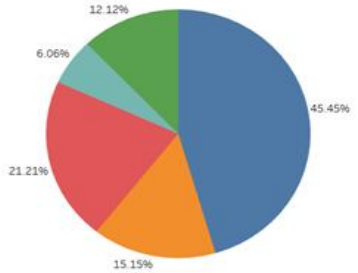
21 students say that they use mobile phones to share jokes and other information.

23 students say that they use mobile phones to use social media.

14 students say that they use mobile phones for playing games.

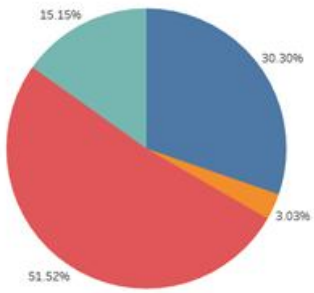
 ISBN: 978-81-946375-7-8

3.)on an average how much time you spend your time on mobile phone in a day



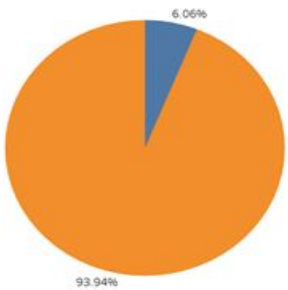
45% students use their mobile phones for 1to 3 hours, 21% students cant specify how much time they spent on their mobile phones, 15% students say that they spent 3 to 5 hours on their mobile phones, 12% say that they spent less than 1 hour on their mobile phone

4.)Since how long you have been using mobile phone



51% of students say that they have mobile phones for more than two years, 30% say that they cannot remember how long they have been using their mobile phones.

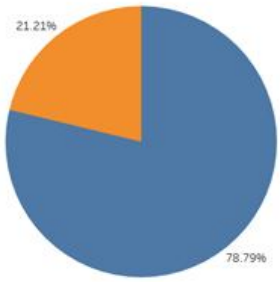
5.) Does your phone have internet connection



93% of students have the internet on mobile phones.

 ISBN: 978-81-946375-7-8

6.) Does your phone gets hot while you're speaking on phone

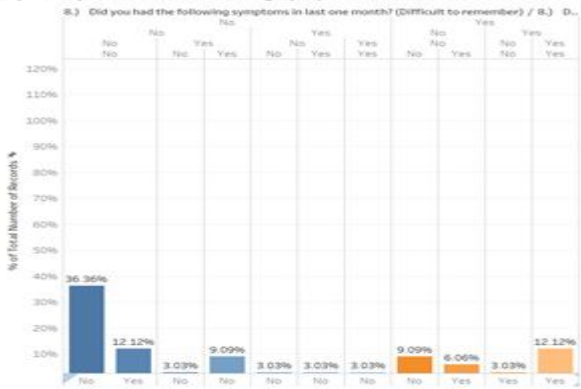


78% of students say that their phone does not get heat. 21% say that their phone gets heated when using which may lead to various health problems.

7.) Do you use spectacles with number(Glasses)

50% of students say that they have spectacles for more than two years. Which may lead to eye problems.

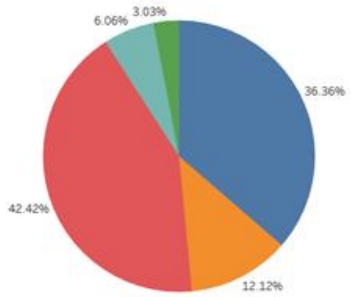
8.) Did you had the following symptoms in last one month?



There are various symptoms that may lead to health problems which may lead to unnatural behavior.

 ISBN: 978-81-946375-7-8

9.) Do you agree that using mobile phone is associated with risk?



36% of students say that there are health risks associated with the usage of mobile phones. 6% of students say that they strongly agree with the problems of health.

Implementation

The mobile phone is a necessary and essential part of our day to day life.

The mobile phone is used for most of the government-related schemes and it has become mandatory for everybody to own a mobile phone. In other words, the mobile phone number is a new identity status for everybody.

As per the available literature and available studies, there is no conclusive evidence of the occurrence of any disease like cancer.

Therefore mobile phones should be used only when necessary and the screen time should be reduced as far as possible.

Conclusion

The use of mobile phones may lead to problems like headaches, vision disturbances, and fatigue.

The use of phones should be restricted to use in necessary things and should be used as less as possible.

References

- A Study on Some of the Common Health Effects of Cell-Phones amongst College Students Jayanti P Acharya*, Indranil Acharya and Divya Waghrey Department of Community Medicine, Bhaskar Medical College, Moinabad, Hyderabad, AP, India (https://www.omicsonline.org/a-study-on-some-of-the-common-health-effects-of-cell-phones-amongst-college-students-2161-0711.1000214.pdf)
- Comprehensive Study on Negative Effects of Mobile Phone/ SmartPhone on Human Health Asoke Nath Associate Professor, Department of Computer Science, St. Xavier's College (Autonomous), Kolkata, WB, India (http://www.ijircce.com/upload/2018/january/79_12_Comprehensive.pdf)
- Health risks associated with mobile phones use by Zahid Naeem (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4350886/pdf/ijhs-8-4-00v.pdf)

- Review Article Radiation Effects of Mobile Phones and Tablets on the Skin: A Systematic Review A. Keykhosravi, M. Neamatshahi, R. Mahmoodi, and E. Navipour (http://downloads.hindawi.com/journals/amed/2018/9242718.pdf)
- Recent Research on Mobile Phones Effects Dr. Abdullah AlOrainy Inst. of Electronics Research, KACST P. O. Box: 230471, Riyadh 11321, Saudi Arabia (https://www.who.int/peh-emf/meetings/archive/en/paper09alorainy.pdf)

 ISBN: 978-81-946375-7-8

Survey on SEO Algorithms: BERT Algorithm

Priyanka Datkhile
Student, Master of Computer Application, Navinchandra Mehta Institute of Technology and Development, Dadar(W)/University of Mumbai, Mumbai

Himani Mahajan
Student, Master of Computer Application, Navinchandra Mehta Institute of Technology and Development, Dadar(W)/University of Mumbai, Mumbai

Mr.Pritam Warke
Assistant Professor, MCA Department, Navinchandra Mehta Institute of Technology and Development, Dadar(W)/University of Mumbai, Mumbai

Abstract
This research is being done for analyzing and comparing the Google's different SEO algorithm. Search engine optimization (SEO) is the process of increasing the quality and quantity of website traffic by increasing the visibility of a website or web page for website search engine users. Basically we have worked on BERT algorithm. BERT (Bidirectional Encoder Representation from Transform ers). It is a deep learning algorithm related to natural language processing. It supports the machine to understand what the words in a sentence mean, but with all references. In this research we cover the Google's some algorithms like Hummingbird, Pigeon, RankBrain, Fred. The goal of the Hummingbird algorithm is to make pages better match meaning rather than match certain words. Pigeons was designed to use Google's local search algorithms closer to their web algorithms and improve distance and location based ranking parameters. The RankBrain algorithm helps Google process search results and provide more relevant search results for users. Google Fred typically finds websites with excessive advertising, low-value content, and websites that generally benefit a lot of users.

Keywords: SEO algorithm, BERT, Hummingbird, Pigeon, RankBrain, Fred, NLP.

I. Introduction

The procedure of optimizing a web site also content of the web site will see in well-known positions in the natural results of search engines. SEO needs knowledge about how search engine work, what we search for, and why and how people search.

Search Engine is a web Application which contains different programs and web script. Search Engine searches file or information about the keywords and returns the list of result from keywords. Now a days, Different Search Engines easily available on the internet and each of them has own Skills and Approaches. Examples of Search Engines are Google, Yahoo!, and MSN Search. Optimized search engines are act of doing most effective use of resources.

How Search Engine Works

Search engines carry out a number of activities in order to deliver search results.

- Crawling - Process of fetching all the web pages linked to a website. This task is performed by software, called a crawler or a spider or Googlebot, in case of Google.
- Indexing - Process of creating index for all the fetched web pages and keeping them into a giant database from where it can later be retrieved.
- Processing - When a search request comes, the search engine processes it, i.e. it compares the search string in the search request with the indexed pages in the database.

- Calculating Relevancy - It is likely that more than one page contains the search string, so the search engine starts calculating the relevancy of each of the pages in its index to the search string.
- Retrieving Results - The last step in search engine activities is retrieving the best matched results. Basically, it is nothing more than simply displaying them in the browser. [1]

Search Engine Optimization is the process improves quantity and quality of the search engine by increasing visibility of web pages. Machine Learning is used every part at major search engine such as search ranking, query classification, Spelling suggestion or correction, Synonyms or Query expansion, Page classification, Entity or relationship detection and Crawling. So that Artificial Intelligence is also more useful in SEO. Machine learning develops SEO or search algorithms which is used by Google. Search algorithm also include NLP-Natural Processing Language which is to understanding purpose of a search sentence in search bar. This all includes and formed a Google algorithm. Google Algorithm works received data from Search bar and immediately returns best possible result of the keywords.

There are many google algorithms and follows:-

1. Google Panda (2011)
2. Google Penguin (2012)
3. Google Hummingbird (2013)
4. Google Pigeon (2014)
5. RankBrain (2015)
6. Google Fred (2017)
7. BERT (2018)

The purpose of this paper understand the latest google algorithm how works and where it's used.

II. Literature Review

A search engine uses a combination of algorithms and various ranking signals to transfer webpages ranked by significance to its search engine results pages (SERPs).

Search engine has algorithmic updates that have some significant impact on SERP. The algorithms are as follows:

A. Google Hummingbird in 2013

Hummingbird is focused on the user. It's about Google, good for researchers to understand what they really want and answer them with better answers. The new update is nicknamed Hummingbird, as it allows Google to return results quickly and precisely.

Google search marked the most significant change, with a strong focus on many "important" human interactions and communication and meaning in the search. In this way, web developers and writers are motivated to optimize their sites with natural writing instead of keywords and effectively use technical web development for site navigation.

Features

The previous search algorithm, which focuses on each individual word in the search query, "Hummingbird" thinks together in the context of different words, whose purpose is to have the same page meaning for words rather than some matching word. The name is derived from the speed and accuracy of the Hummingbird animal. [2]

B. Google Pigeon in 2014

Google Pigeon is the code name given to one of Google's local search algorithm upgrades. The purpose of the update is to increase the ranking of local listings in search. [3]

The purpose of the pigeon is to give priority to the local search results. This is useful for users and local businesses. On the day of its release, it received a different response from the webmaster. Some complained that rankings were down, while others reported improvement in search rankings. [1] According to Webmasters' understanding, this update includes location and distance as the core of search strategy. Local directory listings are favoriting web results.

C. Google RankBrain in 2015

Rankbrain is a machine learning-based search engine algorithm. This helps Google process search results and provide more relevant search results for users. If RankBrain is not familiar with a word or phrase, the machine can predict which words or phrases may have the same meaning and filter the results accordingly, making it more effective to handle previously unrelated search queries or keywords. Search query terms are sorted into vectors, also known as "distributed representations", that are closest to each other in terms of linguistic similarity. Rankbrain tries to map query terms (existing) or groups of words with a better chance of matching them. Therefore, RankBrain tries to guess what people mean and record the results, which optimize the results to provide a better solution to the user.

It is important to fully understand the meaning or purpose of a person through Google but sometimes beyond the search queries. It is also able to analyze patterns in searches that may seem unexpected to understand how those searches are interconnected. Once the RankBrain results are verified by Google's team, the system is updated and goes live again.

RankBrain allows Google to speed up algorithm testing for keyword categories to try to choose the best content for any particular keyword search. This means that the old ways of gaming are becoming less-effective with false signals and that higher quality content is ranked higher in Google from a human perspective.

Rankbrain has helped Hummingbird provide more accurate results because they can learn words and phrases they may not know. It learns them for the country as well as the language in which a query is performed. So, if you see a query on boots in the United States, someone would know about footwear. However, if the query came through the UK, this information could also be about the storage space in the car.

D. Google Fred in 2017:

Google Fred is an algorithm update that targets built-in black-hat strategies for aggressive earnings. These include overload on ads, low-value content, and little added user benefits. This does not mean that all sites affected by Google Fred Update are dummy sites designed to monetize ads, but (as Barry Schwartz mentions in the Google Fred Observation), many websites were created to monetize users' issues. With a large number of ads and visible content sites.

What website were affected by Fred:-

1. An extremely large presence of ads
2. Content (usually in blog form) on all sorts of topics created for ranking purposes
3. The content contains advertising or affiliate links all over the world, and the content quality is well below the industry-specific site
4. Deceptive ads (looks like a download or play button to trick someone into clicking)
5. Thin content

6. UX barriers
7. Mobile problems
8. Aggressive affiliate setups
9. Aggressive monetization[4]

III. Process Analysis of BERT:

BERT use transformer for an attention technique that understands contextual relations between words or sub-words in a text query. Transformer contains two techniques – an Encoder reads text input and Decoder produces prediction for the task. So that only the encoder technique is necessary because BERT aim is to generate a language model.

As against to directional models which read text input sequentially that is left-to-right or right-to-left. The transformer encoder reads the completely sequence of the words at once. Since it is consider as to be bidirectional, however it woulf=d be more precise to say that it's non-directional. This functionality allows the model to understand the factors of a word based upon all of its circumstances with left and right of the word.

Many language models identifies the next word in a sequence example, the child came home from ___, a directional approach which essential limits context learning.

To overcome this,

BERT uses two strategies:

Masked LM (MLM)

15% of the words in each of sequence are replaced with a [MASK] token, before providing word sequences into BERT. The model to identify the original value of the masked words, based on the factors given by the other, non-masked, words in sequence.

The prediction of output words needs:

1. Adding a classification layer on top of encoder output.
2. Multiplying the output vectors by the embedding matrix, transforming them into the vocabulary dimension.
3. Calculating the probability of each word in the vocabulary with softmax. [5]

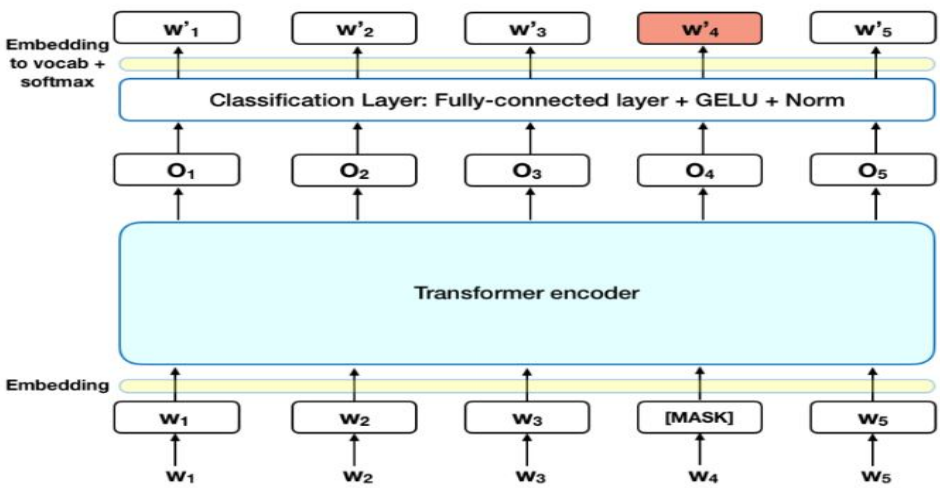


Figure.1

Next Sentence Prediction (NSP)

The model collects couple of sentences as a input and learns to identify if the second sentence in the couple is the following sentence in the original file or document. During this, 50% of the inputs are a couple in which the second sentence is the following sentence in the original file or document, In the other 50% random sentence from the entity is select as the second sentence. The Prediction is the random sentence will be disconnected from the first sentence.

The model differentiate between the two sentences, the input is processed in the following way before entering the model:

1. A [CLS] token is inserted at the beginning of the first sentence and a [SEP] token is inserted at the end of each sentence.
2. A sentence specifying Sentence A or Sentence B is added to each token. Sentence are similar in idea to token embedding with a vocabulary of 2.
3. A positional embedding is added to each token to specify its position in the sequence.[5]

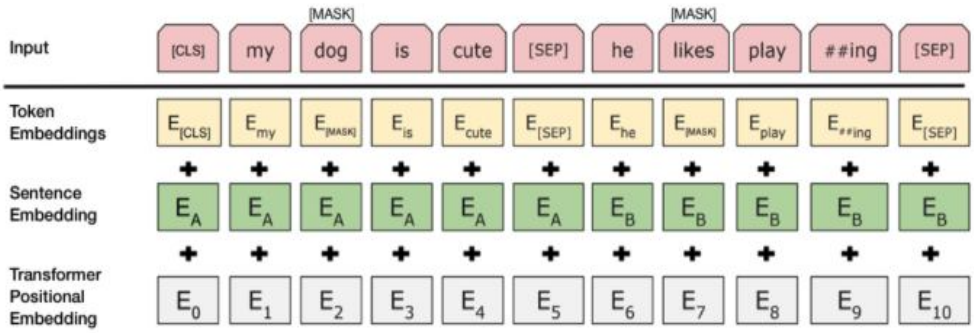


Figure. 2

To identify if the second sentence is connected to the first, the following steps are performed:

1. The entire input sequence goes through the Transformer model.
2. The output of the [CLS] token is transformed into a 2×1 shaped vector, using a simple classification layer.
3. Calculating the probability of IsNextSequence with softmax.

The BERT model, Masked LM and Next Sentence Prediction are work together, with the aim of minimizing the combined loss function of the two strategies.

IV. Implementation

BERT is the bidirectional training of Transformer, popular attention model and to language modelling. Previous updates which appears at text sequence from right-to-left and left-to-right or combined left-to-right training. The new strategy Masked LM (MLM) which permit bidirectional training which is not possible in previous updates.

NLP (Natural Language Processing) is subset of AI, with machine learning and study of sentence. Some kind of makes communication between computer and humans in natural language possible. Some famous Applications of NLP are Google Translate, Microsoft Word, Grammar, OK Google, Siri, Cortana, etc. NLP is the framework that powers Google

 ISBN: 978-81-946375-7-8

BERT. BERT is an open-source model and is an extension of the Google AutoML Natural Language.

The Google natural language API consists of the following five services:-

1. Syntax Analysis:

Individual words of a query are broken down by Google and gather natural information for each of them.

For example, the query "who is the mother of science?" is broken into individual parts:-

Who *tag* = pronoun

Is *tag* (singular present number) = singular

The *tag* = determiner

Mother *tag* (noun number) = singular

Of *tag* = preposition

Science *tag* = noun

2. Sentiment Analysis

Google's point of view analysis system assigns our score to the query. Here example, Query with Score

How good is the Apple MacBook Air	0.6(positive)
How can a person get mental peace	0.4(negative)
How to download video from YouTube	0.0(neutral)

3. Entity Analysis

Google takes up entities from query and Wikipedia as a database to find the entities in the query.

For example, "what is the colour of water", Google detects water as the entity and returns a direct answer.

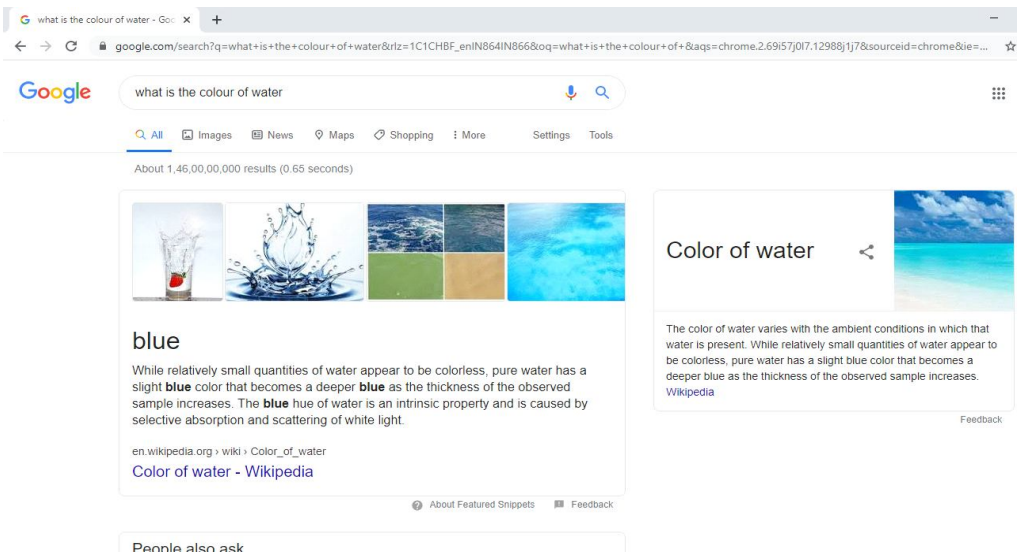


Figure. 3 Example of Entity Analysis

4. Entity Sentiment Analysis

Google goes a step forward and recognize the view in the overall document containing the entities. Processing web pages, Google allocate score to each of the entities depending upon how they are used in the document.

5. Text Classification

A website classified into categories and subcategories and more subcategories in the large database. So that does the text classification. Google matches the nearest subcategory of webpages from the query.

For example, Design of a butterfly,

Google identify different subcategories such as modern art, digital art, artistic design, architecture, etc. and then choose the nearest matching sub category.

V. Conclusion

The keyword counting of a website can significantly affect the ranking of the page itself and its ranking Google search engine. This simply means that when Google is more appropriate it also shows if a particular keyword is searching but doesn't exactly match related data. Rankbrain algorithm available for use in machine learning Knowledge of graphs helps users enhance their experience of getting results. Fred seems to be the only one who has influenced many sites across the web, especially ones with ads featuring low-quality content focused on revenue generation. Hummingbird's results have diminished diversity in search results, and this is especially true for keywords that have the same meaning. Knowledge graph integrations may not be common, but they are becoming more and more specific. Google is getting better at understanding user search intent and creating quick and accurate search results. Pigeon is an algorithm which provide more useful, relevant, and accurate local search results more relevant to traditional web search ranking signals.

BERT is a common technology not only for NLP but also for Google search. However, it is important to understand that BERT does not change the way websites are ranked, but improves Google’s understanding of natural languages.

VI. References

[1] Survey on Search Engine Optimization Tools & Techniques Prashant Ankalkoti Department of MCA J N N College Of Engineering, Shimoga, Karnataka, India

[2] https://www.entrepreneur.com/article/229926

[3] Google Pigeon Update: Local Results Updated Again https://www.seroundtable.com/google-pigeon-update-18938.html

[4] https://www.bluecorona.com/faq/what-is-google-fred-algorithm-update/

[5] https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270

[6] https://en.wikipedia.org/wiki/RankBrain

[7] Search Engine Land Names new Google Local search update.

 ISBN: 978-81-946375-7-8

Dynamic Risk Scorecard for Privileged Users

Himanshu Narvekar
MCA, DES'S NMITD, University Of Mumbai

Riddhesh Jadhav
MCA, DES'S NMITD, University of Mumbai

Ms. Swapnali D. Mahadik
Assistant Professor, MCA Department, DES's NMITD, Mumbai, (Affiliated to University of Mumbai)

Abstract

In today's growing economy where the companies tend to increase their security measures towards protecting their valuable data but attackers simultaneously try to explore various methods for manipulating the security measures and thus bypassing them. So many companies tend to get help from third party companies for keeping their data protected from external attacks. Sometimes there is a possibility that the employees of the company may try to manipulate the information so the company uses Privileged Access Management (PAM) for protecting the data from insider attacks and monitoring the users.

Instead of using PAM, the users still might find a way to manipulate the information systems so there is a need to implement Risk Scorecard. This Research Paper will focus on displaying risky users to the organization with the help of Risk Scorecard so it will be easy to monitor the risky users for the safety of the information systems.

Keywords: Privileged Access Management (PAM), Privileged Identity Management (PIM), Risk Scorecard, Privileged Users, malicious activities, insider threats.

1. Introduction

Every IT environment consists of numerous devices, servers, valuable databases, systems which are all used or managed by many kind of privileged users which are powerful based on the permission given to them in an organization. The terms Privileged Users, Privileged Accounts and Privileged Identities are basically terms which hold the same meaning. Privileged users are users that are granted administrative privileges to the system. Basically, Privileged Users are a particular type of users that have special authority or permission in an IT organization like system id or admin accounts. These users differ from one another based on the rights or permission for performing security, administrative or system modifications. Privileged Users are one of the most common forms of user accounts which grants access on an enterprise domain, allowing users to have admin rights on, for example, their local desktops or across the systems they manage. Privileged Users have extremely powerful passwords which enable them to login into the system. If the passwords are weak and are not encrypted properly then it can potentially result in huge financial loss and reputational damage to any IT oriented organization. In an IT oriented organization if there are several unmanaged Privileged Users then it can result in devastating risks such as leakage of data, system abuse, financial loss, damage to the business continuity, customer loss and brand damage. Privileged Users hold elevated rights for accessing files, installing and running programs and configuring settings in an operating system, databases and system devices of an organization.

 ISBN: 978-81-946375-7-8

PAM stands for Privileged Access Management. It is used for managing Privileged Accounts/Users within an organization to ensure every action/activity of these users is auditable.

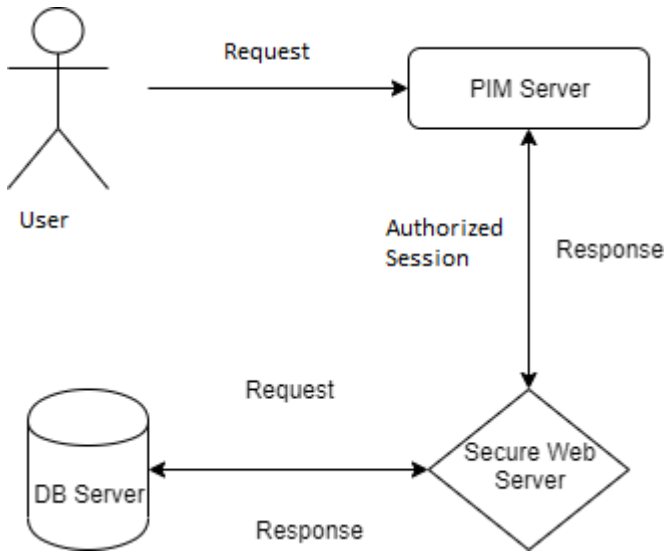


Fig. 1: Basic Implementation of PIM/PAM

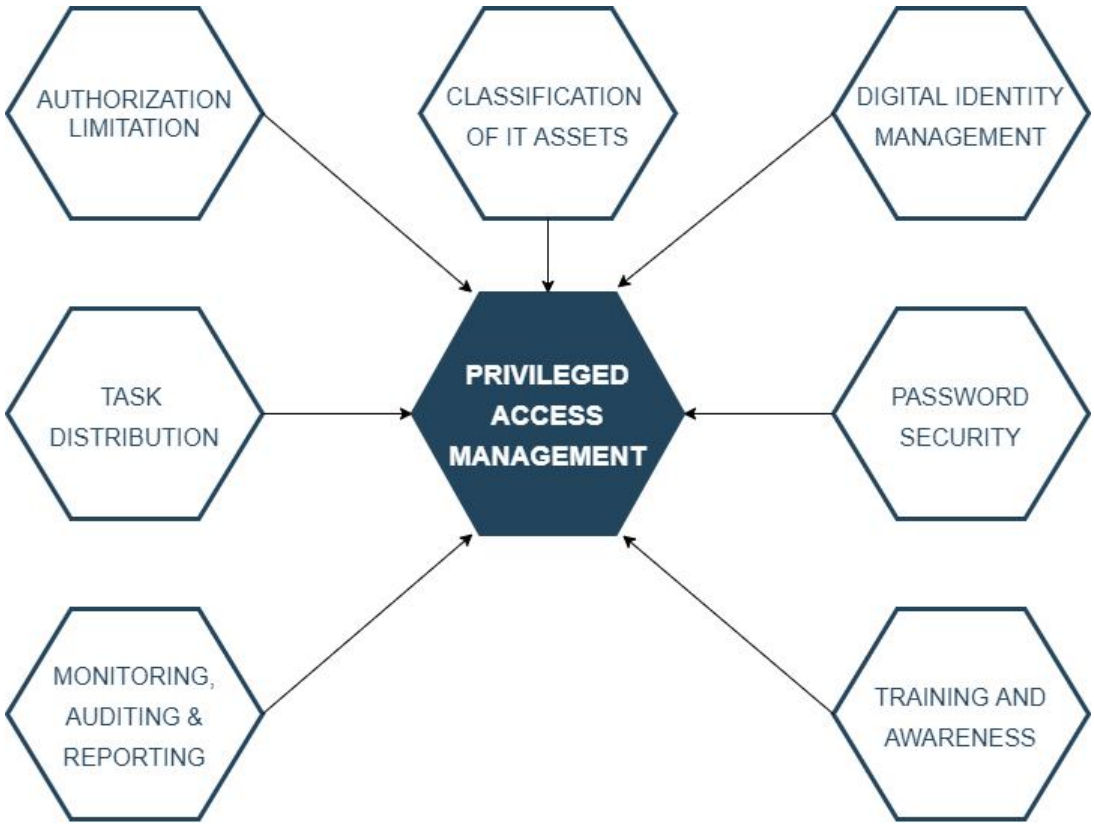


Fig. 2: Components of PAM

There are seven components of PAM. In Monitoring auditing and Reporting, it should be executed on a daily basis to keep a track of user activities so that a company can know if the

employees are performing proper procedures mentioned by the company protocols with insider threat and especially preventing tamper of privileged accounts, the behaviour of user's in information systems should be monitored, audited, and reported. Task Distribution should be not only made for employees of the organization but also for the IT executives who are monitoring the reports as they will easily get access to the file servers, databases, logs, services, etc. Every employee or IT executive should have different authorization access/limitations according to their ranking in the organization. In Classification of IT Assets, the devices should be assigned to a different group as every employee is assigned. IT assets consist of a list of devices used in a network with their IPs, ports and users. Password Security is a very important component as passwords are a collection of letters, characters, special characters and symbols to ensure safe and authorized entry onto a web-page, server etc. A password is a key to any confidential data and is personal to any human and hence is protected so that no unauthorized access is available to anyone. The main agenda of Training and Awareness is that users using the software or hardware should be aware of real life threats and their possible solutions to such threats. It tends to reduce misuse of technology and human error due to lack of proper knowledge. Digital identity management will help to identify the authorization limit of each user and will block access if a user does not have the authority to bypass it [1].

Privileged Users can be categorized into four types such as 'Common\Shared Administrative Users' which generally have super user rights and are commonly shared among the IT team (for example: Windows Administrator user, Linux root user), 'Privileged Personal Users' which are basically used by business users and IT personnel which can also be considered as super users by taking into consideration there higher level of privileges and their use of these privileges can affect the organization, 'Application Users' are generally users who can only have access to a specific application (for example: A .Net developer can only have access to .Net and Database application) for protecting the information, 'Emergency Users' are users which have certain privileges and are used only when a problem needs to be fixed immediately(for example: During disaster recovery) but to access these users we need approval of higher authorities of the organizations.

The proposed monitoring system which is based on Privileged Access Management (PAM) makes use of Remote Desktop Protocol (RDP). The term Remote Desktop Protocol refers to any software or any operating system which allows a personal computer's desktop environment to be run on any other system while being displayed on a separate client device. In the proposed monitoring system enables the user access the PAM server from a personal desktop and then access to any system where they want to carry on with their work. As there are several organizations who only lets the employee to perform on a certain application. This helps for protecting the data as the employee can work on a particular application. For example, an employee of a software company who is a coder is only given access to use particular programming software so that the employee won't be able to communicate with any external world.

The main problem in PAM is that there is a possibility that some insider may try to manipulate or tamper with the information system. The users are needed to be monitored for this the PAM monitoring system should grade the privileged users based on various conditions which will help the organization to know about the users which may pose a threat to them. This report will be generated in a graphical format on a daily basis by the product. This report is called as Risk Scorecard.

2. Literature Review

It is necessary to control and manage privileged users, which are one of the main reasons for insider attacks in organizations. There are various solutions provided by various researchers for this problem.

Sarkar [2] and Padayachee [3] have discussed in their studies why it is important to provide authorized users with as minimum privileges as possible for accessing the valuable information. Baracaldo and Joshi have developed templates regarding malicious use of privileges by the users. They have developed an algorithm based on the activities which they have addressed for the insider attacks by using their created templates [4]. Manideep, Brajendra and Yi have developed a model which focuses on the access of users to the company's databases. Their model generates a graphical interface which is based on Petri net on the basis of the activities which are performed by the authorized users in the company's database [5]. Agrafiotis, Nurse, Buckley, Legg have carried out survey on 120 insider attacks and have thus developed a template which is used for detection of insider attacks. The template is basically developed on the basis of the activities of privileged users according to their job description and malicious performance of the privileged users [6]. Bishop, Gates, Frincke and Greitzer have developed a model called Attribute Based Group Access Control (ABGAC) model which is similar to Access Control List (ACL) where they suggested to divide the user in different groups based on their job descriptions and every user group is given access to different resources located in the information system [7]. One study which was conducted by Kara Nance and Raffael Marty have designed a model which provides a graphical interface and analysis of the actions performed by the respective authorized users [8]. Lance has suggested a model which in based on honeypots i.e. he has implemented honeypots in his proposed system for preventing and detecting the authorized user from performing malicious activities [9].

As per the given studies which focused on detection of malicious activities or insider attacks performed by the privileged users, it has been seen that there are numerous solutions which include both human-oriented as well as technology-based approaches. As there are various studies which focus on preventing the insider attacks but there are no researches which focus on implementing the risk scorecard and thus monitoring the privileged users by rating them on the basis of reports.

3. Research Design

Various logs are generated based on the actions of any user who is part of the proposed monitoring system and are stored on the PAM server. Whenever any organization face any issue regarding manipulation of information system the super users or the admin users have to go through the list of reports in order to find about the user who has misused the information system of that organization. For this purpose, we need to implement a risk scorecard in the proposed monitoring system. The risk scorecard will be generated based on the following factors such as access to sensitive devices given to the user, number of devices accessed by a user, access to devices after during odd hours by the user, usage of sensitive commands by the user, entitlement to numbers of devices, failed login attempts by a user. On the basis of these factors a graph will be generated which will indicate the risky users to the organization and then the organization will monitor them.

	0-2	2-4	4-6	6-8	8<x
access to sensitive devices by the user [Type 1]	0	0.5	1	1.5	2
number of devices accessed by a user [Type 2]	0	0.5	1	1.5	2

access to devices after during odd hours by the user [Type 3]	0	0.5	1	1.5	2
usage of sensitive commands by the user [Type 4]	0	0.5	1	1.5	2
entitlement to numbers of devices [Type 5]	0	0.5	1	1.5	2
failed login attempts by a user [Type 6]	0	0.5	1	1.5	2

Table 1: The score will be given to the users based on the following conditions

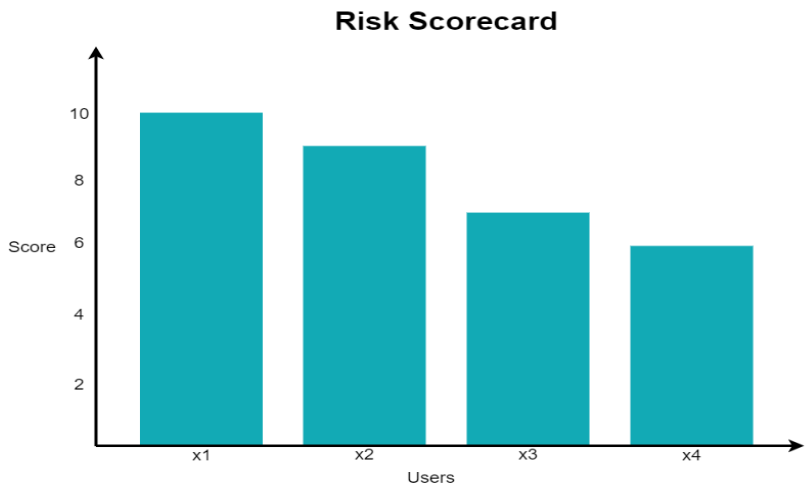


Fig. 3: Representation of Risk Scorecard for all the users.

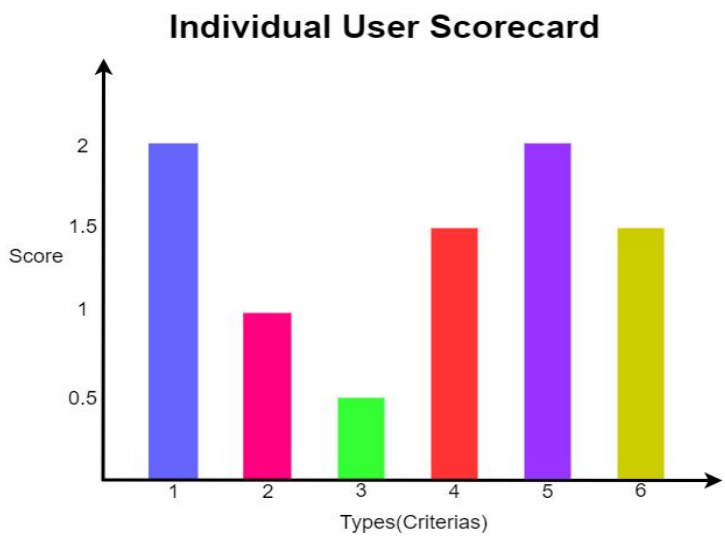


Fig. 4 Representation of different types of malicious activities performed by a certain user.

Every user is given rating on the basis on the following factors as mentioned in the above table. 'Type 1' refers to the sensitive devices that are accessed by the user i.e. the user has gotten access to the valuable information system which can affect the organization in both ways (positively and negatively). 'Type 2' refers to the number of devices accessed by a user. 'Type 3' refers to the devices which are accessed by the user during odd hours (i.e. after office hours). 'Type 4' refers to the usage of sensitive commands by the user (for example, various restricted

command that are used by the user to remove vital data from the system or commands which can be used to tamper the servers). ‘Type 5’ refers to entitlement to numbers of devices i.e. the number of devices which the user is already assigned to. The only difference between ‘Type 2’ and ‘Type 5’ is that ‘Type 2’ are the list of devices for which the user has to acquire rights from the higher authorities to access the devices which are not assigned to him. ‘Type 6’ refers to the failed login attempts made by a user for accessing a number of devices which are not under his domain and still the user wants to access them for retrieval of valuable data.

4. Implementation

This proposed model is supposed to generate a risk scorecard for privileged users based on 6 factors mentioned above. One of the factor is shown below.

Unauthorized Access | Jan/01/2020-Jan/31/2020 | Search for... | Search

Word | Excel | PDF

Device IP	Source IP	Computer Name	Login Time	Logged User	Alert Time	Notified To
10.137.176.4	10.137.166.172	INSVC-MUM01215	31-Jan-2020 17:12	CEINSVC-MUMPRD1\Magiceye	31-Jan-2020 17:12	Amol.Sawant.exe@sodexo.com,Durgesh.Mankar@sodexo.com,Minal.Shah@sodexo.com,Shekhar.BHOSLE@sodexo.com,Vikas.pawar@sodexo.com,bhaskar@swansol.com,deepak.prajapati@sodexo.com,poonam.bhingare@iraje.com,pravin.chavan@iraje.com,
10.137.176.4	10.137.166.172	INSVC-MUM01215	31-Jan-2020 17:10	CEINSVC-MUMPRD1\Crisis.Admin	31-Jan-2020 17:10	Amol.Sawant.exe@sodexo.com,Durgesh.Mankar@sodexo.com,Minal.Shah@sodexo.com,Shekhar.BHOSLE@sodexo.com,Vikas.pawar@sodexo.com,bhaskar@swansol.com,deepak.prajapati@sodexo.com,poonam.bhingare@iraje.com,pravin.chavan@iraje.com,
10.137.172.20	10.137.172.88	INSVC-BLR12019	31-Jan-2020 15:17	CE-SDXCORP\Itsupport.BLR.IN	31-Jan-2020 15:17	Amol.Sawant.exe@sodexo.com,Durgesh.Mankar@sodexo.com,Minal.Shah@sodexo.com,Shekhar.BHOSLE@sodexo.com,Vikas.pawar@sodexo.com,bhaskar@swansol.com,deepak.prajapati@sodexo.com,poonam.bhingare@iraje.com,pravin.chavan@iraje.com,
10.137.172.20	10.137.172.88	INSVC-BLR12019	31-Jan-2020 15:07	CE-SDXCORP\Itsupport.BLR.IN	31-Jan-2020 15:07	Amol.Sawant.exe@sodexo.com,Durgesh.Mankar@sodexo.com,Minal.Shah@sodexo.com,Shekhar.BHOSLE@sodexo.com,Vikas.pawar@sodexo.com,bhaskar@swansol.com,deepak.prajapati@sodexo.com,poonam.bhingare@iraje.com,pravin.chavan@iraje.com,

Fig. 4: Data of Unauthorized Access by Users in Iraje PIM Product [10]

	DeviceIP	Source IP	Computer Name	Login Time	Logged User	Alert Time	NotifiedTo
5	10.137.176.4	10.137.166.172	INSVC-MUM01215	31-Jan-2020 17:12	CEINSVC-MUMPRD1\Magiceye	31-Jan-2020 17:12	Amol.Sawant.exe@sodexo.com,Du
6	10.137.176.4	10.137.166.172	INSVC-MUM01215	31-Jan-2020 17:10	CEINSVC-MUMPRD1\Crisis.Admin	31-Jan-2020 17:10	Amol.Sawant.exe@sodexo.com,Du
7	10.137.172.20	10.137.172.88	INSVC-BLR12019	31-Jan-2020 15:17	CE-SDXCORP\Itsupport.BLR.IN	31-Jan-2020 15:17	Amol.Sawant.exe@sodexo.com,Du
8	10.137.172.20	10.137.172.88	INSVC-BLR12019	31-Jan-2020 15:07	CE-SDXCORP\Itsupport.BLR.IN	31-Jan-2020 15:07	Amol.Sawant.exe@sodexo.com,Du
9	10.137.172.20	10.137.172.88	INSVC-BLR12019	31-Jan-2020 14:44	CE-SDXCORP\Itsupport.BLR.IN	31-Jan-2020 14:44	Amol.Sawant.exe@sodexo.com,Du
10	10.137.172.20	10.137.172.88	INSVC-BLR12019	31-Jan-2020 11:09	CE-SDXCORP\Itsupport.BLR.IN	31-Jan-2020 11:09	Amol.Sawant.exe@sodexo.com,Du
11	10.137.170.20	10.137.170.34	INSVC-HYD12001	31-Jan-2020 10:00	CE-SDXCORP\Itsupport.HYD.IN	31-Jan-2020 10:00	Amol.Sawant.exe@sodexo.com,Du
12	10.137.170.20	10.137.170.34	INSVC-HYD12001	30-Jan-2020 10:13	CE-SDXCORP\Itsupport.HYD.IN	30-Jan-2020 10:13	Amol.Sawant.exe@sodexo.com,Du
13	10.137.172.20	10.137.172.88	INSVC-BLR12019	29-Jan-2020 13:32	CE-SDXCORP\Itsupport.BLR.IN	29-Jan-2020 13:32	Amol.Sawant.exe@sodexo.com,Du
14	10.137.170.20	10.137.170.34	INSVC-HYD12001	29-Jan-2020 09:58	CE-SDXCORP\Itsupport.HYD.IN	29-Jan-2020 09:58	Amol.Sawant.exe@sodexo.com,Du
15	10.137.172.20	10.137.172.88	INSVC-BLR12019	29-Jan-2020 09:34	CE-SDXCORP\Itsupport.BLR.IN	29-Jan-2020 09:34	Amol.Sawant.exe@sodexo.com,Du
16	10.137.170.20	10.137.170.34	INSVC-HYD12001	28-Jan-2020 17:45	CE-SDXCORP\Itsupport.HYD.IN	28-Jan-2020 17:45	Amol.Sawant.exe@sodexo.com,Du
17	10.137.170.20	10.137.170.34	INSVC-HYD12001	28-Jan-2020 09:57	CE-SDXCORP\Itsupport.HYD.IN	28-Jan-2020 09:57	Amol.Sawant.exe@sodexo.com,Du
18	10.137.190.52	10.137.166.29	INSVC-MUM01137	27-Jan-2020 17:41	CEINSVC-MUMWEB7\Administrator	27-Jan-2020 17:41	Durgesh.Mankar@sodexo.com,dee
19	10.137.172.20	10.137.172.88	INSVC-BLR12019	27-Jan-2020 16:52	CE-SDXCORP\Itsupport.BLR.IN	27-Jan-2020 16:52	Amol.Sawant.exe@sodexo.com,Du

Fig. 5: A Screenshot of the Data in Excel Sheet

As seen in the above Fig. 5 it shows unauthorized access of users who have tried to access the proposed model without proper access. This particular module will display the Device IP, Source IP, Computer name, login time, logged user, alert time (at what time the alert has been sent to the super admin), notified to (to which super admin the notification has been sent). Based on this data, a risk score card will be generated. The above data will act as an input to the proposed model.

5. Conclusion

The purpose of this research paper topic is to implement a risk score card in a particular organization to identify and monitor risky users under the PAM (Privileged Access Management). A risk score card will help any organization to easily recognize list of risky users in a graphical format. This score card will be generated on a daily basis and will show top 10 risky users of the organization. The organization can then take necessary actions against those users as per their protocols.

References

1. Erhan Sindiren and Bünyamin Ciylan: 'Privileged Account Management Approach for Preventing Insider Attacks' in International Journal of Computer Science and Networking, VOL.18 No.1, January 2018.
2. Roy Sarkar: 'Assessing insider threats to information security using technical behavioural and organisational measures' in Information Security Technical Report, 2010.
3. Keshnee Padayachee: 'An assessment of opportunity-reducing techniques in information security: An insider threat perspective' in Decision Support Systems, 2016.
4. Nathalie Baracaldo and James B. D. Joshi: 'An adaptive risk management and access control framework to mitigate insider threats' in Computers and Security, 2013.
5. Manideep Chagarlamudi, Brajendra Panda and Yi Hu: 'Insider Threat in Database Systems: Preventing Malicious User's Activities in Databases' in Sixth International Conference on Information Technology: Next Generation, 2009**, ISBN:** 978-1-4244-3770-2.
6. Ioannis Agarfiotis, Jason R. C. Nurse, Oliver Buckley and Phil Legg: 'Identifying attack patterns for insider threat detection' in Computer Fraud and Security, 2015.
7. Matt Bishop, Carrie Gates, Deb Frincke and Frank L. Greitzer: 'AZALIA: an A to Z assessment of the likelihood of insider attack' in IEEE Conference on Technologies for Homeland Security, 2009, **ISBN:** 978-1-4244-4178-5.
8. Kara Nance and Raffael Marty: 'Identifying and Visualizing the Malicious Insider Threat Using Bipartite Graphs' in *2011 44th Hawaii International Conference on System Sciences*, Kauai, HI, 2011, **ISBN:** 978-1-4244-9618-1.
9. Lance Spitzner: 'Honeypots: Catching the Insider Threat' in **19th Annual Computer Security Applications Conference, 2003, ISBN:** 0-7695-2041-3.
10. Iraje PIM Admin Manual, January 2019, Version 1.0.

 ISBN: 978-81-946375-7-8

AI meets AR: The Blend for New Innovations in Applications

Burhanuddin Harianawala
Student, Master of Computer Application, Navinchandra Mehta Institute of Technology and Development,Dadar (W)/University of Mumbai, Mumbai

Junaid Khalfay
Student, Master of Computer Application, Navinchandra Mehta Institute of Technology and Development,Dadar (W)/University of Mumbai, Mumbai

Abstract
A sub domain in Computer Science is Artificial Intelligence, is used to learn or train things& make machines work like humans. Augmented Reality is used to add or overlay digital content onto the physical world, as if they're actually there with you, in your own space. So combining both AI & AR, recognition of objects or things using AI and overlay graphical things upon that using AR. Recent research shows that how integrating the two technologies can be useful for many sectors like, entertainment, for education and in various other fields. This paper showcases how significant it is for integrating the two hot upcoming technologies together.

Keywords: Augmented Reality, Artificial Intelligence, AI Assistant, AI meets AR

I. Introduction

Artificial Intelligence is simulating human intelligence into machines so that machines are programmed to think and work as human beings. This will eliminate the long thinking cap and enhance problem-solving capabilities. Implementing AI in machines will help machines execute tasks be it simple or complex on its own without any human intervention. There is AI-based Face Recognition wherein the machine captures a face at first (training) then whenever the system detects a face in the camera it will search the database for the same and if match found it show results like attendance marked, etc.

Augmented Reality is imposing or integrating of information in digital form like pictures, videos, sounds or text, etc. onto the real world. AR helps in seamlessly blend virtual objects with the real world using the device's camera. The use of AR in entertainment and education field has grown as a result of advances in mobile technology over the years and development of AR application in mobile devices has risen.

Combining the two together will make objects recognition using AI faster than before (as earlier it was necessary to train AI for the objects and then use it) and can make things train on the go.

Example: Using the FaceApp mobile application, the faces are swapped in real-time, changing a person's age. AI and AR work together seamlessly to create immersive mobile experiences when rightly done. Another example is Google live maps (upcoming feature). It is available on ARCore (by Google) and ARKit (by Apple) enabled devices in locations where concept of Street View is present. With Live View, you can use augmented reality (AR) to better understand which way to go. Artificial Intelligence recognizes the location and Augmented reality shows arrows and directions which are placed in the real world to guide your way.

II. Literature Review

The literature review includes papers published in recent years since between 2017 to ensure that the information included was recent and relevant.

1. AR in Education

The value of AR is closely linked to the way in which it is designed, implemented, and integrated into formal and informal learning environments. [1]

2. AR for Training purpose

In the field of aviation and car manufacturing, it's been 40 years since head-up displays (HUDs) were implemented for the first time. Today, HUD devices are empowered by augmented reality. Widely used in today's military applications.HUD is a transparent display which is positioned directly in the fighter pilot's view. This provides more efficiency and secures conditions. [2]

3. AI and AR in Retail Sector

With the use of augmented reality in retail sector artificial intelligence is finding its way in physical retail area too. One of the famous furniture store IKEA has implemented the concept of AR. IKEA Place offered by IKEA , a unique mobile application that allow customers to see how a piece of furniture would look and fit in a given space.

ARKit is the platform for building IOS based application that supports 2D and 3D image recognition and image tracking. This technology allows users to move still photos of furniture into a still image of a room. The user first takes a photo of a room on the application, IKEA app has the AR capability which is able to measure the dimensions of the room and allow the user to drop an image of a piece of furniture into the photo from IKEA's item catalog. [2] In today's world all the sectors have developed and are developing by converting their manual processes to digital which makes it more secured. It is said that to compete in today's business world along with being financially strong you also need to be digitally strong.

Aisle411 is the company that claims it is developing AR technology for retail supermarkets which has successfully created a shopping application for Walgreens which integrates AI into AR technology to help shoppers find products more efficiently. This application is mobile friendly, which also runs on a tablet. The user can attach a tablet to his shopping cart and make use of Google Tango's computer vision-driven 3D-mapping service. Google Tango provides users the complete guidance in-store shopping using a map downloaded on the tablet. This provides customers to search product names, item numbers, and product categories by voice and command and gives them product's exact location.[3]

Lenskart Solutions Private Limited owns and operates an Online Shopping Portal for eyewear in India. They have a 3D Try On feature on their mobile application where in a customer can check some cool spectacles on the application itself before buying it. The 3D Try On operates as following:

- It learns your face shape with accurate scale using AI 3D face modeling.
- It digitizes the glasses photo-realistically.
- It renders the glasses onto your face model with proper and accurate fitting and scaling using AR.

For the above, Lenskart uses technology provided by DITTO (3D virtual try-on technology) which provides virtual try on and frame recommendation technology for eyewear retailers making the overall user experience better. They are also getting a decent number of customers on board via their mobile apps.

 ISBN: 978-81-946375-7-8

Figure 1: AI and AR implementation

III. Research Design

Augmented Reality increases the potential of buying products and coupling it with Artificial Intelligence can boost it to higher levels. This kind of implementation makes ecommerce attractive. Building components through image recognition in turn is becoming more and more important on the manufacturing plant to identify tools, objects and helps in product manufacturing. [4]

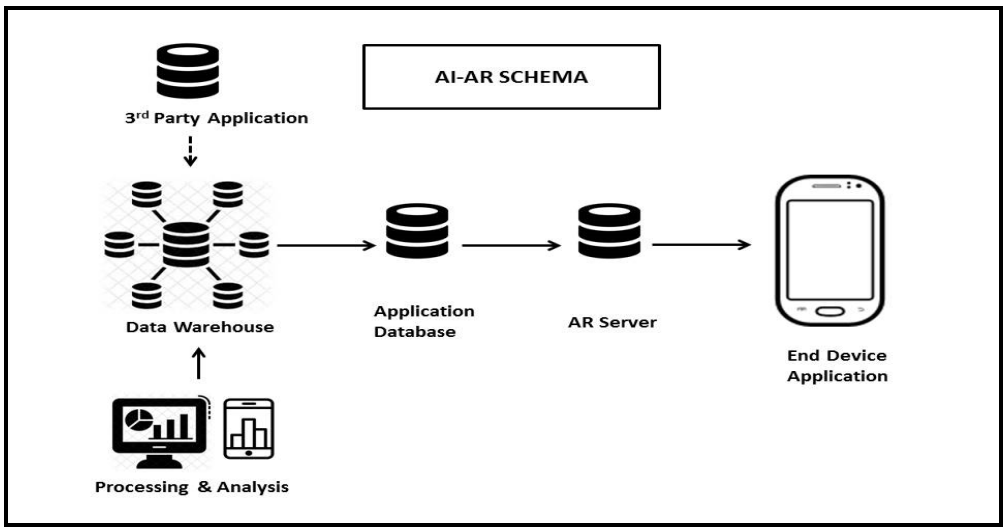


Figure 2: Diagrammatic representation of schema for AI and AR [5]

The above diagram shows the end application (Mobile Device) which displays information which is saved in the organization's Database. The information is originally stored in Organization's Database. The data is processed by Analysis Database on the while, were it gathers information by comparing with other similar products. This data is then sent to AR Database Server which stores this data and produces whenever needed to the user using an AR enabled mobile application on end devices. This way a user can get assisted by voice command via AI Assistant to find the product's information while using the AR application.

IV. Implementation

1) AI Based AR Assistant at Retail Stores

The idea of implementation is that where in a user can get assisted by voice command AI model to find the product's information while using the AR application. This is using

 ISBN: 978-81-946375-7-8

Augmented Reality Assistant through AI. The AI-based Augmented Reality assistant will be an advanced approach that integrates AI into AR technology. The AI model will assist the customer/s in a mall or store to any product information like its details, price, ratings, etc. either by speech or by written text over the end device integrated.

Image 1: AI-Based Augmented Reality Assistant Demo

[AI-Based Augmented Reality Assistant Demo - https://www.youtube.com/watch?v=7CzdA5sEC7U]

2) AI using AR based on a Magic Application for Mobile devices.

Effect: You ask your spectator to hold a card. The spectator doesn't know the card he is holding onto. You take a picture with your cell-phone. You put the card back into the deck. All the time the spectator has no idea which card he was holding onto. Then you ask your spectator to freely think of any playing card from the deck. Let's suppose he says "the Eight of Clubs" **(8 ♣)**.

Now comes the best part of this effect, you reveal to the spectator that photo you clicked before has the spectator holding the thought of the card - Eight of Clubs!

Routine: You will click a picture of the spectator holding a random playing card (see picture below) and set the red outline border on the center of the playing card. It will get trained and then you can overlay the spectator thought of the card over it. This is achieved by AI learning the card on the go, like the center of the card design and the borders. Then user enters the card details (secretly, without spectator knowing it) and then after that, AR helps in overlaying the spectator's card on top of it using AR functionality.

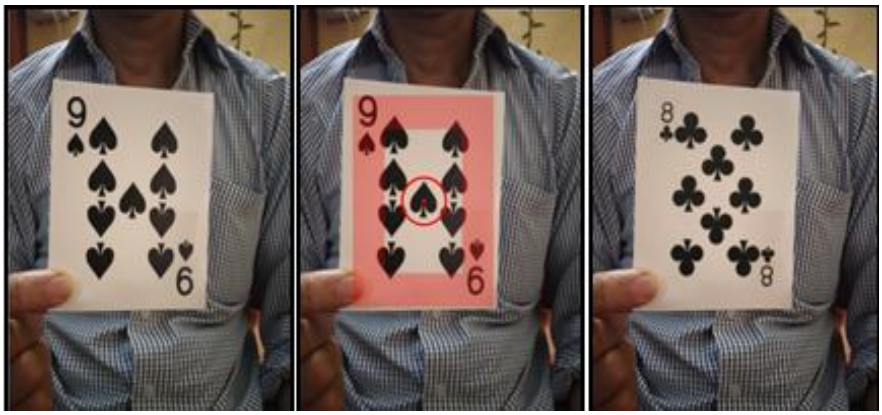


Image 2: AI and AR Implementation on mobile app

 ISBN: 978-81-946375-7-8

3) Skin Rash Diagnosis System [6]

Technology has a vast impact on IT sector but along with IT sector there is a vast impact of technology in healthcare sector too which can prevent and cure any disease if its symptoms are detected at early stage. If symptoms are detected earlier one can have have the complete chance to cure the person as soon as possible. An early detection of disease is a sign of relief for the person suffering from the diseases and for his dear ones too. A skin disease is a disease which can affect your image too. Now we can cure this type of skin disease with a blend of augmented reality and artificial intelligence. Blending these two robust technologies will prove to be a boon to the early diagnosis system too.

A patient will always have a fear if he is suffering from any skin diseases and its natural because any small reaction on skin can bring the patient into panic state especially in rural areas where you do not have much facilities. Think of a mobile application which can do medical diagnosis where predictive analysis can be done using AI and superimposition technique of AR. AR will let you know about the intensity of the skin rash and AI will tell you the further steps to cure it. With this application you can have preventive results.

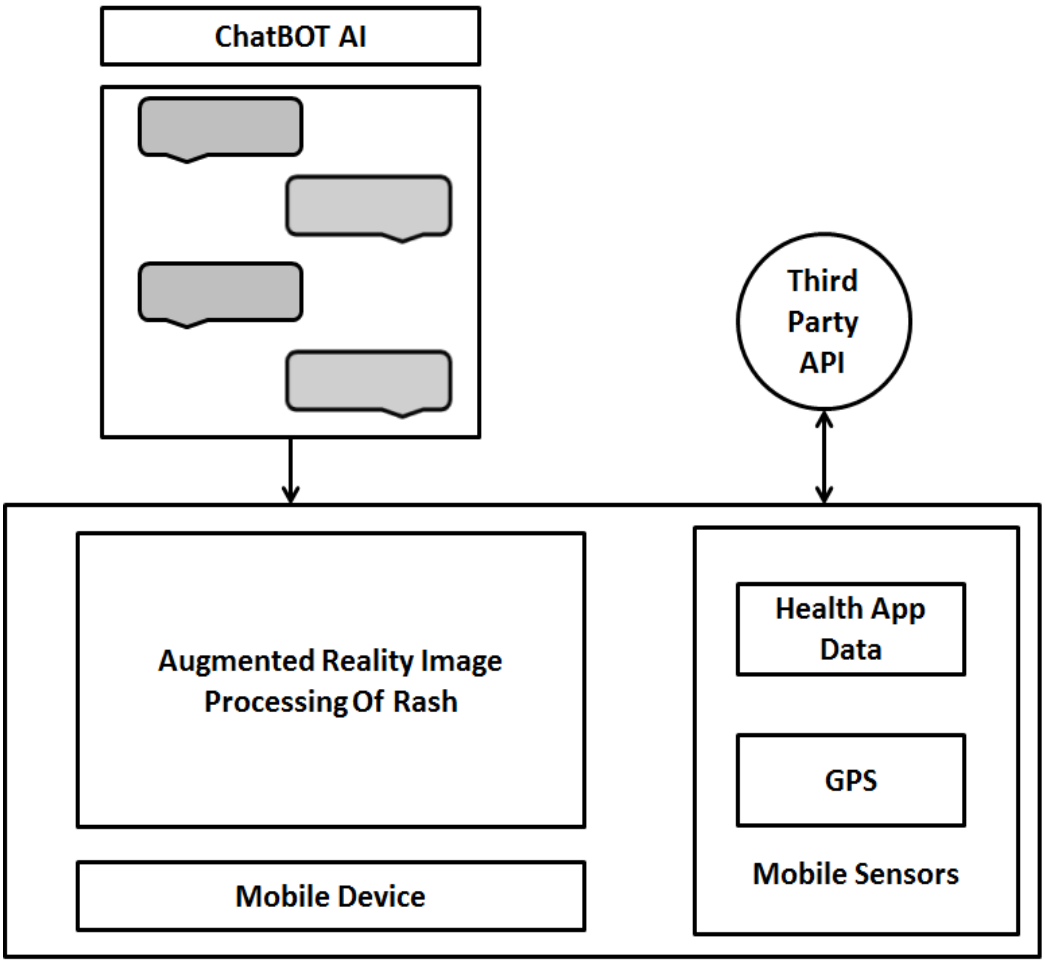


Figure 3: ChatBot AI and Augmented Reality Image Processing Of Rash

System Components

1. ChatBOT AI

A conversation via textual or auditory methods is done using it. It is designed in a human conversation pattern where few questions are asked to user about their skin problems. Based on the conversation diagnosis suggestions are given.

2. Augmented Reality

Additional details about the skin disease are collected by image processing using AR. The details are compared with the database of the system and suggestions are provided through ChatBOT.

3. Mobile Sensors

Patient's location is traced using the mobile sensors. This application also helps in providing nearest doctor address by comparing the lat long through the database of system via the third party applications. Details about the patient's condition are forwarded to the doctor that is suggested to the patient through the ChatBOT.

V. Conclusion

AI & AR when both combined together in business can provide a better marketing tool that can help gather customer's information like his preferences and in-turn let companies learn more about customer's preferences. As the technology is still new, the retailers can implement this and this can help customer/s to get a digital representation of product before purchasing it; product information, comparison (if possible), and ratings about what they are purchasing and if possible the customer gets preferences, the customer gets recommendations about the related products.

This research paper objective was to understand the impact of an AI & AR in mobile application on daily use for the exploration purpose. This literature also indicated that there is insufficient research on the impact of using both AI & AR in many domains, and there is room to explore the potential of AI & AR (as combined) in order to explore more domains & to improve learning and contribute to other literatures ahead.

VI. Future Scope

AI & AR technology has much more to offer to the industry domain than just the retail or entertainment. It will be very helpful for travel industry for the purpose of tourism; teaching for students and much more. Once the technology starts blooming, the possibilities will be endless.

VII. References

1. Tasneem Khan, Kevin Johnston, and Jacques Ophoff, "*The Impact of an Augmented Reality Application on Learning Motivation of Students*". Hindawi. Advances in Human-Computer Interaction Volume 2019, Article ID 7208494. *https://doi.org/10.1155/2019/7208494*

2. Burhanuddin Harianawala, Saifuddin Harianawala, Junaid Khalfay, "*Reality of Interactive World*". IOSR Journal of Engineering (IOSRJEN). ISSN (e): 2250-3021, ISSN (p): 2278-8719. *https://www.iosrjen.org/Papers/Conf.19015-2019/11.%2063-69.pdf*

3. Augmented Reality Shopping and Artificial Intelligence – Near-Term Applications, *https://emerj.com/ai-sector-overviews/augmented-reality-shopping-and-artificial-intelligence/*

4. IEEE Future Directions: The near Future of Augmented Reality – V, *https://cmte.ieee.org/futuredirections/2019/08/15/the-near-future-of-augmented-reality-v/*

 ISBN: 978-81-946375-7-8

5. What is the difference between AI and AR? *https://www.quora.com/What-is-the-difference-between-AI-and-AR*

6. Avadhanula, Alekhya & Morale, Prajakta & Shah, Pooja & Kasetty Sudarshan, Sneha – *"Artificial Intelligence and Augmented Reality based Skin Rash Diagnosis System"*. San Jose State University. *https://www.researchgate.net/publication/320703991*

 ISBN: 978-81-946375-7-8

Use of Piezoelectric Technology for Charging Electric Cars

Rasika Ramesh Kerkar
MCA, DES'S NMITD, University of Mumbai

Gayatri Manoj Kulkarni
MCA, DES'S NMITD, University of Mumbai

Abstract
According to research it is estimated that more than 1 Billion vehicles are running across the streets worldwide. As the technology and development is growing at rapid rate it is assumed that by 2036 we can expect to see more than 2.8 billion vehicles on planet and half of them would run on electricity rather than using traditional fuel sources. In this research paper is a review of a model which uses the concept of piezoelectric generators embedded in the tier which captures and converts the vibrations caused by vehicles into useful energy which then can be used to charge electric cars while they are running. This would save considerable amount of time which is needed to charge electric vehicles at charging points.

Keywords: Electric car, Vibrating energy, Piezoelectric devices, springs, Electric Energy, Piezo cells, vibration energy

I. Introduction

The vehicles that run on electricity are known as an Electric Vehicle also called as EV. EV's get their power from the rechargeable batteries installed inside the car. These batteries are not only used to power the car but also used for the functioning of lights and wipers as a result the batteries run out quickly.

Compared to the vehicles that only require few minutes to fill gas at the gas station, recharging your EV requires much more significant time investment. While most of the EV's can be fully charged in about four hours, some take a whopping 15 to 20 hours.

Electric charging stations are still under the development stages. Not all places where you go to on daily basis have the electric vehicle charging stations, meaning that if you're on a long trip and run out of a charge, you might get stuck in the middle of the road.

The solution for these drawbacks could be installing large batteries and faster charging stations but the problem still exist The best solution for this can be the use of a piezoelectric array mounted on the tires of the vehicle that will charge the vehicles while they are running .This will eliminate the extra time required to charge the vehicle as the vehicles can be charged while they are running on the roads. Even if the batteries run out of charge it can be easily charged anywhere and at any time.

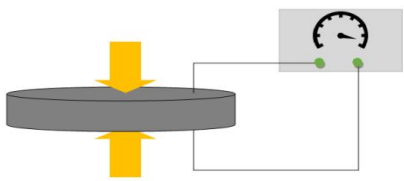


Fig.2.1 Phenomenon Of Energy Generation

Piezoelectric Materials:

Piezoelectricity, also called the piezoelectric effect, is that the ability of certain materials of generating an AC voltage when subjected to mechanical stress or vibration. The most commonly used piezoelectric material is quartz[1]

The piezoelectric effect occurs when the charge balance within a material's crystal matrix is disturbed. When there's no applied stress on the material, the positive and negative charges are evenly distributed so there's no potential difference. When the lattice is modified slightly, the charge imbalance creates a difference. This current is extremely small and would only cause a little electric shock.

The piezoelectric material through the mechanical stress like load from humans foot, vehicles, railway tracks, dance floors, etc. creates vibrations through spring which is produced on positive and negative charge centre's shift which results in an external electrical field .[5] For such conversion to take place, mechanical strain generates mechanical energy which is then converted into AC voltage and then again converted to DC voltage. Thus electronic charge gets generated in response to mechanical strain applied. When the crystal is strained, polarization takes place. The amount of polarization and the stress applied are directly proportional to each other. And therefore the piezoelectric crystals are used to harvest the energy produced by the vibrations from traffic on the road or moving vehicles on highways.[5]

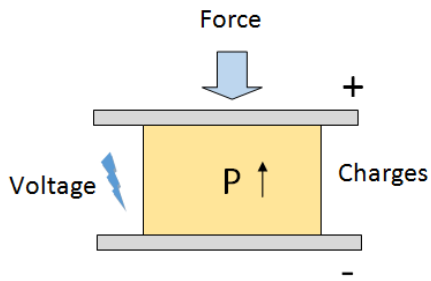


Fig.2.2 Generation Of Piezoelectricity[2]

II. Literature View

Vijaylaxmi Kalyani, Shruti Lohiya, Kritika Gupta, In this paper authors are focusing on charging electric cars via vibrating energy which totally replaced electricity and made use of vibrations present around us which are harvested economically.[4]

Noaman Makki and Remon Pop-Iliev , in this paper they have discuss the use of piezoelectric materials, both PZT and PVDF, within a commercial vehicle's pneumatic t ire to harvest power that can be used for power sensors or even run onboard devices. Differen t ways of harvesting energy using highly bendable piezoelectric elements, both PVDF and PZT, are explored. Three (3) energy harvesting are explored, with each relying on the deformation of tire's Treadwall and Sidewalls due to vehicle weight acting on it. These methods are comp ared on the basis of their power production capacity as well as other significant factors for u se within the tire. [5]

K. Anil, Neriyanuru Sreekanth , in this paper the authors are reviewing about the power generation using piezoelectric pulse generator and discuss the use of piezoelectric material with in a commercial vehicles pneumatic tire to harvest power that can be used to run onboard devices or to recharge the electric vehicle with the matlab/simulink simulation results . Power harvesting in two types of vehicle tires are explored in bike and car, with each having a different normal forces acting on the contact area of tire due to the vehicle weight acting on it.[7]

III. Research Design

A piezoelectric array is mounted inside a tire of a vehicle and is flexed or distorted during each revolution of the tire. Due to the inherent nature of piezoelectric devices, distortion generates an electron flow. The outputs of the piezoelectric devices are connected to an electric circuit to transfer the high voltage, low amperage electricity produced by the piezoelectric devices into low voltage, high amperage electricity that is compatible with the electrical system of motor vehicles.

Mechanical energy is converted into electrical energy by the piezoelectric materials, and vice versa. For this, inner liners of the tires are bonded with piezoelectric units. When wheels rotate, the tire continuously distorts and slackens as and when they interact with the road. Because of this the periodic voltage with a speed-dependent frequency is produced by the piezoelectric units. The energy generated, can be used as auxiliary power to increase the life of batteries.

Each piezoelectric vehicle tire is equipped with an array of highly bendable PZT (lead (Pb), zirconia (Zr), Titanate (Ti)) benders covering entire inner surface of the tier.[3] The output is in an AC waveform, because of this deformation of tires, has to be converted into DC signal before it can be stored into a capacitor. Each row of benders, running across the width of tire is taken as single generator and will be rectified separately with all PZT lines connected in parallel. At any given time while the vehicle is running, only 3 or 4 rows depending on the length of the arrangement of the contact patch power will be generated. The use of sensors within the wheel eliminates the need to rely on a limited power source such as a battery hence allowing an increased usable sensor life.[3]

IV. Implementation

In the model, when the tiers are not moving it will maintain its circular shape. Movement of the vehicle causes a part of the tire to conform to the shape of the road it flattens out. The sidewall just above the contact area also undergoes deformation by bulging out. When the vehicle moves, a new section of the tire continually deforms and relaxes in a cyclic pattern whose frequency depends on the speed of the vehicle. The piezoelectric bender deforms and relaxes with the tire and this deformation will allows collection of energy.[3]

To send this energy to the battery, it uses a connection running from the tire to the hub and then hub to the battery or the energy can be transferred with an induction current. Power is obtained from the tire using an assembly that constantly maintains the electrical contact between the chassis and the wheel while letting the wheel spin freely. Such a setup allows the extraction of tire-generated power to run the onboard electronics. The energy obtained by bonding piezo benders depends on three factors—tire surface area, rpm at 60 mph, and bender end-to-end deflection.[3] These factors are a function of tire radius with a bigger radius leading to a bigger area , lower rpm, and lower deflection. Different tire sizes will produce a different power output depending on the combination of these factors.
The future of advanced driver assistance systems will be contingent on the vehicle making all its decisions based on the assumption that the tires can withstand it all. And maybe they can, if the tires are factory produced, and haven't worn out the tread. This calls for sensing construction far more advanced than that of a tire pressure monitoring system, bringing us into the scope of advanced tire intelligence. The sensors needed for such intelligence would require power .[3]

These sensors would need some special characteristics, such as being able to operate on miniscule pulses of power, and the ability to be powered by the motion and pressure of the tires and the vehicle's interaction with the road.

The amount of pressure that we are dealing with is only transitory, since the vehicle is a moving object. The piezoelectric energy harvesting will need to be done just under the tread. A crystalline layer lies between two different tread layers would serve the purpose. The batteries in an electric vehicle can be self-charged as the car is in motion, due to the power generated by the tires. Therefore, only a one-time external charging will be required. This is exactly what is needed when you're in the middle of nowhere and your car might not reach the next charging station.

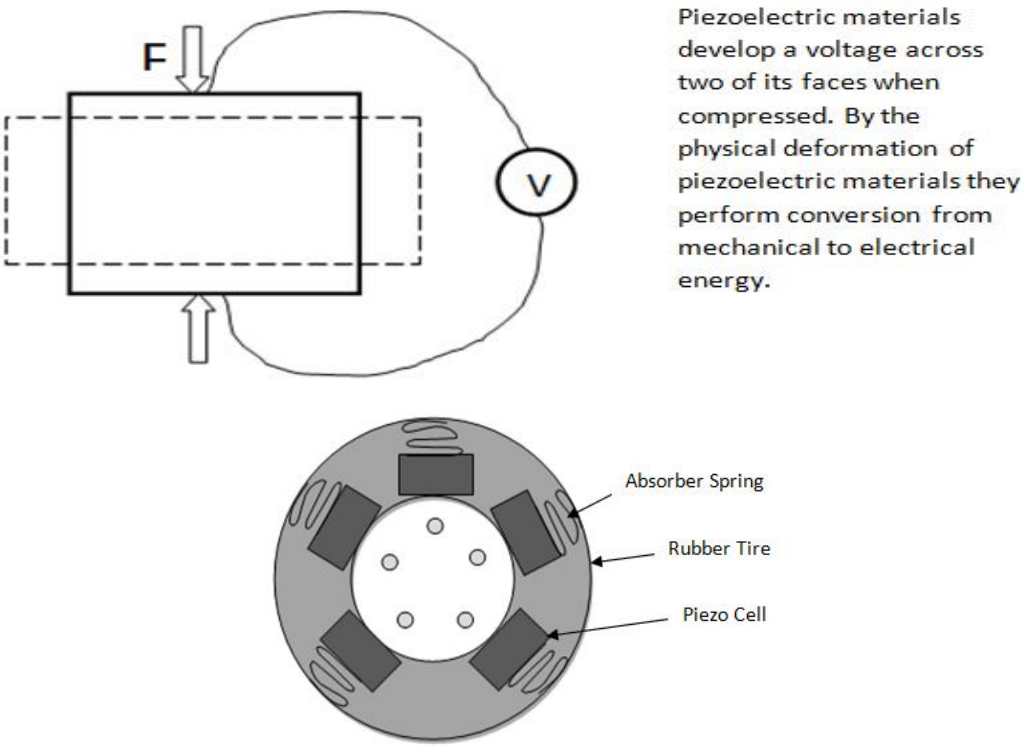


Fig.4.1 Structure of Piezo cells integrated Car's Wheel

V. Application

The Power can be stored somewhere and it can be used to charge small mobile devices or digital watches.

VI. Conclusion

In this Research paper we have tried to gain attention towards a new Technological way which can make use of unused energy generated by cars and piezoelectric materials to charge electric vehicles. The technology is based on the concept of energy which is usually considered as waste or error or unused.

As we know that electrical energy is non renewable source of energy there is huge need to save and use this energy efficiently. Thus by using this vibration energy along with piezoelectric materials we can save great amount of energy which would otherwise go waste. The saved electricity can used to light up many houses in rural areas or small villages.

The Piezoelectric material has the great ability to sense the mechanical strain and transform it into mechanical energy which is converted to electrical energy. The amount of energy produced is dependent on the number of piezoelectric elements on tier. Then this produced energy is stored in batteries. This idea is very cost effective and it can be used anywhere easily to generate and store electricity for future use.

 ISBN: 978-81-946375-7-8

VII. References

1. Piezoelectric_power_generation_in_automotive-tires. https://www.altenergymag.com/article/ 2017/12/1-article-for-2018-piezoelectric-power-generation-in-automotive-tires/27642/
2. Piezoelectric Materials: https://www.iaeme.com/MasterAdmin/uploadfolder/IJMET_07_02_009/IJMET_07_02_009.pdf
3. Piezoelectric-mechanism. https://www.altenergymag.com/article/2017/12/piezoelectric-power-generation-in-auto motive-tires/27642/
4. Vijaylaxmi Kalyani, Shruti Lohiya, Kritika Gupta, "Electric Car Charged By Vibration Energy: A Future Technology", Journal of Management Engineering and Information Technology (JMEIT). https://www.slideshare.net/iaeme/use-of-vibration-energy-for-charging-electric-car
5. Noaman Makki and Remon Pop-Iliev , "PIEZOELECTRIC POWER GENERATION IN AUTOMOTIVE TIRES", International Workshop SMART MATERIALS, STRUCTURES & NDT in AEROSPACE. https://pdfs.semanticscholar.org/aa97/69bab50324192b12a1ec7e406c 6cdf49c6e9.pdf
6. Ghaithaa Manla, Neil M. White, John Tudor, Harvesting energy from vehicle wheels TRANSDUCERS 2009 - 2009 International Solid-State Sensors, Actuators and Microsystems Conference
7. K. Anil, Neriyanuru Sreekanth , "Piezoelectric power generation in tires" 2014

 ISBN: 978-81-946375-7-8

Review of IOT in Automobile

Nikhil Singh and Abhishek Pandey
Students, Masters of Computer Applications, Navinchandra Mehta Institute of Technology and Development, Dadar (W), Mumbai

Abstract
Due to accidents major public Injuries occurrence has been increased nowadays. For maintaining a successful healthcare for the bikers one of the important confrontations public awareness and protection as well as safety. One among the increasing fatality rates for the previous year is bike accidents in highway and motorways. In this research paper the initiative which are taken in order to reduce the possibility of people dying due to road accident. In this research paper the the progress of two wheeler using framework by the sensors that will monitors and keep the track and as well as control the speed.

In this research paper, the sensors which are used is the speed sensor which will control the speed. The other sensor is the ultrasonic sensors which is used in this research paper which will be used to maintain the distance between the bike which will avoid collision between the bikes. With the help of this type of sensors the rider will get an notification when the bike speed exceed over 90km/hr. All the sensor and other devices which are connected in speedometer of the bike and also the gear box of the bike are controlled by the microcontroller. The notification light starts blinking in speedometer when the speed goes above 90km/hr. So in this way the accidents, collision and the rate of death will decrease. The main aim of this research paper is making bike which can help the bike rider aware about the speed control, distance between two vehicle in order to avoid road accidents.

Keywords: Prediction and detection of collision, sensors & ultrasonic sensors and speedometer.

1. Introduction

The collision of bike and vehicles often result in occurrence of death by accident to vulnerable bikers. Such scenario can be protected or avoided with the use of technologies. The next generation bikers with the capability of sensing, computing and communication has the potential to be integrated in Internet of thing environment. On the other hand, IOT is getting a better boost from various automobile industry. The positive feedback and reviews have led to brainstorming of number of IOT based industries to come up with number of different ideas and planning.[2] In consideration of IOT in automobile the architecture, framework and designing has also evolved. IOT is integrated with a mind set to avoid bike-vehicle collisions, offer bike safety and comfort level to user. Moreover, it leads to create a biker's safety management framework which integrates misbehavior detection scheme (MDS), collision prediction and detection scheme (CPD).[4]

The sensors, actuators and technologies in IOT, used towards building the autonomous Vehicles are examined in this paper. The stage of evolution is implemented from old normal bikes to fully interactive and responsive to user.

In the fast running life of human many of them prefer bike in order to save themselves from traffic. But Bike accidents have been increased across in which lots of humans lost their lives which are difficult to measure. So to reduce the number of bike accidents we have to implement Internet of thing in bike so in order to operate the bike using number of sensors. The vision of Internet of thing involved in transportation and bike was evolved long ago and was a part of

future implementation. From Last two year the internet of thing technology is been implemented in real world.

It is possible to make advanced the two wheelers system by the sensor that monitors and control the speed. The sensor used in this research paper to control the speed is the speed sensor. Here ultrasonic sensor is also placed in order to maintain the distance between the vehicles to avoid collision between the vehicles.[7] The riders will be given indication when the control of the bike exceeds the particular limit.

The Whole devices and sensors connected the speedometer and in gear box is controlled by microcontroller used in the vehicle. If the speed goes above 90km/hr the buzzer alarms. With the help of this death, collisions and accidents can be avoided. The main aim is to check whether or not the person who is riding the bike is aware or not about the speed , distance control between the vehicles in order to avoid collision. The bike is always in secured network as the emergency help is connected directly to the system.[5]

2. Literature Review

The Government has legally taken the field of "IOT(Internet of Things)" as strategic research paper for nation, announcing that it is a digital revolutions for leading the country and achieving the master plan . The authorities has a promotional approach of reinforcing the competitiveness in software sensor and its utility primarily based factor and gadgets. IoT basically refers to the net environment where human beings and machines all are connected to the guided and unguided network for you to together create, collect, utilize and share records and services that consists of sensors i.e. Input records, devices for acquisition, Clyde sharing of records i.e utilization for software.[6]

The future scope of IoT, that allows you to transform any actual world object into intelligent digital item. The IoT ambitions to unify the whole thing on this world below a not unusual infrastructure; giving us not simplest controls of the things around us, however also keeping us informed of that the state of the matters. In this discussion, that we've got presented a take a look at that addresses IoT concepts through systematic evaluate of company white papers, scholarly research papers, professional discussions with specialists and on-line databases.

Moreover, this studies article makes a specialty of definitions, geneses, fundamental requirements, traits and aliases of Internet of Things.. The IoT is an innovative idea in order to remodel any real world object into intelligent virtual objects within the future. It allows consumer to identify the whole lot on this world uniquely, take manipulate over identified the things (e.g. Door automatic Locks, Microwave, sensor Lights, Television, Coffee making machine etc.) and keep knowledge about state of the things. This time period IoT describes several technology and studies disciplines which tells that the internet has reached to every real world physical items in the world.[6]

Everyone ought to be aware of the protection measures and security mechanisms which are to be followed in the roadways. The number of highway accident rate is very high and is increasing consistent with every years when compared to other hazards which is unpredictable. Survey says that human beings waste more than million hours in traffic which is also major cause to the prevalence of injuries and hazards. Two automobiles same time at a identical point using at the same range of speed will purpose anyone of the automobile to be collide to each other. In order to avoid this form of collision two automobile must communicate with every other approximately the directions, velocity and other information approximately the path they are travelling. So that there will be a possibility of reducing collisions between the automobiles.

Wearing helmet is the most essential and obligatory for the riders. Many deaths and disabilities AND injuries are due to no longer carrying the helmets. In this mission helmet is made smart for the bikers. If the person who is driving the bike ought to put on the helmet. If he is no longer carrying the helmet the engine will now not start till he wears the helmet. The sensor placed inside the helmet is hooked up with the clever smartphone which identifies the area of the rider and tell to his/her family if accident passed off. The LED mild within the helmet glows if he goes to apply BREAK or he turns left or right. Thus this assignment is proposed. Here on this assignment collision can be controlled but no longer completely but partially, it informs to their recognized ones after the accident came about . The research paper completely aims to make riders ride safer and make the ride simple for the rider by guiding them the route.

For the welfare of transportation system IOT related devices are being build which can be connected to smart phones. Mostly the economic people uses their bike as a mode of transportation so it should be safely and securely maintained . So for secure mode of travelling GPS and various other sensors are used to detect the speed of the vehicle. The Speedometer of bike are made which can connect with smart phones using Bluetooth . The speedometers are made fully Digital console and user interactions are made easy to use it. The Sensors and actuators help in detecting the Best Lap and the last Lap of the bike travelled . The Speedometer once connected to the smart phone then it display the Phone signal strength and Phone's Battery strength . The smart phone also helps in auto syncing the clock of bike accordingly. One more feature that is added is the user can find its last parked location of bike as well as the current parked location of the bike at any location. Trip report generation is also possible which indicates how much fuel was used and time taken to travel as well as the top speed driven at the time of travelling. The accidents are usually occurred when the rider gets distracted when they receive a call or a message so using IOT auto reply message can be setup so that the rider can enjoy their ride with no disturbance.

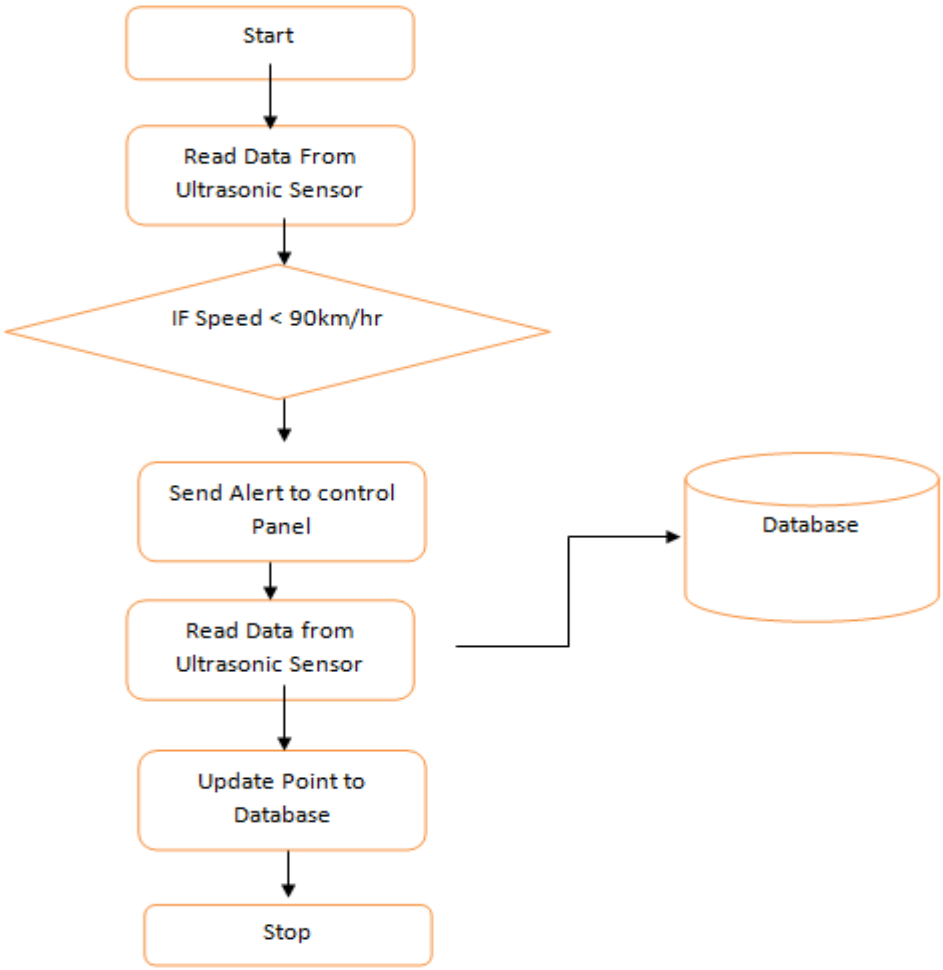
TVS NTORQ 125 is the TVS company's first scooter which have Bluetooth support and smart connectivity features which has brought an idea for making a ride secure and simple for riders at a low cost compared to other scooters which have IOT integrated in it. The Unique feature of TVS NTORQ is that it comes with company's new SmartXonnect Platform technology to make it smart features to support.[3] The SmartXonnect platform allows various features like Display phone battery strength, Top Speed recorder, navigation, give alert for service reminder, There is also a feature of DND mode in which the Application automatically sends an SMS that the rider is driving the bike. TVS NTORQ user needs to download the android application provided by the bike company providers. The minimum requirement for running the application on smartphones is Android 4.4 KITKAT and above.

It is too much critical to make bike secured Bike being important thing in day to day life Awareness and safety should be an important aspect to be placed first in order to have effective welfare of bikers.[8] Riding a bike that has more or less featuring of just riding has become the cause of accidents resulting to death and disability. It is been observed the graph of the accidents on road and highways are certainly increasing on a higher node which is not a acceptable. The desire of research paper is to eliminate and prevent the root cause of accidents and to bring the bikers on a comfort and a secure path. The requirements brought news idea of implementing automation and IOT in bike which in turn brought the package of security and safety that can be implemented and brought into practice. Therefore, advanced automobile tools such as ultrasonic sensor is placed which plays the functionality of avoiding collision between the vehicles, connectivity of mobile phones that is easily connected and displayed on vehicle screen and user responsive, indication of danger when bike has exceeded or about to exceed a particular limit. the a threshold value is set with respect to the vehicle capability. These systems

 ISBN: 978-81-946375-7-8

could benefit riders by minimizing their distraction on road. And sensors will alert us when some obstacle is there through feedback signals delivered to the smart phone. In other words, it alerts when there's a car or another bike coming up from behind you, and lets you know with notification so you don't have to look away from the road ahead and calculates all the points where the distance between our bike and vehicle on the road is minimum than normal distance and is send to smart phone.

Fig 1: Data flow diagram for Over Speed Detection[6]



The above is the flow chart that represents data flow for over speed detection. The over speeding can be termed as one of the root cause or the important factor for accident. Hence initiatives are taken to represent it functionally and eliminate the cause of it. The flow chart mainly includes the role of ultrasonic sensor, throwing alert and database. This functionality gets activated as the bike starts along with the start the sensors are automatically loaded. The user is driving and at point he certainly exceeds the average speed or a speed that can be prone to accidents, this were the alert signal is activated. This activity of monitoring the data i.e. the speed is dynamic and quick. The alert is given to control panel which in turn read it from the ultrasonic sensor. This data or the point is updated to database were it gets classified and delivers the further action to be implemented.

The output that is delivered is based on condition that is implemented in the backend. The code in the backend is designed in such a manner that it calculates the speeding continuously and

with respect to that the alert is calculated. Suppose the user is riding at a speed that touches 90km/hr and comes back to less than that in such case the alert is not activated, if the biker is continuously riding at the top speed in that case the alert is activated and popped up to the screen.

Another alert that is included is the distance that needs to be maintained between the vehicles well this too is a important factor that needs to be included in the alert box. This alert indicates the biker to slow down and maintain a speed lower than current. Normally the distance maintained by the vehicles should be 10 meters, with respect to the road environment offered in India there is some flexibility maintained in backend so if the distances among the motors comes under 10 or 7 meters then the indication might be given to the rider who rides the bike.This helps the biker to take the necessary action in order to prevent collision or accident. This mainly results to the user applies a controlled brake to low down the speed or the user may change the lane of road. Suppose the user is driving at a high speed and applies a sudden brake this may results to losing the balance or something bad may occur hence the alert is activated at a minimum distance of some meter so that biker has enough time to control down the situation and get out of accidental prone zone.

The basic flow of the flow chart is given out in theoretical manner to get the user the sequence either the flow in which the functioning of bike will be performed, while in practical there are number of activities performed by the technical instrument that is attached or configured in the vehicle for the smooth functioning. Before implementing a flow, chart is made mandatory to explain how the things will be automated and brought into reality. A serious discussion is set up with the engineers and management to study and analyse that how beneficial it could be with respect to successful model and a better business. The marketing strategy is very important before a product for a brand is introduced in the market because the product is directly related with the name of company in case the product fails all the blame goes to the company which in return creates a very bad impression in market also a trust factor is scaled down among people and goodwill of company is crashed shares lows down and much more incident takes place simultaneously hence it becomes very complicated before a strong decision is taken. Once the implementation starts other factors comes into existence such as design and modelling also the engine, mileage is dependent on it.

It requires a skilled worker and engineers to carry out the difficulties and other stuff in order to have a successful completion. Each and every phase of this requires a professional and technical strong team that can overcome the problem and other difficulties with a better solution and conclusion. Once the goal is achieved a better outcome is delivered on the other hand the bikers come out with good review leading to the growth of IOT in automobiles.

3. Benefits & Consequences

Considering security is a major concern in bikes. Bikes are most probably parked in open places where people can temper your bike parts. So, using the IOT technology the rider can monitor their bike through their mobile phones & smart devices. The rider can not only monitor the bike but they can also control various part of bike such as locking their bike handle, locking bike tire, etc.

Leading to have a upper hand to secure your bike.

The other advantage is that the rider stays connected to their known once. The feature is implemented in new generations bikes, through which the rider will be connected to their known once's so incase if the rider took a wrong way for destination the known once can trace him and guide the rider for their way to reach the destination. This is a biggest advantage of

getting connected virtually with their known people, also not getting a feeling of leftover. This scenario can be much related when the rider is travelling a good distance from one place to another the tracking functionality acts as backend for the rider and gives a feeling of less prone to error.[5] This helps the rider to gain the will power and continue the riding with a good mental health which helps to cover a good distance and safe ride.

The functionality of speedometer is also introduced in the bike giving you the feel of connectivity all through the way and act as an assistance for you.

The speedometer in a bike plays a role of assistance which can generate alarm for each and every action on your smart phone as well as bikes such as smart phones battery strength, miscall alert, SMS notification, DND mode, etc. This acts as an external device to communicate and brings up the alert frequently this provide a better use of technology to cooperate. With the use of AI intelligence an assistance can also be setup in speedometer which can assist the rider with all the help which riders need extending the level of comfort zone and betterment. This brings the improvement in riding quality and brings a good connectivity between user and other functionality. This is further responsible to road safety and minimizing the risk of another dangerous accidental scenario. The bike was already inbuild with the brake system in order to control the bike the use of drum brake was introduced for the betterment of rider still the engineers were focused on bringing much more advance functionality to improve the performance hence the ABS was introduced in market. The trial of ABS was liked by many bike companies and users so the companies further aim to bring this on road for the betterment of the biker and good performance. The betterment scenario continued further and lead to revolution of introducing IOT in automobiles.

The IOT based smart bikes can detect various measures to be taken by the rider during their ride or while riding. The IOT based bikes have various monitoring systems that acts dynamic for the user to act when necessary the bike which ca detect various aspects such as if bike gets crashed or any part of bike gets tear down the notification is being send to the rider as well as their known ones this creates the necessary scenario were rider is compelled to take action also the notification is brought under the view of biker and their known ones this lead to keep a track of action that needs to be taken.

A lot of tech may develop its own disadvantages such like Complexity, Privacy issues, Compatibility etc.

The IoT is a diverse and complex network. Any failure or bugs in the software or hardware will have serious consequences. Even the network or the hardware failure can cause a lot of inconvenience.[2] Hence complexity can term as the category that should be taken under consideration.

Due to the upgraded technology there is always a way through either the loop holes that are found by the expert in the field. There is always the possibility of hackers breaking into the system and stealing the data. This may result to the misusing of data and other personal details. There is a possibility of misusing your information and using it against you which may in turn results something bad.

In IOT based system there is no standard rules for devices so there are different compatibility issues.[6] While creating an application it should be created for every version of devices which is a major issue. Mostly it is difficult to make application similar in both android as well as IOS. So while making the application similar most of the time there is compatibility issues to run it on various versions in smart phones.

4. RESULT AND CONCLUSION

This system is an IoT-based solution, here we include the concept of producing the historical data with help of sensors this data is stored in database. The concept of storing the historical data allows bike riders to make informed decision for the route. In addition, this big data can be also used by the government for better monitoring of the environment. Future works will be oriented towards studying data management and analysis for optimal application of the big data. From the result of experimental section, it is clear that the proposed intelligent system is the best in its kind for providing a better alerts and can detect more type of miss-driving as compared to normal driving skills. The basic idea of this system is to minimize the road accidents which are increasing day by day by alerting and warning the driver of their ride styles and providing them the best security necessary and also send the alerts to the parents and one concerned person about the driving behaviour of the driver.

This data can be used in various area of studying and analysing which may help to overcome with current situation and difficulties also the brainstorming ideas of people will help to build up new technologies. The data may help to analyse the most accidental prone areas, roads and may include other scenarios that cause the biker to undergo the risk for life. The State authority can use this data for grabbing, finding and analysing the driver with these unfit driving skills and can be used to study the driving behaviour or a specific area.[1]

The rider always has a impression somewhere in the backend of there thinking that it is under the category of risk due to bad road uncontrollable speed and other factor, this impression is the first point were the biker gets the feel of risk and finds it difficult to ride which in turn may result to accident, with the help of data collection and technologies the biker may develop the feeling of less possibility of accidents which in turn may help to reduce the rate of failures and accidents. To avoid this our research paper proposes this methodology. road safety authorities to get better visibility about the automobile's performance as well as the overall traffic situation, allowing for timely actions.

Many regulations and ideologies are given to the public but they are not ready to follow those and they don’t have awareness about these principles. In future further technologies will be proposed for the better improvement of the smart system.

5. FUTURE SCOPE

In upcoming time bikes can be connected with various sensors which will be connected to strong bandwidth in order to control all the data directly to the android phone. IoT based Key Less Bike that could start by Android Phone.[8] There can be bike which can be actually handled by smart phone’s, a bike ignition can be start and stop using smart phones, A rider can start the bike as well as kill the ignition using smart phone. There can be indicator light pattern of blinking which can be controlled by the smart phone using some application. There can be a auto sensing technique if the bike is being stolen, Which can sense that the bike is being forcefully Moven from one place to another, If such incident is detected the alarm gets started and a notification is sent to the owner of the bike. With the help of sensors various defects in bike can be detected which helps the rider to break down in middle of their ride.

Some concept bike which are going to be build are Honda Riding Assist motorcycle, a concept version of our self-balancing motorbike, became exhibited at CES 2017, the world's largest consumer electronics and patron generation tradeshow.[3] The Honda Riding Assist motorcycle won three pinnacle awards at CES 2017. BMW Motorrad VISION NEXT 100 : All new – and yet the BMW Motorrad VISION NEXT one hundred nevertheless still captures elements from the BMW Motorrad timeline: black triangular frame, white lines, classical opposed-twin engine

type. The design of the front is minimalistic and refined with extraordinary detailing. Above the front wheel, a big metal reflector is integrated within the frame.

6. REFERENCES

[1] A Survey on Internet of Things From Industrial Market Perspective. Author :- Charith Perera ; Srimal Jayawardena ; Chi Harold Liu ; Min Chen. IEEE Access, 2014.

[2] An approach to IOT based parking and reservation system on cloud. In Green Computing and Internet of Things (ICGCIoT), pages 352–354, 2015 Author :- Jatin Kinra , Parminder Singh Sethi and Vaibhav Hans.

[3] https://indianexpress.com/article/technology/tech-news-technology/tvs-ntorq-125-scooter-with-bluetooth-smart-features-launched-in-india-price-in-india-rs-58750-5052172/

[4] A security-aware safety management framework for IoT-integrated bikes,2015. Author:- M.A.Razzaque , Siobhan Clarke.

[5] Mission on! Innovations in bike systems to provide a safe ride based on IOT, 2017 2nd International Conference on Computing and Communications Technologies (ICCCT), 2017. Author:- Archana D. ; Boomija G. ; Manisha J. ; Kalaiselvi V.K.G.

[6] International Journal for Research in Applied Science & Engineering Technology (IJRASET), IoT based Smart Bike Monitoring System, Apr 2019. Smita yawle, Sushrut Wadnerkar, Vikas Raibole, Gauri Jane, Sahil Radke, Prof. Komal Bijwe

[7] COLLISION CONTROL FOR BIKERS USING IOT, 2018. Author :- Ramya .D, Sujatha.K, Devi Kala Rathinam. D, Sherin.J, Shilpa.A.

[8] Design and Implementation of A Bicycle Parking System Based on Internet-of-things. National Conference on Information Technology and Computer Science (CITCS 2012). Author :- Yujia Huang, Zhongliang Yang and Shuhua Xiong.

 ISBN: 978-81-946375-7-8

Mango Plant Leaves Disease Detection by Using Image Processing

Omkar Mayekar
Student, MCA, DES Navinchandra Mehta Institute Technology and Development, Dadar (W) /University of Mumbai

Shruti Salvi
Student, MCA, DES Navinchandra Mehta Institute Technology and Development, Dadar (W) /University of Mumbai

Dr. Rasika Mallya
Associate Professor, MCA Department, DES Navinchandra Mehta Institute Technology and Development, Dadar (W) /University of Mumbai

Abstract
Mango is the famous delightful fruit and cash crop. When diseases influence the mango plant there is a significant reduction in the production of mango due to which agriculturists suffer in selling their yield. This difficulty caused me to acquire different ways to identify and diagnose the diseases influencing the mango plant and devise the expert system to check those. To improve product and quality it is essential to restrain all such dangerous diseases at the earlier stage. In our country most of the farmers are uneducated. So, they cannot get accurate knowledge about the disease. It needs an expert advisor or a farm officer. But it is difficult for an expert advisor or an agricultural officer to communicate every farmer.In the specified system for image enhancement "Imad just" function is especially applied for contrast enhancement. This is an inbuilt function in MATLAB. Contrast-enhanced is performed by linearly scaling pixel values, and characteristics used are Correlation, Contrast, Entropy, Energy, Cluster prominence, Homogeneity, Variance, and Dissimilarity. The classifier applied is SVM. The purposed system gives an accuracy of 90% while testing on 92 samples.

Keywords: mango disease detection, production, image processing, diseases

I. Information

Now, the growth of mango fruit decreases because of climatic situations and environmental affairs, like heavy raining, high humidity, reducing soil nutrients, a difference of associated diseases and problems.Typically, the exposure of mango plant diseases is done by simple eye observation, which gives less assurance. Low productivity of mango fruit is due to the various diseases affecting mango plant which are not recognized by the farmers as they are illiterate. In some rural areas, farmers may have to go long distances to visit experts, this may be too expensive and time-consuming and even farmers are unaware of diseases.Kisan call canters are available but do not service 24x7 hours and sometimes communication will be failed. Sometimes, reply getting from the agricultural officer has hesitated. Hence, proper disease detection is not done within a less time which results in a reduction of mango fruit production. This research paper presents a basic idea about the detection and identification of various diseases that happened in various parts of the mango plant like stem, fruit, leaves, branch, etc. also, proper defensive care is given for the affected area of the mango plant so, that production of mango fruit will gain as well as the quality of the mango fruit will increase.

 ISBN: 978-81-946375-7-8

Mango Plant Leaves diseases examination and its symptoms

The technique like RGB image feature pixel counting is largely applied to farming science. Image examination can be put in for the following purposes:

1. To identify mango plant leaf diseases.
2. To assess affected region by disease.
3. To assess the color of the affected region
4. To find the borders of the affected region.
5. To control size & shape of leaves.

There are some common symptoms of mango plant leaf diseases, they are as follows:

1. **Anthracnose in Mango crop:** Colletotrichum gloeosporioides

Symptoms

- Leaf of the product spots, wilt tip, sprigs canker, blossom canker and fruit rot.
- On the leaves and sprigs there is small bleb like spots grow. Also warm sprigs wilt and die end symptom emerge on it.
- Necrotised tissue drop out leads to shot hole symptom
- Branches were affected ultimately dry up & black spots appear on fruits..
- The fruit mash becomes very firm, Split and rot at mellowing. Infected fruits drop

2. **Powder type mildew in Mango batch:** Acrosporum mangiferae

Symptoms

- It attacks the leaves, stem, flowers of raceme and fruits.
- Grimy or white powder type growth emerge on leaves.
- Discarding od infected leaves come when the disease is acute.
- The fruits which are affected does not grows in size and fall before obtaining pea size.
- Survives as inactive fungus in infected leaves.
- Secondary expand by air born conidia

3. **Mango miscreation in Mango batch:** Fusariosis moniliform var. subglutinans

Symptoms

Two major types of miscreation are vegetative & floral/blossom.

- Vegetative miscreation is loss of acute dominion and limping of vegetative sprout in the leaf acute angle or at the end.
- The sapling gives little shootlets bearing, cancroid leaves gives bunch like appearance.
- Saplings normally sprout vegetative buds all over the inter-nodes making ropes besom like symptoms distinguish.
- Vegitative buds also sprout in trees which are growing up where little bunches are constructed at the terminals.
- Number of flowers were reduced or no flowers.

- The terminals which turns into bunch eventually dry leads to expiry of the shoots.
- Floral miscreation emerges with the exposure of blooms where the flowers are much large.
- In some scenarios, the flower buds not at all open and remains dull green.
- Branches with diseased bloom can cause both miscreated panicles which give leafy appearance unlike typical panicles.
- Miscreated panicles usually do not carry and if there is a fruit in the season, it does not extend more than the size of pea.
- Flower buds are converted into vegetative buds and huge number of little small leaves and stems, decreased internodes compressed together by giving a witches broom like emergence.

4. **Sooty mould disease of mango:** Capnodium mangiferae Cke. & Brown, Capnodium ramosum Cke.

Symptoms

- Black velvety covering on the leaf surface.
- The whole surface of leaf might be covered or initially the symptoms may be present only as flakes on the leaf.
- In serious cases, the whole tree turn completely black due to the presence of mould over the entire surface of sprigs and leaves
- Under dry spell such affected leaves curl and shrivel.

5. **Leaf spots disease of mango:** Xanthomonas campestris pv. Mangiferae indicate

Symptoms

- On leaves, te disease first emerge as a regular to angular lift up little water-soaked wound measuring 1-5 mm in dia.
- Normally packed at the top which increase in size and become brown to black in color.
- These wounds are covered by yellow halo at first.
- Sometimes these wounds combine and form huge, rough and lifted necrotic spots.
- In serious infections, the leaves become yellow and drop-off.
- On leaf stalks, the blight sometimes develop seemingly along the mid-rib.
- On branches, sprigs and stem, the fresh wound are water-soaked, which later become increased and dark brown with up-and-down fissures discovering the vascular tissue filled up with gum which flows out.

II. Literature Review

The main purpose of why we have taken this approach of recognizing mango plant leaves diseases is to help the production of mango fruit in India, which is exhausted due to various diseases. This research paper represents how various diseases attack mango fruit fertility. The foremost diseases that attack mango fruit fertility are Anthracnose, Powdery Mildew, and Red Rust according to research paper thus, we have collected those four diseases along with a subsidiary disease called Black Banded.This research paper also represents preventative techniques and antidote for the disease [1].

This research paper introduces MATLAB based disease exposure operation. This way shows it's the capacity to identify leaf-based diseases though many influential diseases influence the fertility of mango crops. For illustration, Black Band which majorly transpires on a stem and yet affects fertility. few diseases are there which transpire on the fruit itself. Precision is also the foremost constraint that requires advance.

The initial level of digital image processing is "Image Preprocessing." which is 1 of the foremost steps for maintaining originality in an image that may be corrupted due to some noises or due to other purposes. This paper is a review of several Image Preprocessing techniques that can be employed [2].Image segmentation is important for choosing the proper area of interest in an image. This is necessary to decrease processing time that may be needed to prepare rather a whole image. This research paper presents a different segmentation method called Fast and Robust Fuzzy C- Means algorithm.

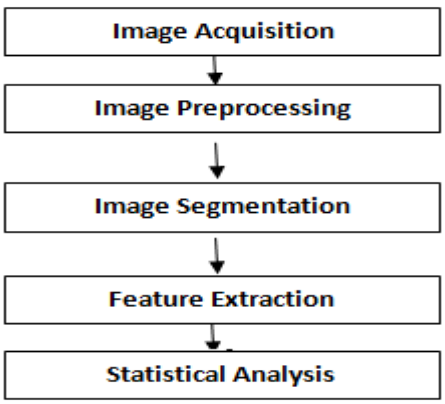
The advantage of this algorithm is that it automatically chooses a group that includes disease. Also, it differs showing benefits over the conventional FCM algorithm [3]. This research paper presents a summary of the characteristics to be obtained. Characteristics that can be procured from the image may of various types like global or local.The MATLAB having an in-built function for rendering GLCM.This research paper presents equations from GLCM matrix for calculating different texture-based articles. The matrix function is GLCM of viewpoint and distance thus GLCM computational cost is very high.

III. Methodology

There are five important steps which used to detect mango plant leaf diseases.The processing plan consists of image acquisition via digital camera or web, image pre-processing involve image enhancement and image segmentation where the affected and useful areas are segmented, property extraction and classification. At last the presence of diseases on the plant leaf will be identified. In the beginning step, RGB images of leaf samples were picked up. The step-by-step procedure as shown below:

1) RGB image acquisition
2) convert the input image into color space
3) Segment the components
4) obtain the useful segments
5) Computing the texture features

System Architecture: The basic steps for mango plant leaf disease detection and classification using image processing are shown.



 ISBN: 978-81-946375-7-8

a) Image Acquisition

Image acquisition is the first step of image processing.it is defined as seizing the image through the camera and save it in digital media for performing further MATLAB operations. It is also an act of retrieving an image from hardware, so it can be passed through the further process. In our work, using the digital camera we captured healthy and diseased images of the leaf.

Fig.1 Healthy Mango Leaf Fig.2 Diseased Mango Leaf

b) Image Pre-processing

The major goal of image pre-processing is to enhance the image data held rejected distortions or to enhance some image features for more processing. The change image size and shape, noise filtration, image conversion, enhancing the image and morphological operations this is different techniques of the pre-processing method.In this work, we used different "MATLAB" codes to resize an image, to enhance contrast and RGB to grayscale conversion as shown in fig.3, for extra operations like making clusters in segmentation.

Fig.3 contrast enhanced

c) Images Segmentation

The third step is the image segmentation. in this step image is segmented into multiple sections, i.e., sets of pixels. It also analyzes the complex image into a simple image. Segmentation can be categorized as follows: Region-based, Edge-based, Threshold, Feature-based, Clustering, Model-based. The output of the pre-processed image is accepted as input to find the affected part of the mango leaf section.

Fig.4 Diseased leaf image clusters

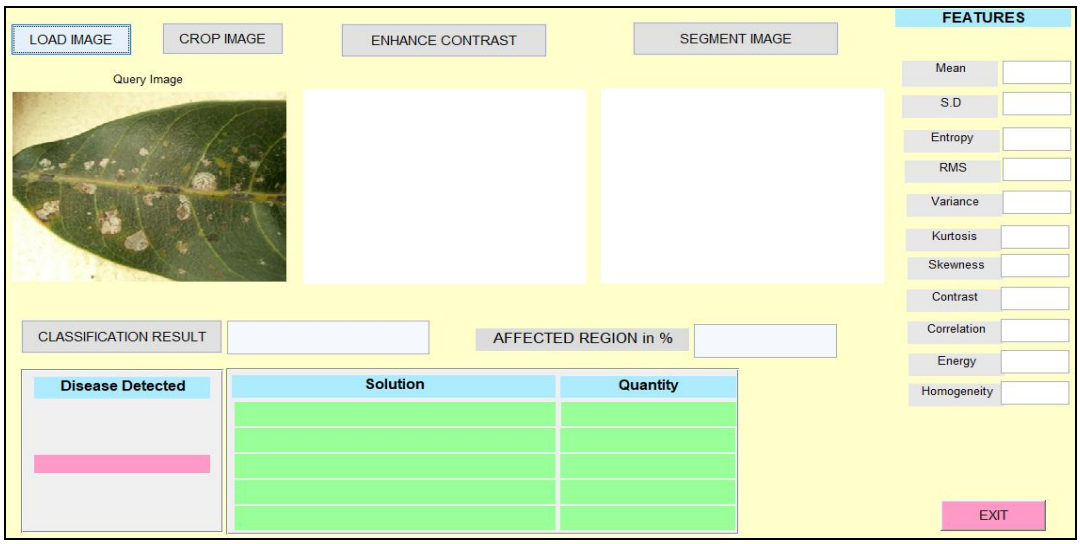
 ISBN: 978-81-946375-7-8

d) Feature Extraction

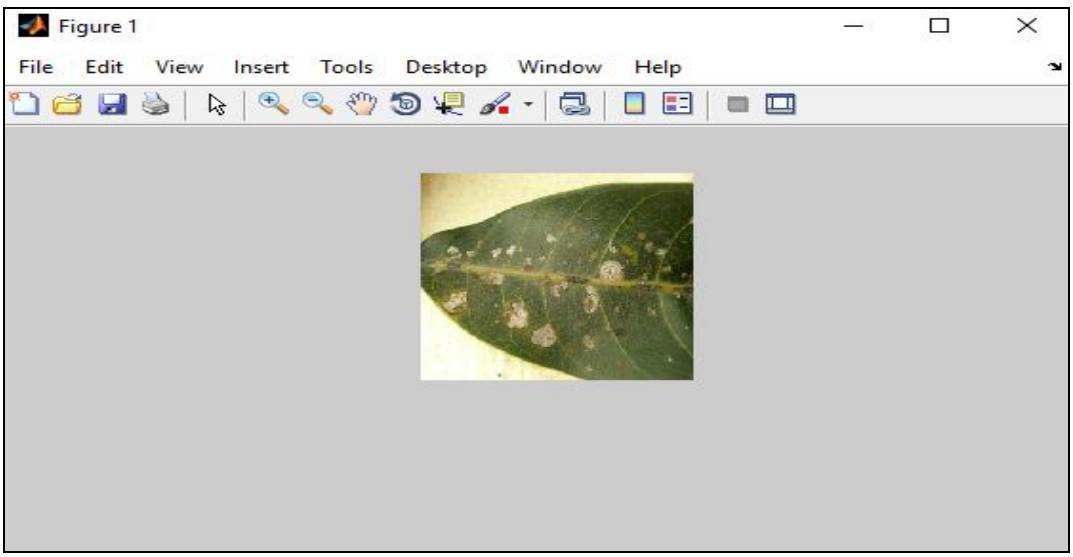
The main aim of this step is to detect and extract features that define the significance of a given image. Image feature extraction involves color, texture, and shape in image processing, which is the most basic visible feature. In the feature, extraction craved feature vectors such as color, texture, morphology, and structure are plucked. Feature extraction is a method for including the number of resources that are needed to describe a huge data correctly. Statistical texture characteristics are obtained by Gray level co-occurrence matrix (GLCM) formula for texture analysis and texture features are calculated from the statistical distribution of observed intensity combinations at the defined position related to others. Numbers of gray levels are essential in "GLCM" also statistics are classified into an order of first, second & higher for the number of intensity points in each combination.

IV. Implementation

1. Load Image

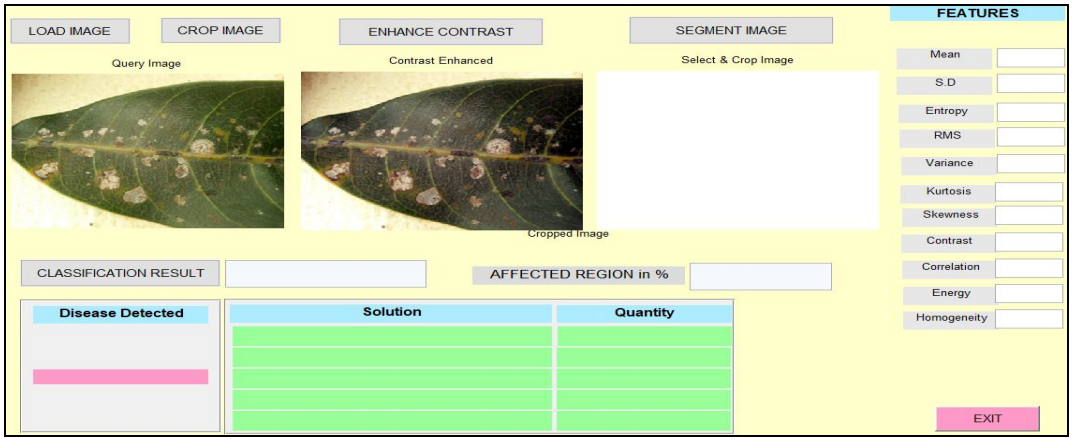


2. Crop image if necessary

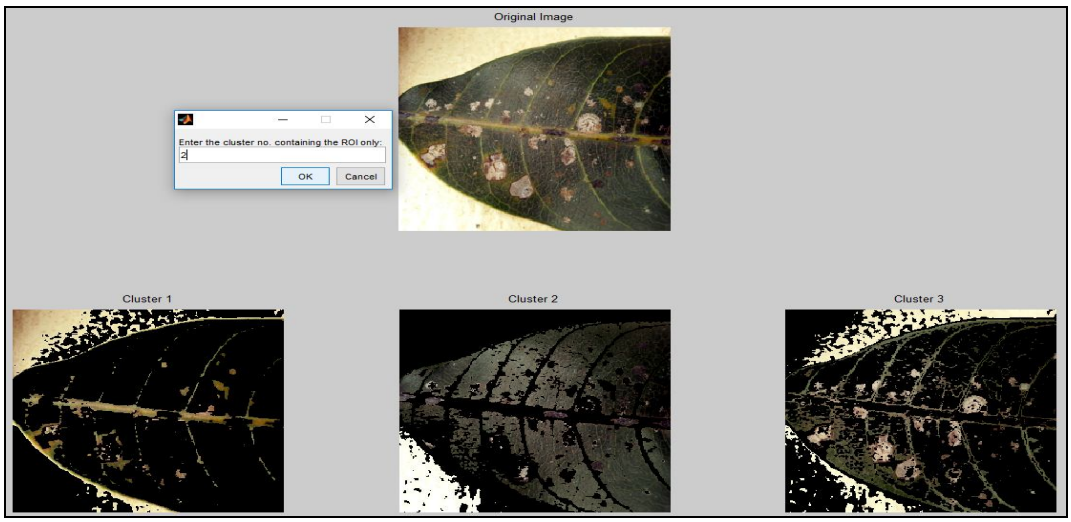


 ISBN: 978-81-946375-7-8

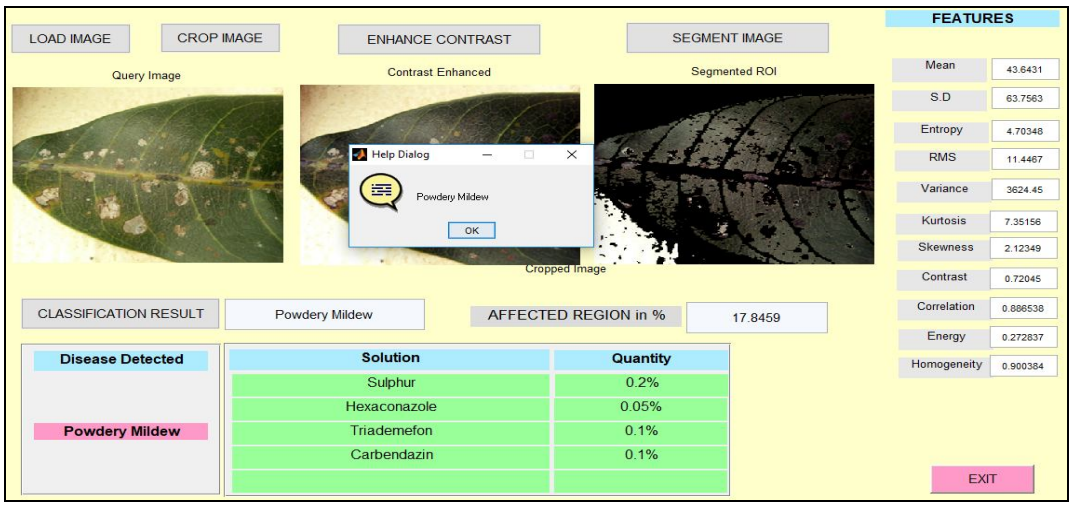
3. Enhance image contrast



4. Click on Segment Image And select appropriate cluster



5. Click on Classification Result. You will get the detected disease name with an appropriate solution.



V. Conclusion

This research paper demonstrated proper and an efficient mango plant leaves disease detection and analysis technique by using MATLAB image processing. The recommended system permits us to detect and identify diseases influenced by mango plant leaves and to give appropriate defensive care and solution for it. K-means and Multi SVM techniques which are configured for mango leaf disease detection. MATLAB software is perfect for digital image processing. K-means clustering and SVM algorithm provides high efficiency and uses very less time for complete processing. In future work, we will enlarge our database for more plant disease associations.

References

1. O. Prakash and A. Misra, "Important diseases of mango and their effect on production," Biol Memoirs, vol. 18, pp. 39–55, 1992.
2. K. N. Shukla, A. Potnis, and P. Dwivedy, "A review on image enhancement techniques," International Journal of Engineering and Applied Computer Science (IJEACS), vol. 2, no. 7, pp. 232–235, 2017.
3. T. Lei, X. Jia, Y. Zhang, L. He, H. Meng, and A. K. Nandi, "Significantly fast and robust fuzzy c-means clustering algorithm based on the morphological reconstruction and membership filtering," IEEE Transactions on Fuzzy Systems, vol. 26, no. 5, pp.3027–3041, 2018.
4. J. M. Patel and N. C. Gamit, "A review on feature extraction techniques in content-based image retrieval," in 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET). IEEE, 2016, pp. 2259–2263.
5. R. M. Haralick, K. Shanmugam et al., "Textural features for image classification" IEEE Transactions on systems, man, and cybernetics, no. 6, pp. 610–621,1973.
6. Y. Hu, C.-x. Zhao, and H.-n. Wang, "Directional analysis of texture images using grey-level co-occurrence matrix," in 2008 IEEE Pacific-Asia workshop on computational intelligence and industrial application, vol. 2. IEEE, 2008, pp. 277–281FLEXChip Signal Processor (MC68175/D), Motorola, 1996.
7. D. L. Padmaja and B. Vishnuvardhan, "Comparative study of feature subset selection methods for dimensionality reduction on scientific data," in 2016 IEEE 6th International Conference on Advanced Computing (IACC). IEEE, 2016, pp. 31–34.
8. O. Prakash, Compendium of mango diseases and disorders Capital Publishing Company, 2003

 ISBN: 978-81-946375-7-8

Social Distancing Alert Using Internet of Things (IOT)

Mr. Prathmesh R Lokhande
Student, Master of Computer Application, Navinchandra Mehta Institute of Technology and Development, Dadar (W), Mumbai

Abstract
The current state of emergency due to COVID-19 has affected our world in several ways. In matter of days world have had to change the way of live. Community Transmission has begun in the world and cases would be expanding and this will lead to increase of people's death. In this critical situation, a technology like IOT has become a powerful tool. Outpace is a method to reduce the crowd interactions and to put a stop from spreading the viruses within groups of people. The focus of this paper is to develop a device using IOT and to separate oneself within groups of people.

Keywords: Internet of Things (IoT), Ultrasonic Sensor, Social Distance, COVID-19, Corona virus

I. Introduction

The current outbreak of the novel corona virus also known as COVID-19 was declared as a public health emergency by the WHO were over a million people have been affected by the disease. Due to this pandemic situation, all has been lockdown in their homes. Several Countries has declared lockdowns. Initiated by the country China, Wuhan is the first city to declare the lockdown. As the spread of virus started from Wuhan and gradually speeded to other parts of countries like Asia, Europe, Australia, etc. they also started to follow the same process of Social Distancing and Isolation. Social distancing is a self-discipline act that should be used by all so that it will be a barrier between the infected people and those who are not infected by the disease. As it will lead to stop the spread of virus.

In matter of days world have had to change the way of live. Even now when things are slowly getting back to normal, it is important to maintain safe distance between people to prevent spreading of the virus. No matter where we are: universities, libraries, banks, etc. problem of maintaining safe distance remains great challenge. The World Health Organisation (WHO) has recommended a minimum of 1 metre distance between oneself and anyone who is facing the symptoms of coronavirus disease. This device alerts you if anyone comes closer than 1 metre distance.

II. Literature View

Triax Technologies, Inc., Norwalk, CT, a provider of Internet of Things (IoT) worksite technology, has introduced Proximity Trace, a tool that delivers proximity-distancing alerts via a wearable device for workers across many industries, including manufacturing. This technology offers protection for essential workers during this situation, so that the workers can come back to offices, while addressing recommended social distancing practices. Proximity Trace devices can be worn on the body through a device which attach on workers helmet. This device will emit a louder alarm if the workers are too close to each other. In this way, the workers can work without being worrying about keeping social distancing [1].

Ravi Pratap Singh, Mohd javaid, Abid Haleem and Rajiv Suman has reviewed and discussed different ways to fight against these situations using Internet of Things applications. These IOT

implementation will results on decreasing the healthcare costs and improve the condition of the infected patient. Therefore, this study based research is endeavour to survey, discuss, and spotlight the overall applications of the IOT philosophy by offering a perspective protocol to tackle the COVID-19 situation [2].

Mark Terry has published the Research Roundup- Social Distancing Works and More in that he shares view on how Social Distancing helps to control the spread of virus.

Fahim Aslam has published the COVID-19 and Importance of Social Distancing. In the modern world interactions are done on regular basis using technology available but majority activities are carried out on a daily basis which relies on social interaction. Socialization is necessary from an economical perspective for the country as workers need to perform their duties on a regular basis so that the company can functions smoothly [3].

Anisha Sinha has reviewed and discuss IoT and its uses for various applications. IoT is an innovation, that utilizes web for controlling electronic, mechanical, autos and other physical gadgets associated with the web. IoT can be utilized for various purposes like industry, designing and foundation, home, office, security, wellbeing and medication. IoT can be utilized for checking air and water contamination, outrageous climate, business cultivating [4].

III. Proposed System

The components used are Arduino UNO board, HC-SR04 Ultrasonic Sensor and Buzzer. The programming language used is Arduino IDE. The proposed method uses IoT systems for Social Distancing Alert device. In this method the HC-SR04 Ultrasonic sensors send waves. These waves are absolutely invisible and come back after hitting an optical. The Trig pin activates the output pin. In which LED and Buzzer are connected. A distance of 1 metre has kept as per WHO guidelines.

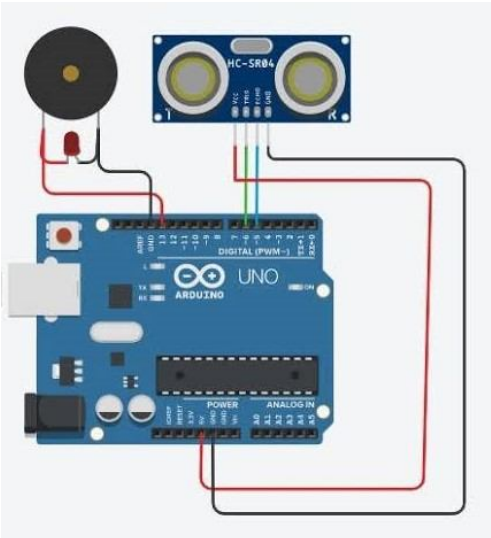


Figure no 1: Circuit Diagram

As shown in figure no. 1, HC-SR04 Ultrasonic Sensor connected with Arduino UNO board. HC-SR04 Ultrasonic sensor identify the object in 1 metre distance passed that data to Arduino UNO board. Then Arduino UNO board checks the distance between object is less than 1 metre or not. If condition satisfies, then Arduino UNO board generates output in form of buzzer and LED.

 ISBN: 978-81-946375-7-8

Figure no 2: Social Distancing Alert device

As shown in figure no. 2, Social Distancing Alert device how it looks. People can wear this device as ID card in several places such as universities, libraries, banks, industries, market, etc. People can carry this device everywhere. This device is useful for everyone to protect yourself in crowded area where chances of spreads of the virus is high. This device design is made for students as well as employee perspective. When things are getting normal places likes schools, colleges, industries, etc. has high chances of spreads of corona virus. With the help of this device they maintain social distance between them and work comfortable. They can also attach their ID cards with this device.

Material Used

Hardware required for implementation:

1. Arduino Board: The Arduino Uno is a microcontroller board based on the ATmega328 (datasheet). It has 14 digital I/O pins (of which 6 can be utilized as PWM outputs), 6 simple analog inputs, a 16 MHz ceramic resonator, a power jack,a USB connection, and a reset button, , an ICSP header.

Figure no 3: Arduino UNO board

Arduino can be utilized to communicate with a PC, another Arduino board types or different microcontrollers boards. The ATmega328P microcontroller gives UART TTL (5V) serial communication medium which should be possible utilizing as digital pin 0 (Rx) and digital pin 1 (Tx). The ATmega16U2 firmware utilizes the standard USB drivers, and no outer drivers are

required. The Arduino programming has a serial monitor which permits straightforward printed information to be sent to and from the Arduino board. A SS library allows for serial data exchange on any of the Uno's pins. There are two RX and TX LEDs on the arduino board which will streak when information is being transmitted by means of the U2S chip and USB connection with the PC.A SoftwareSerial library takes into account serial medium on any of the Uno's digital pins. The Arduino programming incorporates a Wire library for utilization of the I2C transport bus.

2. HC-S04 Ultrasonic Sensor: The HC-SR04 Ultrasonic sensor is a 4 pin module, whose pin names are Vcc, Trigger, Echo and Ground respectively. This sensor is an exceptionally well known sensor utilized in numerous applications where estimating and measuring or detecting objects are required. The module has two eyes like activities in the front which shapes the Ultrasonic transmitter and Receiver. The sensor works with the straightforward secondary school equation that Distance = Speed * Time. The Ultrasonic device transmits a ultrasonic wave, this wave goes in atmosphere and when it gets back hitting by any material it gets reflected toward the sensor this reflected wave is captured by the receiver device.

Figure no 4: HC-SR04 Ultrasonic Sensor

HC-SR04 distance measuring sensor is regularly utilized with both microcontroller and chip like ARM, Arduino, Raspberry Pie, PIC, and so on. The current consumed by the sensor is under 15mA and consequently can be straightforwardly controlled by the on board 5V pins. The Trigger and the Echo pins are both I/O pins and henceforth they can be joined with I/O pins of the microcontroller. To begin the measurement, the trigger pin must be made high for 10uS and afterward switches off. This activity will trigger a wave at frequency of 40Hz from the transmitter device and the receiver device will wait to that the wave will return. When the wave is returned by getting reflected by any abject the Echo pin goes 1 for a specific measure of time which will be equivalent to the time taken for the wave to return back to device.

IV. Future Enhancement

- In this research paper the ultrasonic sensor is to identify the object which comes from front.
- Need to modify the design or add some another sensor which sense the objects from each sides.
- Need to reduce the size also so people can carry in pockets.

V. Conclusion

Social distancing is an effort to stop the transmission of coronavirus in large crowds in meetings, movie halls, weddings and public transport. To enforce the importance of social distancing, schools, colleges, malls and movie halls are now closed across states. To reopen such places and control the spread of virus people have to used Social Distancing Alert. It helps us to maintain social distance in several places.

 ISBN: 978-81-946375-7-8

VI. References

1. Triax Technologies, Inc., Norwalk, CT.: a provider of Internet of Things (IoT) worksite technology.
2. Ravi Pratap Singh, Mohd javaid, Abid Haleem and Rajiv Suman: Internet of things (IoT) applications to fight against COVID-19 pandemic.
3. Fahim Aslam: COVID-19 and Importance of Social Distancing, April-2020.
4. Anisha Sinha: Study of IoT (Internet of Things) and its Application, International Journal for Research in Applied Science Engineering Technology, ISSN: 2321-9653, Volume 8 Issue VI June 2020
5. Mark Terry: Research Roundup- Social Distancing Works and More
6. http://bweducation.businessworld.in/article/LPU-Researchers-Develop-KAWACH-IoT-Device-For-Safe -Social-Distancing/25-03-2020-187238/
7. https://timesofindia.indiatimes.com/life-style/ health-fitness/health-news/ coronavirus-outbreak-the-rules-of-social-distancing/articleshow/74650718.cms

 ISBN: 978-81-946375-7-8

Review of Virtual Reality Applications in Different Sector

Rohit Jogale and Snehal Pawar
MCA, Deccan Education Society's Navinchandra Mehta Institute of Technology and Development, University of Mumbai

Abstract
In today's world, there are many technologies in which humans are surrounded. Some technologies had acquired in our lives. These technologies are being integrated to provide new services where computer systems are the heart of such integrated systems. This paper will cover various sector and aspects of virtual reality system like how virtual reality system will be automating and provides facilities to user in different sectors. This paper is going to deal with the various application of virtual reality. Current virtual reality is being used in many areas such as advance fields of engineering, education, medicine, training, and entertainment. Virtual reality is a three-dimensional interactive environment technology that is based on a computer used for simulation of reality. The virtual reality is bringing into an imaginary world that is looking like a real world.

Keywords: Virtual Reality, Virtual environment, Application of virtual reality, Sector of virtual reality, Simulation, Three-dimensional system, Training.

I Introduction

Virtual reality technology is based on computer systems that take input and provide outputs to devices by allowing interaction of users and experience which creates an artificial environment. Virtual reality is basically based on our visualizations which enable users to do something with a controlled environment which not really exist in the environment. In a virtual reality system environment, the user can do a similar routine like throwing tennis ball simulation or simulation of flying in space. The system can be made with a hand gesture or a simple nod. The research paper includes working of virtual reality, applications, and their consequences.

The virtual reality is the mimic of the real world. It, not just the flat TV monitor viewing, it gives the 3D visual experience. The origins of virtual reality can be traced as far back at least as "The ultimate display"[11] with the help of powerful microcontrollers. VR technology contributes greatly efficiently in many applications. The VR is of the following type Fully immersive, Non- immersive. Collaborative, Web-based, Augmented reality. Research has result which demonstrate that the virtual reality technology how growing and rapidly acquiring in different applications sector. This technology simply based on visualization of our own. This paper describes our research to design and implementation of VR technology. Every application has its own purpose to include VR system which helps automate the any complex process in simple way.

II Literature Review

Joam McComas, PhD, PT [1] represented their paper on "virtual reality for application for prevention, Disability Awareness, and Physical therapy rehabilitations in neurology". In this paper, the focus is on improving spatial abilities in people with mobility impairments. Paper also explores the importance of virtual reality in physical therapy, brain injury, and neurology.

Yalong Yang, Bernhard Jenny Faculty of IT in Monash University Australia - They are represented their paper on "Maps and Globes in virtual reality". This paper explores different

ways to render worldwide maps with virtual reality. The paper contains the exact use of virtual reality technology which includes four interactive visualizations geographic data which is Flat Map, Egocentric globe, Curved map, Exocentric globe.

Rasika Sonje - Presented research paper on "Virtual reality for Real Estates". This research paper includes the mixing of a physical world with a digital world. The research paper has mixed reality to evolution in humans, computer and the environment.

With the application of Virtual Reality in architecture design, architecture engineer is able to achieve their works in the virtual world, which is visible, touchable and sensible [4]. Virtual Reality technology not only can give a new way of artistic expression for architectural creation but also let the architectural design method and concept greatly changeable [4].

Danial Vogel, Paul Lubs represented a research paper on "Animation Virtual Reality – interactive controller-based animating in virtual reality." Virtual reality is an animated system [5]. The animation creation process is a tedious and complicated process. This technology mainly used in different domains as video games, films, and education. This paper shows how two-dimensional animations transform into three-dimensional animations using virtual reality system software.

Kjell Brunnstrom – The research represented on "Quality of experience for virtual reality simulator. The research paper shows how virtual reality provides quality of experience to users. It is depending on the purpose of virtual reality in different sectors. Quality of Experience is the degree of the delight or annoyance of the user of an application or service [6]. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in light of the user's personality and current state [6].

Hassan Aziz Represent paper on "Virtual Reality Application in Healthcare". The research paper focuses on applications in healthcare that demonstrate the potential to cognitive address, Functional and psychological impairments [7]. Virtual reality creates an opportunity for training and education for clinical practitioners. This paper also covers exposure therapy using virtual reality, Distraction Therapy, Virtual reality in surgical planning and performance tracking, Virtual reality as a diagnostic tool.

Munehiro Nakazato and Thomas S. Huang represented their paper on "3D mars: Immersive virtual reality for content-based image retrieval". Three-dimensional mars are interactive visualization system for content-based image retrieval. In 3D MARS, the user browses and queries images in an immersive 3D Virtual Reality space of CAVE. The results of the query are displayed in 3D so that the user can

See the result with respect to three different criteria such as color, texture, and structure [8].

Bouchlanghem N, Thorpe A. – represented their paper on "Virtual Reality Application in Construction Industry"[9]. This paper contains an investigation of the use of virtual reality in the construction sector. This technology used for creating a three-dimensional image of the construction project model. It is very important to aspect for involving virtual reality in the construction sector is to lower cost and time to make a physical mock model of construction.

 ISBN: 978-81-946375-7-8

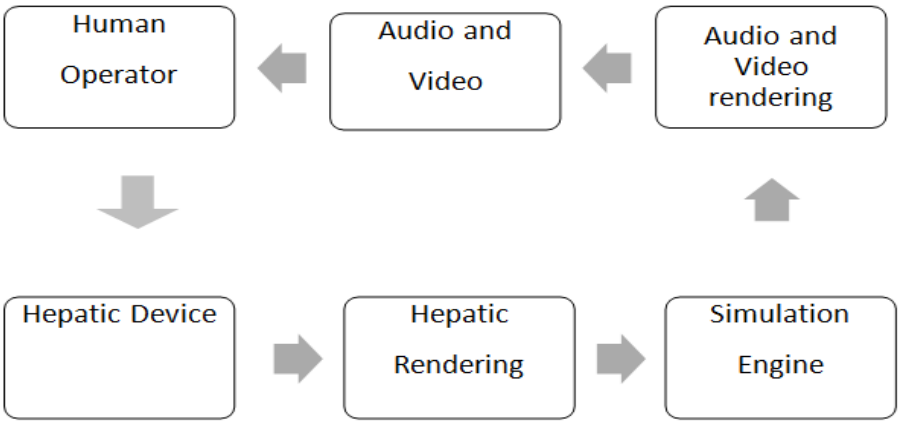


Figure 1- Basic architecture of virtual reality system components [13]

Virtual reality architecture has basic three components Input components, Simulation components and rendering components.

Input Components- The input components has input devices like keyboard, joysticks, 3D position devices (gloves, body suit), voice system or hepatic devices etc. It is also consisting gesture recognition systems which take inputs for virtual reality system [2].

Simulation components (Simulation engine) – This is main component which is iterated each time or frame. Simulation process handled by simulation engine.

Rendering process –The rendering process which is creates sensations for output data. There many rendering processes like video rendering, Audio rendering and hepatic rendering.

III. Applications of virtual Reality in different sectors

1. Gaming

VR gaming is a new engender of computer games in which an individual can experience a 3D environment and interact with it. VR games also introduce immersive experience (360-video content as well as heavy graphics), a new level of interaction (ability to control and modify the game environment), content. There is a different kind of VR games are available [2]. Some of them are as follows-

- Shooter games- In this category, players feel like they are on the real-world battlefield with proper audio.
- Racing game- While using virtual reality in racing games, there is no trouble with movement limitation. These games usually have advanced graphics.
- Horror games- This category of VR games has a very good scope in the market because it provides more realistic graphics. HD based VR headset provides a 360-degree view and object-based audio.
- Adventure games- This one is a well-liked category in the range of VR games. It provides 3D comprehensive graphics, sensors, and controllers.

2. Telecommunication

VR technology is also used in telecommunication, mostly in mobile which allows an easy approach to different VR projects.

The principle ultimatum is that of coping with a medium which mainly depends upon voice tone, pitch, gesture, and visual communication as compared to spoken words.

The conventional way of communicating on the telephone is replaced by video conferencing, Skype and live chat. These communication mediums are operating on the internet and they are inexpensive and suppler. Telecommunications are frequently utilized to assist virtual reality systems like surgery simulation. An example of this is remote surgery which includes images from that surgery that can be transmitted to various locations around the world. Due to these, surgery can be performed in distant locations using robotic and VR technology. Although the goal of tele robotics is autonomous operation, a supervising human operator is still required in most of cases [1].

3. Defense

In the defense sectors of country virtual reality is more important for training purposes. Virtual reality adopted by the military which includes army, navy, air force. It is particularly used for soldiers which is a simulation for a combat situation. It also has other dangerous settings where soldiers are learning to react with different conditions and situations. The virtual reality helps them to visualize the battlefield, where they are completing training without risk of death [2]. This is safer and less costly than the traditional training of soldiers. The military sector contains many types of virtual reality simulation fields which are flight simulation, Medical training in the battlefield, combat situations training, Battlefield vehicles simulations. The system will provide immediate feedback on every soldier, which helps to track the performance of soldiers.

4. Healthcare

Virtual reality in healthcare is one of the biggest adaptation of technology. It covers all aspects of healthcare which includes surgery simulation, skills training. The main purpose of technology is to learn new skills and enhancing knowledge in healthcare in a safe environment. It allowing this without any danger to the patient. There are several types of applications of healthcare which uses virtual reality – Simulation Software's, Diagnostics, Robotic Surgery, Dentistry, Medicine, etc. Simulation software provides a simulation of surgery which is a doctor practicing and gaining knowledge about surgeries [8].

Virtual robotic surgery – The virtual reality use for training purposes and control by the surgeon. It is reducing the time and risk of complications. This field supports remote tele surgery which is surgery performed by a surgeon from a different location to the patient. Mostly think of virtual reality connected with surgery only but this technology also used in non-surgical ways like diagnostics tools. It also used in medical tests like x rays, Scans, testing of blood samples to determines the cause of medical conditions.

5. Fashion Industry

Virtual reality in Fashion is not springing immediately in mind when thinking about this technology. Virtual reality used in the fashion sector in different ways. Virtual reality with the fashion industry includes software that includes building virtual fashion stores and 3D images to help clothing designs. Some fashion application uses virtual reality technology. Lens kart online business company uses this technology. Customer can be visualizing himself/herself without using lenses. It is less time consuming and the customer uses multiple lenses and visualizes how they are looking with that lens.

6. Construction

There is wide use of virtual reality in the construction industry, which has very high in efficiency and low profit. The virtual reality system made three-dimensional models of construction building where they experience like in the real world [10]. Construction of building projects with virtual reality offers many benefits. It helps in making real and actual construction structures within lesser time. Making a physical small scale model is costlier and more time consuming than the virtual reality system model. This model can be having a three-dimensional

view of construction. Virtual reality helps to reduce errors that are present in a constructed building. Virtual reality simulated construction of a building. In the future virtual reality, the system allows for making robust buildings in a smaller amount of space.

7. Automobile

There is continuous improvement needed for vehicle designs. Because of involving virtual reality in automobiles industry improvement, testing and review of vehicle's designing are easier than the traditional process. The designing process is the most time consuming and expensive process for the car manufacturing process. Modifications perform on an existing model which improves car design and creating a new model. The virtual reality enables designers to visualize the car model which saves time and cost [2]. Using virtual reality car manufacturer designs interior as well as the exterior of vehicles. Designing a physical model or vehicle prototype is more time consuming than creating a virtual reality system. Further design can implement on the car exterior in real-time using virtual design. Virtual reality provides a three-dimensional view of a car which is highlighting features of cars and body designs.

This technology also provides training for engineers in the automobile industry.

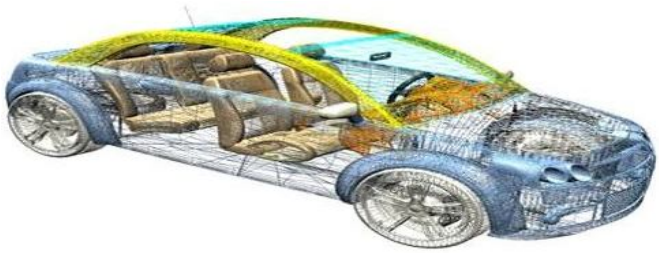


Figure 2 - 3D model of car using virtual reality [2]

IV CONCLUSION AND FINDING

In the future, Virtual reality as per research that found, virtual reality becomes mainly stay at home and working places. In the Evolution of computers, it becomes faster and they will able to create more realistic graphical images for better simulation. This technology already introduced in various sectors, where some sectors are not used in virtual reality. It is also possible that in the future peoples can communicate with each other by using virtual phones. In future virtual reality will be allowing one person to see a three-dimensional view of others. Virtual reality applications are benefitted and improve the design, assembly, engineering, inspection, and production process. In future virtual reality replace all conventional methodologies into perfect visualize model to reduce time, and cost.

REFERENCES

[1] Tomasz Mazuryk and Michael Gervautz; "Virtual reality History, Application, Technology and Future"; Institute of Computer Graphics Vienna University of Technology, Austria.

[2] Ronak Dipakkumar Gandhi, Dipam S. Patel; " Virtual Reality – Opportunities and Challenges"; International Research Journal of Engineering and Technology (IRJET) Volume: 05 Issue: 01

[3] Inrak Choi, Eyal Ofek, Hrvoje Benko, Mike Sinclair, Christian Holz; "CLAW: A Multifunctional Handheld Haptic Controller for Grasping, Touching, and Triggering in Virtual Reality"; Microsoft Research, Redmond, WA, USA

[4] Rasika Dilip Sonje, Revati Ravindra Lokhande, Vaibhavi Vijaykumar Joshi; "Virtual Reality for Real Estates"; Department of Information Technology Volume 8 Issue No.3

[5] Daniel Vogel, Paul Lubos Frank Steinicke ; “AnimationVR - Interactive Controller-based Animating in Virtual Reality"; University of Hamburg

[6] Kjell Brunnströma,b, Mårten Sjöströmb, Muhammad Imranc,b, Magnus Petterssonc, Mathias Johansond ; “Quality of Experience for a Virtual Reality simulator” ; bMid Sweden University, Sundsvall, Sweden

[7] Hassan A. Aziz; "Virtual Reality Programs Applications in Healthcare"; Journal of Health & Medical Informatics

[8] Munehiro Nakazato and Thomas S. Huang; “3D MARS: Immersive Virtual Reality for Content-Based Image Retrieval"; Beckman Institute for Advanced Science and Technology Urbana, IL 61801, USA

[9] Bouchlaghem N., Thorpe A. and Liyanage, I. G.; “VIRTUAL REALITY APPLICATIONS IN THE UK's CONSTRUCTION INDUSTRY”

[10] https://www.explainthatstuff.com/virtualreality.html

[11] https://www.vxchnge.com/blog/virtual-reality-statistics

[12] Stoyan Maleshkov , Dimo Chotrov; “Affordable Virtual Reality System Architecture for Representation of Implicit Object Properties”; Virtual Reality Lab, Technical University Sofia, Sofia, 1000, Bulgaria

 ISBN: 978-81-946375-7-8

Virtual Reality and Augmented Reality in Health Care

Shubham Bansod and Rahul Parulekar
MCA, Deccan Education Society Navinchandra Mehta Institute of Technology And Development, University of Mumbai

Abstract
In the last few years the importance of Virtual Reality has increased due to the voluminous media coverage .In this paper how Augmented Reality(AR) and Virtual Reality(VR) is used in different sections of healthcare is presented and also the various applications of Augmented Reality and Virtual Reality in various domains are included. Virtual Reality speaks to a promising zone with great prospects of enhancing in preparing of health-care experts. It contributes to raise intrigued and inspiration in learners and to viably support skills procurement and exchange, since the learning prepared can be settled inside an experiential system.

This paper has the most point of examining the method of reasoning and main benefits for the utilization of virtual reality in health-care instruction and preparation. Critical re- search and ventures carried out in this field will also be displayed.

Keywords: Virtual Reality, Augmented Reality, Virtual environment, Application of virtual reality And Augmented Reality, Simulation, Training.

I. Introduction

In healthcare technologies Virtual Reality and Augmented Reality are having a huge impact[10]. This application can be used for student to practice surgery. Complicated medical conditions can be explained to the patients using AR and VR. In the field of the AR people have a rich capability to create and provide a rich user interaction experience where people perform drug and medical device virtual experiments.

1.1 Virtual Reality (VR) is the utilize of computer innovation to form a mimicked environment. Not at all like conventional client interfacing, has VR placed the user interior an encounter. Rather than seeing a screen before them, clients are drenched and able to associate with 3D universes. By recreating as numerous faculties as conceivable, such as vision, hearing, touch, indeed scent, the computer is changed into a watchman to this manufactured world. The as it were limits to near-real VR encounters are the accessibility of substance and cheap computing control.

1.2 Augmented Reality (AR) is a technique whereby virtual items are superimposed on this present reality. AR is viewed as a live perspective on physical true condition whose components are converged with augmented computer produced pictures making a blended reality. Just as human variables considerations of AR in health and medical sector. AR should to react continuously with the client's developments in reality and the computer created information is in a perfect world semantically incorporated with what happens in reality. A few distinct methods for coordinating computer created information with this present reality exist. Hand-held transparent presentations and projections are two techniques for overlaying this present reality with virtual items. Some AR frameworks utilize a camera to tape the client in reality and show back the account to the client on a screen, as if it were a mirror with extra data superimposed. It appears that Heads-Up Displays (HUDs) are bound to turn into the prevailing conveyance

technique for AR in future. In present the deployment in commercial market, HUDs will effectively available and generally cheap. The accessibility of this AR equipment will probably inspire VR and AR developers to make and profit by AR in health and medical sectors.

The Virtual Reality operation first performed at the Royal London Hospital in the UK to Removal of Cancerous Tissue Using Virtual Reality [14].

II. Literature Review

Hassan Aziz :- Present paper on "Virtual Reality Application in Healthcare". The term paper focuses on applications in healthcare that illustrate the potential to cognitive address, Useful and psychological disabilities. Virtual reality makes an opportunity for preparing and instruction for clinical practitioners. This paper too covers presentation treatment utilizing virtual reality, Diversion Treatment, Virtual reality in surgical arranging and execution following, Virtual reality as a symptomatic device [1].

Judi Moline :- Present paper on "Virtual Reality for Health Care: a survey" This paper focuses on ongoing and continuous investigate related to applications of virtual environments and related advances within the health-care field. It moreover gives a general introduction to virtual reality particularly because it relates to health. The paper also talk about the application like medical therapy, preventive medicine and patient education, medical education and training, visualization of massive medical databases, skill enhancement and rehabilitation, and architectural design for health-care facilities [3].

Tomasz Mazuryk and Michael Gervautz :- Represent paper on "Virtual Reality History, Applications, Technology and Future" the paper talk about the history of VR And AR, At the starting of 1990s the advancement within the field of virtual reality got to be much more stormy and the term Virtual Reality itself became greatly well known [4]. Ready to listen around Virtual Reality about in all sort of media, individuals utilize this term exceptionally regularly and they abuse it in many cases as well. The reason is that this unused, promising and intriguing innovation captures greater interest of individuals than e.g., computer design.

III. Application of Virtual Reality and Augmented Reality

In market there are various applications which differ in technologies and multimedia to support training in the medical and surgical sector. There is a telesurgical application to interactive and simulator of human body and to virtual world for emergency training[10]. The 3D environments have a full immersion development training platform for psychologists and psychologists 'treatments [6].

3.1 Surgical and Medical Training

The possible categorization of virtual divides and virtual environments medical education based upon individual training medical prototyping and medical cries training. The medical training device for individual medical training are also referred as "partial trainer" [6]. It provides high skill under the simulation environmental with anatomically correct and high realistic. There is a vast discussion of potential of VR-based individual training teaching and task oriented clinical skills. It is one of mostly done research area and there are more common research areas. The simulation for training palpation of subsurface breast tumors, training for arthroscopic knee surgery, virtual endoscopy simulation, simulator for bone dissection and training system for gunshot wound[10].

The recent development and research are virtual environment for esophageal intubation training. The open system surgery simulation developed by Bielserand colleagues is also one of best research development system[6].

Figure-1 Visualizing the brain in VR[13]

3.2 Emergency training

Restorative emergency preparing frameworks center on complex preparing assignments in which the person must act specifically and physically on the environment and in which the reactions to an activity may be exceptionally unpretentious (such as a alter in skin color)[1]. Doing so requires the plan and execution of several interaction and recreation techniques, from distributed-system bolster for high-fidelity recreations to the improvement of clinically realistic virtual patients and, maybe most critically, to the creation of methods that allow a client to act naturalistically upon the virtual environment. a customizable model with practical anatomical highlights speaks to the persistent, whom clients act upon utilizing physical rebellious that interface the mannequin to a computer control framework that drives the fitting clinical state and response. Other analysts created the whole preparing situation by means of program, with energetic virtual patients displaying changing physical condition and responding to the clinician/trainee, who interatomic with the framework by means of an arrangement of menus[1]. Freeman and colleagues as of late created a virtual reality quiet recreation framework for educating crisis reaction abilities to U.S. Naval force therapeutic suppliers [6]. Fast and successful restorative intervention in reaction to gracious and military-related calamities is significant for sparing lives and restricting long-term incapacity.

As suggested by Freeman, “this experiential, problem-based training approach engages the user in a stress-filled, high fidelity world, providing multiple learning opportunities within a compressed period of time and without risk.”[6].

3.3 Surgery

Constancy and realism are to main factors in this fields because the priority is not only for making certain that no unsuitable handling will be learned but also for the comfort of the user recreating stressful condition is also kept in mind but on the other side there are also some cons the purchased cost and maintenance cost is more[12].

Figure-2 Teaching surgery in VR [13]

3.3.1 Laparoscopic Surgery

Virtual reality are differentiated as “hybrid” due to their functionality of having real instruments and virtual visualization projected [12],on display devices feedbacks are received to the user according to their necessity and informative systems are kept to measure parameters like time required for task completion monitoring error during surgery.

The device supports six basic level handling such as camera navigation, knot tying, suturing, cuffing, clipping, hook electrode[12]. It support both the trainings group (team training) and individuals. These device which provides a helpful way to implement new interne problems to have a hands on practical experience of the surgery[12]. It helps to transfer the real surgery procedural to the virtual simulation procedure to advance intern skills. It support 3 level of difficulty with 18 alternative case of laparoscopic cholecystectomy [12].

3.3.2 Robotic Surgery

Training methods are on a rising demand as there is an increase found in robot-assisted procedures. There is a rapid increase in the number of surgeons that are to be trained in robotic surgery [5]. During the construction of training programmer for robotic surgery various training modalities are used. A programmer process starts with knowledge development next step is skills training which is combined with simulation modalities. At the initial phase of robotics surgery bed side assisting, mentoring & proctoring are some roles. To make sure that surgeons have some basic knowledge about robotic surgery an e-learning model is used [5].

Dry lab training due to its higher accessibility it is useful not all skills laboratory could afford a robots are used for training[5]. During the usage of simulation models it’s beneficial that training exercises must be validated and also have goal. Simulator exercises are necessary to have good face and translate well to clinical setting before robotic training programmer is introduced to them. Unfortunately, few reports which are published are used. In recent few years, number of reports of virtual simulators are published [5].

3.4 Pain and physical Therapy

The pain is sometime is a physiological problem. The physical therapy works for the pain in right way if it provided in a right way at a right time [13]. So the VR provides a virtual environment to heal the pain. It makes the human feels they are in happy environment. It is sometimes difficult for human to be in then VR provides pleasant environment to perform pain healing environment. Its real fun and more interactive in the virtual environment. The focus can be greatly enhance in these environment is no distraction. The activity in the VR is best programed to get the best out of it. If help to enhance the recovery and speed in recovery time [13]. The Harbor View Burn Clinic Specialized for the treaties the Physical Burned Patients [13]. The Scares of the patient are grafted so it helps them to heals the scares much faster because of the pleasant mind of for patient.

Figure-3 VR is being used for physical therapy[13]

3.5 Patient Education

Virtual Reality (VR) will be useful for doctors to have a view and examine inside human body of the patient and may show patient about how his/her inside body parts are functioning. It will be helpful for doctors and patients for better treatment during the operations, which may reduce the risk [9]. It will be also helpful for pregnant women for her day to day treatment during the pregnancy. VR may be helpful for doctors to understand the human body before the actual operations. VR may play vital role in the medical field.

3.6 Telepresence

Robotics and VR environment are linked to telepresence application at remote site telepresence system are used in manipulation of medicines SRI International's Grenn Telepresence Surgery System was bulid so that surgeons can perform operations from sites[3]. The operation is done on virtual image. Recently medical school like Jonhn Hopkins are using telepresence[3]. Holds laparoscope for surgeons to monitors and hard-wired connections makes this remote interaction possible[3].

3.7 Mental-Disease professionals

The treatment of mental disease people is very challenging. The realistic approach to treat this patients is very beneficial for that the doctor has to understand what the person in think in his/her mind. The doctor has to see in his mind. There is involvement of lots of sensor are required to simulate what that patients is seeing in his mind and show to the doctor[11]. The patients is provided with virtual environment with the doctor is also provided with the same environment[6]. The doctor and patient need to tag in the same environment to understand the problem of the patient.The treatment of schizophrenia patients is done by the doctor by totally in depth analysis, for that doctor need to look in the schizophrenia patient head [6].

IV. Conclusion

The objective of this research was to understand the impact of Virtual Reality (VR) and Augmented Reality (AR) in the field of Health Care. Virtual Reality (VR) and Augmented Reality (AR) both are getting more attention in the field of Health care. It is historic transformation in the way doctor's and patient interact with the technology. This both technology are gives contribution so that during the operation the chance of the operation failure is become less also, doctors have all information prior of patient. patient will also the idea through education so they clear all their doubts. These difficulties represent challenges to future development involving pedagogical, cultural, economic, interface design issues Also, with the help VR consultation doctors can greatly develop better understanding and modify their medical practices,

V. References

[1] Hassan A. Aziz; Virtual Reality Programs Applications in Healthcare Journal of Health Medical Informatics, ISSN: 2157-7420

[2] Munehiro Nakazato and Thomas S. Huang; "3D MARS: Immersive Virtual Reality for Content-Based Image Retrieval Beckman Institute for Advanced Science and Technology Urbana, IL 61801, USA

[3] Judi Moline United States Department of Commerce National Institute of Standards and Technology Gaithersburg, "Virtual Reality for Health Care: a survey"

[4] Tomasz Mazuryk and Michael Gervautz Institute of Computer GraphicsVienna University of Technology, Austria "Virtual Reality History, Applications, Technology and Future"

[5] HWR Schreuder; University Medical Center Utrecht “Virtual Reality Training for Robotic Surgery”

[6] FABRIZIA MANTOVANI, Ph.D.GIANLUCA CASTELNUOVO, M.S. ANDREA GAGGIOLI, M.S. and GIUSEPPE RIVA “Virtual Reality Training for Health-Care Professionals”

[7] Stephen R. Ellis, What Are Virtual Environments? IEEE Computer Graphics and Applications

[8] Keshavjee K, Clinton J, Troyan S. Unpublished findings.

[9] J. Fischer1 and M. Neff2 and D. Freudenstein3 and D. Bartz1, Medical Augmented Reality based on Commercial Image Guided Surgery.

[10] V. Mehta, Devraj, H. Chugh and P. Banerjee, "Applications of Augmented Reality in Emerging Health Diagnostics: A Survey," 2018 International Conference on Automation and Computational Engineering (ICACE), Greater Noida, India, 2018, pp. 45-51.

[11] K. -. Hsiao and H. F. Rashvand, "Using sensor enabled augmented reality for healthcare," IET Conference on Wireless Sensor Systems (WSS 2012), London, 2012, pp. 1-5.

[12] https://www.intechopen.com/books/medical-and-surgical-education-past-present-and-future/virtual-and-augmented-reality-in-medical-education

[13] https://visualise.com/virtual-reality/virtual-reality-healthcare

[14] https://www.arm.com/blogs/blueprint/ten-ways-vr-ar-transforming-healthcare

 ISBN: 978-81-946375-7-8

Generating Websites with AI - An Analysis of the Sketch2Code Tool

Miss. Snehal Chandrakant Torne
Student, MCA, Deccan Education Society's Navinchandra Mehta Institute of Technology & Development Dadar (W)

Abstract
This paper include computer vision techniques as well as deep learning techniques to work on the sketched wireframes. The mainly focused on how elements can be detected by using Element Detection Method on sketched wireframes which can be done by using the AI deep learning methods. After element detection technique Design process can be done by using the sub two methods Preprocessing of image and Post-Processing of image, In Preprocessing the raw image must be cleaned up before it can be processed and In and Post-Processing it translate the generated code into HTML code. Past ways to deal with address this issue have utilized hand-structured highlights or mechanized highlights removed by profound convolutional neural system models. Right now paper study shows the prediction of images by utilizing the inferential data relying upon the visual substance found in a images. Apparently, this is the principal endeavor to address such issue by utilizing labels anticipated pictures done by the ink2code.Here there much focus on the development sketch2code application which utilize current most recent AI methods AI Orchestrator.

Keywords: Artificial Intelligence (AI),Orchestrator, Custom vision, Neural network,Object detection,REST.

I. Introduction

Artificial Intelligence is an area of computer science which deals with creating intelligent machines that work better than humans. It includes a goals like reasoning, problem-solving, planning machine learning,knowledge representation, natural language processing, perception,motion and manipulation, social intelligence, creativity, and general intelligence. AI also have the capability to learn and think computer programs, these computer programs "take the decisions which normally require a human level of expertise" and help people to anticipate problems or deal with the issues as they come up.

Earlier developing the website with the amazing design is such tedious task we should have the enough knowledge of web developing technologies and much more .All work done by manually by coding each and every page and it is difficult to handle the complex websites.But now it became possible to make changes in one location and it will be available anywhere on the website. By using drag and drop web development technology. Out of the latest technologies that have been in work nowadays is Artificial Intelligence (AI). Artificial Intelligence in Web Development providing the efficient way to design as well as develop the web applications and the websites.Web development using Artificial intelligence can really have a surprising outcome for building web experience.Artificial intelligence uses the machine learning to identify and implement web development trends,technology that can builds the websites on its own. Users just have to state that what they want and the Artificial Intelligence generates a personalized design.

Furthermore,this paper is focus on how deep learning AI methods can be applicable to this taskand how one can easily design websites and web applications in least amount of time period by using the Artificial intelligence.

II. Related work

Now a days the websites look and feel should have a more attractive,user friendly and it must satisfy the customers' needs and the customers' requirements. AI web design is the design of a website with artificial intelligence technology. By using an advanced machine learning algorithm, AI web design tools can provide website that automatically improve their organization needs and business with a variety of options when it comes to designing every page of website.

The web development using Artificial Intelligencefocuses on automatic generationof program source code using AI and Machine Learning algorithms and itanalyze the data which is crucial to all forms of web development .by using this deep machine learning algorithms the system can generate the program source code in human-readable format, And it will help to the humans to understand that code and modify code as they want.For users such as designers and developers, AI and Machine Learning algorithms would save critical time early on a project by enabling ideas quickly, improve iteration cycles, and eventually enable the development of superior apps. As the designers and developers perspective the goal is to save as much time as possible on minor tasks. Most importantly, this would allow designers and developers to focus on what matters bringing value to end-users.

This paper is based on research of converting wireframes sketched on paper directly into code,it provides a number of advantages by overcoming the lengthy process of wire framing and codingwhich needs experienced developers and designers, as well as allowing anyone who can draw a wireframe on paper to create a fully functioning website with no training[1].Further, deep machine learning techniques has shown considerable success in vision problems [4, 5].It is identified a gap in between research of applying deep learning techniques to low delity design to code theapplications based on research [1]. deep machine learning can solve this problem as this is primary vision problem.

III. Literature Review

Techniques

1 Computer vision techniques

Computer vision techniques is used to detect and classify wireframe elements from an image[1],this technique is based on capturing the elements from sketched wireframe and it will help to easily identify the elements from the wireframe.

1.1 Image Denoising

Capturing the elements from wireframe this method uses real world images not the digitally created images as input i.e. images taken from a web camera. These images often contain Gaussian noise due to variations of current over the sensor of a camera [9],this noisy images can give the very poor performance in edgeas such, it is important to reduce this noise.Deep machine learning techniques such as denoising auto-encoders are very effective against noise removal but they are slower then kernel ltering techniques [10].In the gaussian blur ltering technique the gaussian kernel is one of the technique which is use to smooth an image, gaussianblurs technique is highly effectual against gaussian noise,one of the median blur technique is a ltering technique which uses a kernel to preserves the median value in each window. Median blurs are effective against Gaussian noise [11] and preserve edges [12].To develop real time applicationdenoising algorithms are used,edges are an important part of

wireframe symbols and are used as features to detect and classify elements. As such, preserving edges is an important property of our denoising method.

1.2 Colour detection

Colour detection is used in this method to support element detection. Colour detection technique is focused on detection via thresholds as this is an efficient technique for large uniform colour representation.The digital images are represented by using three colour red, green, and blue i.e(RGB) known as a colour space[1],other various colour spaces are present here is saturation, hue and L*A*B or value (HSV) which presents colour in different ways.

1.3 Edge detection

This method is mainly focused on detecting elements from wireframe sketches, because the Wireframe element consist of straight edges therefore edge detection is an important technique for detection of elements.

1.4 Segmentation

In arrange to classify the wireframe components they must be identified first.A wireframe sketch will contain numerous components and thus strategy of recognizing component boundaries is required.There are numerous potential segmentation algorithm but following one is applied for segmentation of elements in wireframe:

Structure based –form detection utilizesthe edge information and boundary following to differentiate segments. These techniques often utilize deep learning methods.structural based techniques give better color invariance.This method uses the contour detection,contours are the curve joining all continuous points with the same colour or intensity along a boundary.

Algorithm Suzuki85 [1] is utilized for fed edges from edge detection method and contour approximation with the Douglas-Peucker algorithm [1] it permits imperfect shapes to be adjusted by approximating a contour over slight mistakes within the shape.

1.5 Text detection

This technique is used for detection of text from sketches, it mainly uses thestroke width transform (SWT) [1] to identify text content. SWT is fast, language independent, lightweight scene text content detector.Deep learning methods,Feature based methods are two methods

2 Machine learning techniques

Machine learning is an effective technique which gives computer systems the power to learn from data without being expresslymodified. machine learning a perfect tool to solve the problems in this domain as there is considerable data and many classi cation and identification tasks. Two key techniquesis used here for this method - multi layer perceptron networks and semantic segmentation networks.

2.1 Deep learning

Deep learning is the eld of machine learning which uses deep neural networks comprising ofnumerous hidden layers. deep learning can be applied to this domain because there is huge amount of data needs to be handle ,a need of most deep learning methods and because this is main vision problem and deep learning has become the standard for top performance in many vision problems.

2.2 Convolutional neural networks

Convolutional neural networks (CNNs)is important to understand semantic segmentation which is core for this method. CNNs are especially valuable over MLPs for image processing tasks as MLPs suffer from the curse of dimensionality due to the full connectivity between neurons and therefore do not scale well for high resolution images[1]. Further, MLPs don’t take into account

the spatial structure of data, so pixels on opposite sides of an image would be treated the same way as pixels close together.

Step1: Breaking down of an image

The first step of breakingthe image into smallerpieces, overlapping images of equal sizes and it can process the one single tile of image in following way:

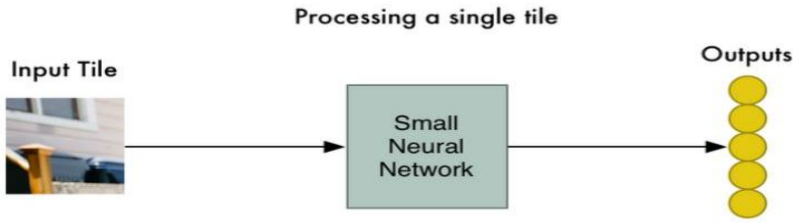


Fig.1:Image Breaking into Tiles[15]

Step 2: Feeding of a tiny image to a small neural network:

In this step, all the tiny images are given to a small neural network that is going to be analyzed and identify if the expected elements is present or not. there will not be any backpropagation reason that all the tiny images are going to be given to the neural network with the same weights. Keeping the same environment between images helps to process them all the similar way.

Step 3: From an array of images to an array of results after processing:

The output of each image fed to the neural network are stored in a new array. Reason of that, the sequence & arrangement of the original tiles is kept. It is necessary to keep the sequence & arrangement of the tiles like the original given image because machine should have to think the same way as humans and give the result faster. The nearest two individual images are, the more related to each other. If images are shuffle the order and sequence at this point, the correlation between the relative position and the relation between pictures would be lost.

Step 4: Downsampling

In this Downsampling step the order has been kept of arrays, the array is divided in 2x2 grid squares and from those small parts, only the most interesting output, being the one with the best similarity to the expected output, is kept.This process helps the machine to reduces the size of the array that it has to process while keeping the important part.

Step 5: Full Convolutional Neural Network

The last step comprises of feeding the reduced array to a fully-connected Neural Network which will be responsible of processing it like any image only composed of numbers for the computer and give the final output result.

Processing of each step

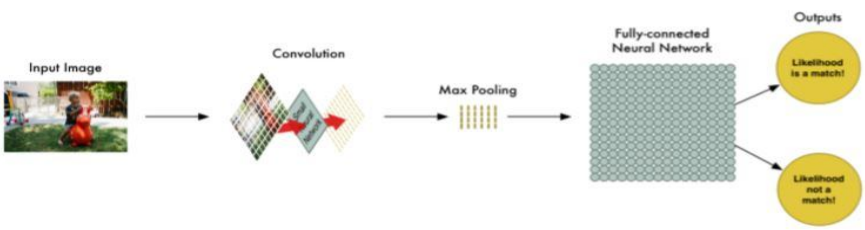
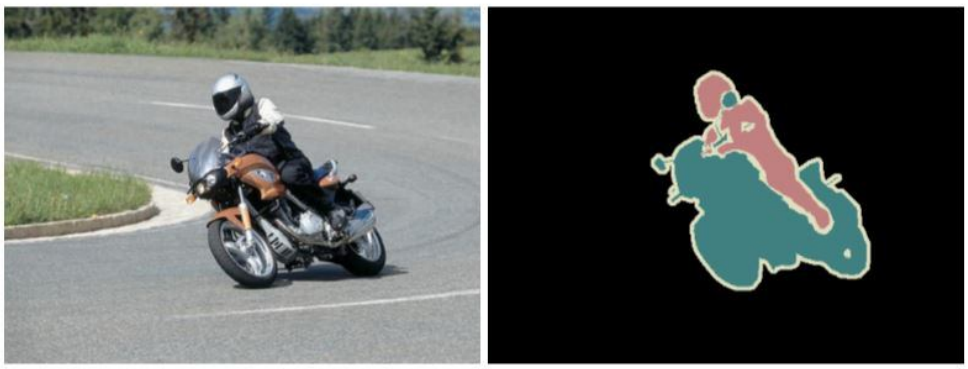


Fig.2[15].

2.3 Semantic Segmentation:

In this technique Image segmentation is a image processing task, the objective is to rearrange the image presentation into something more important and meaningful format. It includes labeling to the each and every pixel such that pixels which share theirfeatures are labeled together. There are a numerous methods such as Texton Forests ,clustering methods and edge detection but more focused is on trainable segmentation networks as these techniques are best performing [14]. trainable segmentation uses trainedArti cial Neural Networks (ANNs) set ofdata forthe image, and a labeled segmented map of the image.



(a) Input Image (b) Segmented map

Fig. 3: Semantic segmentation[1]

There are two main architectures segmentation networks are based onto get around this issue:

• Encoder Decoder architecture

• Dilated / atrous convolutions

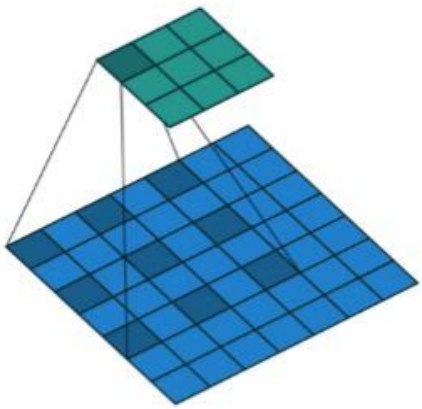


Fig. 4: Dilated/Atrous Convolutions[18]

Semantic segmentation networks uses the large-scale dilatedconvolutionsince they can increment the receptive field of the layer it means the area of the input which the layers can see without increasing the number of parameters or computations. dilated convolution with de ned gaps as shown in gure 4, Dilated convolutions support exponential expansion of the receptive. It means that it allows a larger receptive eld with the similar computation and memory costs as an encoder decoder architecture[1].

Elements

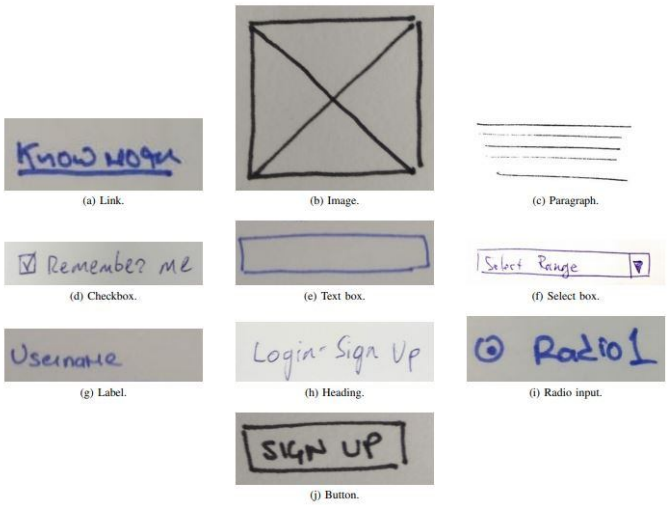


Fig.5: Sample Sketch of all ten classes of Dataset[16]

Detection Method

This section primarily uses computer vision to solve the task of converting an image of a sketch into code. This approach includes four key stages:

1. **Element detection:**Itusesthe computer vision technique to identify and classify the position, sizes and types of all elements from the sketch it isrequired for replicate similar elements in HTML. It can be image,Paragraphs,Titles**,** Buttons,and Inputs.
2. **Structural detection:**In this module there is use of recursive algorithm which used to parsed the input image and returned a list of all elements with their labels, positions, and sizes from previous phase. In this phase the list of the elements taken from previous phase and infer the hierarchical tree structure.
3. **Container classi cation:** In this phase hierarchical tree taken from the previous phase and classify container elements. While HTML uses a number of different tags to represent containers, such as or , and use types which have more semantic meaning. The ve classes of elements: Rows , Stacks ,Forms ,Headers, Footers .

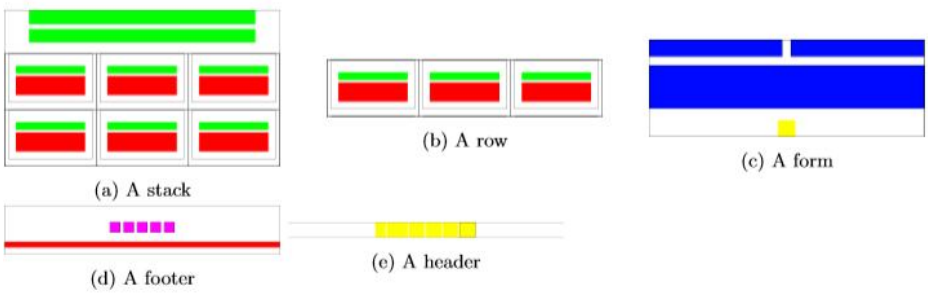


Fig.6: Container classi cation[1].

Identify container can be arranged based on structural features. Structural features contains the containers bounding box with size and position which is relative to the page. It is necessary to use relative positions and sizes so that elements can be compared between wireframes with different dimensions.

4. **Layout normalization:** From the previous phases tree hierarchy of elements is taken. All leaf nodes have element types and all branch nodes have container types[1]. since, sketched wireframes many times contain human errors,types of errors are:Rotation, Imperfect shapes, Translation, Scaling.

Design Process

1. Preprocessing

Preprocessing is required to translate an image from a camera into an image which can be fed into the experimental method. Due to the positioning of the camera or lighting conditions the raw image must be cleaned up before it can be processed.

2. Post-Processing

The output of the experimental method is code which represents the structure of the wireframe. There are three post-processing steps: 1) distribute the generated code from the experimental approaches to clients, 2) translate the generated code into HTML, and 3) live update the website.

```
{
    'type': <the element type>,
    'x': <relative x position>,
    'y': <relative y position>,
    'width': <the element width relative to the page width>,
    'height': <the element height relative to the page height>,
    'left_padding': <relative left padding>
    'top_padding': <relative top padding>
    'contains': [<elements>],
}
```

Fig.7 :JSON Tree Post-Processing result [1].

Figure 8: JSON tree-like structure used to represent structure of a wireframe. This can be directly translated into HTML.

IV. Sketch2Code

Sketch2Code is One of the most popular application of Microsoft AI which converts the sketched wireframes into HTML code.Sketch2Code is a web-based offering that uses machine learning to turn handwritten designs into working HTML code,The application uses various machine learning elements, including computer vision.It uses AI to translate hand-written drawings to working HTML prototypes,designers share ideas on a whiteboard, then changes are reflected in the browser instantly.

Early step in creating an application or any website, first designers should have to sketch a wireframe on paper blocking out the structure of the interface. When designers converting their sketched wireframe into source code they will face more challenges because converting wireframes into code is a complex process, this often involves passing the created application design to a developer and then developer should have to implement the graphical user interface (GUI). This work is very costly and time consuming for the developer. An application which converts wireframe sketches directly into code which have considerable more benefits, Faster iteration a wireframe can move to a website prototype with only the designers involvement, Accessibility allows non developers to design and create the applications, Removes requirement of developer for initial prototypes design of application and allowing developers to only focus on the application logic rather than GUI code.

4.1. Process Flow

The process flow for generating code from the sketched wireframe:

1. The users can uploads an sketched wireframe images through the website.
2. A custom vision model which is present in the architecture can predicts the HTML elements are present in the image and their position on image.
3. A handwritten text recognition service reads the text inside the predicted elements from custom vision model.
4. A layout algorithm uses the spatial information from all the bounding boxes of the predicted elements to generate a grid structure that accommodates all elements [17].
5. An HTML generation engine utilize all this information to generate HTML markup code reflecting the result.

4.2. Architecture

The Sketch2Code solution is made up of Custom Vision Model, Computer Vision Service, Azure Blob Storage, Azure FunctionMicrosoft components

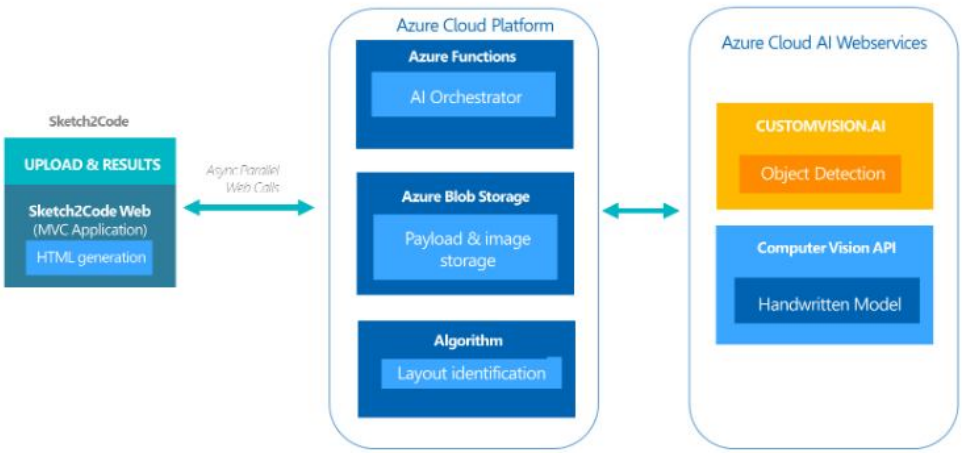


Fig.8 Architecture of Sketch2Code [17].

4.1.1 Azure Custom Vision Service :The Azure Custom Vision Service easily build custom image classifiers, without requiring any knowledge of the fundamental machine learning models. It gives web interface that can help to upload and tag training data, as well as to train, analyze, and make predictions. It can also interact with custom models via a REST API. Custom Vision service recently included enhancementof Object Detection, which allows to train the service to identify and detect numerous objects inside an image with their positions.

4.1.2 Object Detection: Image classification with positioning is the process that takes an image as input, and outputs a label and also draws a bounding box around the object to find it in the image.Object Detection goes one step further, and can identify multiple objects in an image. This has multiple well known applications including face detection, or autonomous driving.

4.1.3 Tag an image: Custom Vision web interface is used to tag images.

4.1.4 Azure Blob Storage : An Azure Blob Storage account is used to store all the mediator steps of the orchestration process. A new folder is created for each step with contents Original.png, Slices, results.json. Original.png:The original image uploaded by the user, Slices :This step contains the cropped images that can be used for text prediction, results.json:This step processes the results from the prediction processrun against the original image.

V. Artificial Intelligence Orchestration

With General Artificial Intelligence still some years away, most Machine Learning algorithms all have niche focus and do small jobs, and thus to achieve use case may need to use and coordinate between multiple artificial intelligence agents to achieve your goal.

Artificial Intelligence Orchestration is the combination of Neural Network algorithms through logic. Utilization of purpose-built algorithms puts humans in control and allow machines to think about small things without the global risks of man/machine war and robotic overlords relatedwith General AI and computers developing consciousness.

Azure AI Orchestrator

The azure functions serves as the backend entry point that coordinates the generation process by interacting with all the services[17]. In the azure functions AI Orchestrator plays the vital role to interact with the services, Orchestrator functions describe how actions are executed and the order in which actions are executed.

VI. Conclusion

The goal of this paper is the analysis of creating an application which translates a sketched wireframe into a HTML code to design website, and to explore how deep learning differentiate to computer vision methods for this task.

By analyzing Sketch To Code technique we came up with the conclusion on that there are number of techniques are available to translate the sketched wireframe into HTML code from them one of them is most popular and fastest technology is Microsoft Sketch2Code,This technology help designers to convert our sketch images into HTML code by using their Azure Functions with AI Orchestrator and it also uses the Azure Blob Storage to store their dataset. AI orchestration has an instance identifier and reliable.

Along with that we also analyzed the applications similar to the Sketch2Code technology one of them is uizard is the application which also converts the sketched wireframe into HTML code by using the pixTocode technique, the application translated code of UI is now supportedin Angular as well as React framework. There are some applications similar to Sketch2Code which is Pix2Story and Pix2Music in this application the preprocessing is same as the Sketch2Code technology but post-processing is as per their application technique.

Reference

1. Robinson, A. (2019). Sketch2code: Generating a website from a paper mockup. *arXiv preprint arXiv:1905.13750*.
2. Landay, J. A., & Myers, B. A. (2001). Sketching interfaces: Toward more human interface design. *Computer*, *34*(3), 56-64.
3. Seifert, J., Pfleging, B., del Carmen Valderrama Bahamóndez, E., Hermes, M., Rukzio, E., & Schmidt, A. (2011, August). Mobidev: a tool for creating apps on mobile phones. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (pp. 109-112).
4. Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, *38*(2), 295-307.
5. Varadarajan, B., Toderici, G., Vijayanarasimhan, S., & Natsev, A. (2015). Efficient large scale video classification. *arXiv preprint arXiv:1505.06250*.

6. Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). Show and tell: A neural image caption generator. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3156-3164).

7. Karpathy, A., & Fei-Fei, L. (2015). Deep visual-semantic alignments for generating image descriptions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3128-3137).

8. Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*.

9. Boyat, A. K., & Joshi, B. K. (2015). A review paper: noise models in digital image processing. *arXiv preprint arXiv:1505.03489.*

10. Xie, J., Xu, L., & Chen, E. (2012). Image denoising and inpainting with deep neural networks. In *Advances in neural information processing systems* (pp. 341-349).

11. Shukla, H. S., Kumar, N., & Tripathi, R. P. (2014). Gaussian Noise Filtering Techniques using New Median Filter. *International Journal of Computer Applications*, *95*(12).

12. Arias-Castro, E., & Donoho, D. L. (2009). Does median filtering truly preserve edges better than linear filtering?. *The Annals of Statistics*, *37*(3), 1172-1206.

13. LeCun, Y., Touresky, D., Hinton, G., & Sejnowski, T. (1988, June). A theoretical framework for back-propagation. In *Proceedings of the 1988 connectionist models summer school* (Vol. 1, pp. 21-28). CMU, Pittsburgh, Pa: Morgan Kaufmann.

14. Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & Garcia-Rodriguez, J. (2017). A review on deep learning techniques applied to semantic segmentation. *arXiv preprint arXiv:1704.06857*.

15. Cheng, J. (2018). Artificial Intelligence and auto-generation of code.

16. Jain, V., Agrawal, P., Banga, S., Kapoor, R., & Gulyani, S. (2019). Sketch2code: Transformation of sketches to ui in real-time using deep neural network. *arXiv preprint arXiv:1910.08930*.

17. https://aischool.microsoft.com/en-us/services/learning-paths/sketch2code/sketch2code-lab/sketch2code-introduction

18. Dumoulin, V., & Visin, F. (2016). A guide to convolution arithmetic for deep learning. *arXiv preprint arXiv:1603.07285*.

 ISBN: 978-81-946375-7-8

Talk To Me

Rushab Ambre
Student, Master of Computer Application, Navinchandra Mehta Institute of Technology and Development,Dadar (W)/University of Mumbai, Mumbai

Saifuddin Harianawala
Student, Master of Computer Application, Navinchandra Mehta Institute of Technology and Development,Dadar (W)/University of Mumbai, Mumbai

Dr. Rasika Mallya
Associate Professor, Master of Computer Application Department, DES's Navinchandra Mehta Institute of Technology and Development, Dadar (W)/University of Mumbai, Mumbai

Abstract
NLP is growing in the worldwide using it in the service base process giving better results. So in this research paper what is NLP, how it is leveled and dataflow of NLP is depicted. The manual analysis is time consuming and the response time is also slow. So with the help of NLP, AI and ML we can provide better and faster response is stated in this paper. The paper mostly focuses on the health sector and education sector. The education sector has many positive and good results such as feedback analysis and descriptive answer evaluation model. The paper provides our view of chatbots in the health and education sector. Chatbots provide better and faster response in real time. The real time response is important and mandatory in the health care sector. The healthcare sector is critical area to implement technology so more research is going on in this sector for betterment.

Keywords: Natural language processing (NLP), Artificial intelligence (AI), Machine Learning (ML)

I. Introduction

Natural language processing is the term used to explain the process of providing and identifying the elements of day to day life and extracting, understanding the beautiful insight with hope of computer algorithm. The computer algorithm process the unstructured data or written data The NLP is branch of computer science which requires many other skills of computer science such as Artificial intelligence computational linguistics and machine learning disciplines. The NLP is a process which is focused on providing automated response based on understanding and processing voice recognition service.

The algorithm also tries to mimic human like responses. Many systems are such that provide responses on learning from the previous reabsorbing feedbacks. NLP is an effective way to provide student to learn the scientific learning. It not only helps in learning it also helps in understanding and increasing the academic grades.[1] NLP assist in developing effective process of learning in the educational setting by developing scientific approaches which will help in using Internet and computer in the effective manner. The approach in NLP is more focused on development of educational software systems and educational strategy that can assist in utilizing the natural languages for education for example e-ratter and Text Adapter [2].

There is also a larger clinical health records in database that will provide a dramatic effect on the development health industry. Owing to the lager amount of free text documentation now available EHRs (e health records or HER) there has been constant increase in research to advance Natural Language Processing methods and applications for the clinical

domain[3].Further more enhancements in the NLP development show a development of open source NLP software are specific tailored with the help of Artificial intelligence and Machine Learning to specifically process the clinical data which leads to increase in the adoptability. The maturity in the development of NLP leads to the successful development of complex clinical resources.

The major aim of study is to:

- Analyses and understand NLP
- The objective to study the use of NLP in the educational and health sector.

II. Literature Review

Architect level of NLP:

1. Phonology

Phonology is the part of linguistics which termed as to the systematic arrangements of sounds. '**Phono**' means voice or sound and logy means word or speech .The phonology is concerned with function, behavior and organization of sounds as linguistic items [4].

2. Morphology

There are different parts in the words the small unit which are termed as morphemes an example of it could be the pre installation. The words Morphemes are '**pre**' which is prefix and '**tion**' is suffix and root word is installing. The interpretation of morphemes of the words is staying same across all the words.

3. Lexical

In NLP system or in humans the interpretation of the individual words are processed at word level understanding-the things begins with the parts of speech. The each word has different part of speech in the context of the statements. In the NLP system the various deployed mode has various nature of representation of words.

4. Syntactic

The Syntactic level focused on the structure of the grammatical sentence. The parser and grammar are checked at this level. The output and the syntactic are to process relationship between the words. There are various parsers for various grammatical models. Syntax of many languages convey the meaning in order and dependency contribute to connotation For e.g. "**Lion chased deer"** and "**The deer chased the lion**" they are different only in terms of syntax yet convey different meanings.

5. Semantic

In this level many people think of these level determine the meanings of the sentence. But this is not the case in this each word is individual mangled and the meaningful meaning is taken out.

6. Discourse

The above two levels Syntactic and Semantic determine the meaning till a sentence is framed. The discourse level work on the multi-sentence level .There is two common approaches:

Anaphora Resolution - Anaphora Resolution is the replacing of words such as pronouns which are semantically stranded with the pertinent entity to which they refer [4].

Text Structure Recognition – adds to the meaningful representation of the text [4].

Generally it is classified as the three different levels of the linguistic analysis:

- Syntax: Analysis the grammatical part.
- Semantics: Meaning the text is identified.

- Pragmatics: The actually context of the text is identified and its purpose.

There are two mechanisms of NLP

- Natural language Understanding
- Natural language Generator

Natural language Understanding tries to identify and understand the text given .It is the process of automatically producing text from the structured data in a readable format with meaningful phrases and sentences [5].

Big data for NLP

Around 80% of the data is available in the raw format [5].The data is stored in big organization which further becomes information. Though the organization human uses different language which is ambiguous and unstructured to be interpreted by computer system .NLP solve many significant problems using big data. It can be in any sector retail, healthcare etc. Normally now each and every organization has to implement big data kind of operation. The data generated by medium and high level organization is huge and can be from any sector. A pharmaceutical company or any low level firm they have to keep track of the huge amount of data of the customer such as documents etc. They are normally Natural human language documents.

Today many interactive organizations have been developed by using NLP technologies.in banking or any retail self service sector various tools have been generated to automate the translation. User can now ask question in his own natural language there is no specific language barrier to it. And it normally reduces the task of manual call assistance and provides the response.

Big data has leverage the performance of business intelligence with the NLP. Now the knowledgeable person is not restricted to search query with specific keywords. They can ask normal query in their own language NLP provides a very short summary on the content and provides the inference. It provides the real time response with speedup processing and faster response.

AI for NLP

The term AI or machine intelligence which gives the leanings and perception to such specific task like chess game proving mathematical theorem etc.AI libraries provided a best success chances for the problem statement. The AI system focuses to design and implements the best answer for the given problems. NLP plays a very crucial role in the AI Data processing with AI. It can process the meaning of language and answer question. But it fails to understand the meaning of words.AI allows user to comment in their own words with the computer.

Chatbot

The working of Chatbot's the user has to enter the query .The data of query is send to the chatbot machine learning engine. The Engine return the phrased data related to the relevant data. The data is provided with the response to the chatbot. The chatbots are built upon the domain knowledge specific .It help to build better chatbots then the generic chatbots .These chatbots are used in the education, healthcare and service or customer service industry etc. The normally user input is provided to the semantic mapper, it mapped to the semantic elements. There might be conflicts in the generated phrase. But this data is provided to the conflict meditator to resolve the conflicts. If there is no further conflicts or the conflicts are resolved are further provide to the topic chatbots a] Basebots b] Dominated and c] repbot. According to the survey results, conversational basebot yield better faster results and satisfactory results.

A lot of factors that affect the understanding of a chatbot and may factor itself become challenging in the natural language processing are [9]:

- Synonyms, homonyms, slang
- Misspellings
- Abbreviations
- Omitting punctuation rules
- Different accents

Implementation a machine learning chatbot includes following stages:

- **Business logic analysis**

It is required so that the developer's team can understand the client's needs.

- **Channel and technology stack**

In order to create a voice chatbot, it is good enough to use the Twilio platform as a base channel and in the case of text chatbots the Telegram, Viber, or Hangouts can be easily used.

- **Development**

The making of the machine learning chatbot consists of two steps which include: the development of client-side bot and providing chatbot connection to the provider's API (Telegram, Viber, Twilio, etc.).

- **NLP integration**

After development, it can be integrated with artificial intelligence.

- **Testing**

Finally bot can start asking the questions as per the training provided to answer for. As there are not many scenarios to be validated and verified so manual testing can be done.

Artificial intelligence chatbot can attract more users, save time and raise the status of site. Therefore, the more users are attracted to the website, the more profit is obtained [9].

Applications

1] E-Paper Evaluation:

Normally to test online or in the computer systems are conducted the MCQ to evaluate the marks of the students. But it failed to provide the analytical or summarized answer evaluations.

To get better insights about the students' knowledge, problems solving skills and methodology it evaluator must analysis the full detail answers. Or that it requires the best word processing tools. The modern word processing tools are supported with machine learning algorithm. The machine learning algorithm is used to evaluate the grades of students on the particular subjects. Attempts to build an automated essay grading system dated back to 1996 when Ellis B page proved on the **'Phi delta kapppan'** that a computer could do as well as single human judge [6]

The Evaluation of descriptive answer has three techniques:

A] Statistical process.

It is the poor process because it only analysis or just match the words in the context .It failed to match the synonyms in the student answer context .the also can't deal with the lexical variability.

B] Information extraction process.

In this process the context is analyzed for the structure data in it. The extracted structure data is that match with the dependencies between concepts means the sentence dependencies are analyzed and the relationship between the sentences is depicted .Afterwards whole analysis result is evaluated by the human expert for the rating or marking.

C] NPL based process

These basically an interaction between the computer and the human written context [7].the NLP has two techniques used for generating summary. The Extracting techniques is the where the whole document of summary generated by selecting the some sentences from the document only. The Abstractive summarization is the just invert of the extractive techniques process. The abstractive generate totally new summary it make sure that it holds the meaning of the old context [8].

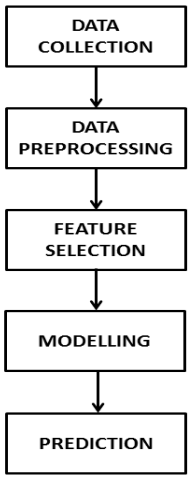


Fig-1 NPL based process

2] Feedback Evaluation and assessment by professor:

The most of the modern Scholl have automated there feedback collection process through online .The sentimental analysis is done for evaluation of student feedback. The feedback approaches sentiment analysis is basically three main approaches a) lexicon based b) Machine learning based c) hybrid. The modern NLP evolved an evaluation drive discipline, empirical, which state that the feedback evaluation has to wait till less time and instant results are provided by the system. The Manual evaluation has very hectic process and time consuming. The manual evaluation has two limitations the result is inconsistent and it is slow process. The humans are notoriously in their judgment they are inconsistent about in what is bad and what is good they sometime may be basic towards the result. So it's best to use the Modern NLP approach.

3] Chatbot

A chatbot is a computer program that communicates with a customer via textual or sound methods. Such programs are often designed to help clients via websites or via phone. The chatbots are generally used in messaging applications like Slack, Facebook Messenger or Telegram. They can be used to order food, buy tickets, or show the weather podcasts. [9]

Examples of chatbot use cases:

- **HealthCare Sector**

Natural language processing chatbot can help in booking an appointment and specifying the price of the medicine [10].

Examples include:

Sensely: is an easy access to medical advice, Sensely gathers information from the patients and refers patients with an appropriate next course of action. Patients can use the bot all by themselves and get required guidance on self-care or suggestion on service. Sensely is obtained with verbal and textual conversational capabilities; this makes Sensely as a unique option to be used. Basic health conditions such as blood pressure can be monitored by connecting this bot to Bluetooth enabled health devices

Ada: Follows similar rules, where patients can mention their issues and symptoms to receive information on the possible condition. Ada is very highly ranked by users and researchers regarding its accuracy; can help the users in managing to get an appointment at their preferred medical center.

Babylon: It is a doctor in your hand. Babylon is programmed for checking symptoms through video consultation and provides recommendations for personal appointments, it can do it all. The bot is a clear presentation of what can be achieved by careful utilizing AI Chabot's in healthcare. And it is only available in London.

- **Education Sector**

Botsify a chatbot lets teachers around the globe to teach their students online using chatbots. Students can converse with the chatbot on a specific topic; learn that topic by text and images or by just watching a video. After learning a topic, students can pursue some quizzes and submit their results to their respective teachers assigned. Teachings with chatbots are not an easy task.

Duolingo is the most popular language learning chatbot platform in the U.S (and possibly around the world). In order **to understand the user context and respond to users contextually and uniquely this chatbot platform is fully equipped with AI algorithms to understand meaning that different users get a different response for a similar inquiry.**

Andy the chatbot is the new English teacher to practice English. Andy can give lessons in grammar, expanding vocabulary with a built-in dictionary, and even providing a simple interface for casual discussion. The Andy chatbot app is available for **Android** and **iOS users [10]**.

III. Conclusion

NLP is strongly growing in many sectors .In these paper some feature of health and education sector are analyzed. Many current Chatbot in education and health sector are trending which are stated here. The NLP is providing and will provide much useful insight for the business and service base industries. The NLP has the best used or automated response it may be contextual or by the voice response. Paper also overviewed sentimental analysis with machine learning, artificial intelligence and NLP

IV. References

[1] Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. Journal of the American Medical Informatics Association, 18(5), 544-551

[2] Burstein, J. (2009). Opportunities for natural language processing research in education. In Computational Linguistics and Intelligent Text Processing (pp. 6-27). Springer Berlin Heidelberg. http://link.springer.com/chapter/10.1007/978-3-642-00382-0_2#page-1

[3] Velupillai, Sumithra & Suominen, Hanna & Liakata, Maria & Roberts, Angus & Shah, Anoop & Morley, Katherine & Osborn, David & Hayes, Joseph & Stewart, Robert &

Downs, Johnny & Chapman, Wendy & Dutta, Rina. (2018). Using Clinical Natural Language Processing for Health Outcomes Research: Overview and Actionable Suggestions for Future Advances. Journal of Biomedical Informatics. 10.1016/j.jbi.2018.10.005.

[4] Natural Language Processing: State of The Art, Current Trends and Challenges - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Phases-of-NLP architecture_fig1_319164243 [accessed 21 Jan, 2020]

[5] https://www.xenonstack.com/blog/artificial-intelligence/

[6] Page, Ellis B."The imminence of... grading essays by computer." The Phi Delta Kappan 47.5(1966)

[7] K. Meena and L. Raj, "Evaluation of the descriptive type answers using hyperspace analog to language and self-organizing map," 2014 IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, 2014, pp. 1-5.

[8] Rahman, Md & Siddiqui, Fazlul. (2018). NLP-based Automatic Answer Script Evaluation. 4. 35-42.

[9] https://sloboda-studio.com/blog/how-to-use-nlp-for-building-a-chatbot/

[10] https://mobisoftinfotech.com/resources/blog/how-ai-chatbots-revolutionizing-healthcare/

 ISBN: 978-81-946375-7-8

Analysis of Face Detection System using MATLAB

Sejal V Tamore
Student, MCA, Navinchandra Mehta Institute of Technology and Development

Abstract
With the increase growth of Digitalization and multimedia, Face Detection System(FDS) has been getting lot of attention. FDS helps in the security of our data that needs to be kept secure and confidential. Nowadays there are various biometric techniques like fingerprints, iris, palm, voice, face etc. The main objective of this paper is to identify whether the person is present or not, or instead some image is shown of that person to gain access to the system. Security is the important part of this paper. Due to which the reliability will be increased along with the security of the system.

Keywords: Digitalization, Biometric, Security, Face Detection System(FDS)

I. Introduction

Over a past few year, there is a wide demand for detecting a face for security purpose, from then face detection has been in the limelight and it has been one of the most successful system for analysis and understanding. With the help of FDS there is a possibility that the task can be achieved faster for detecting a face, then it will be helpful for various applications like security system, Attendance system [1]. Face detection is a stepping stone to all facial analysis algorithms, including the face recognition, face verification/authentication, facial expression tracking and many more. So that computers can understand face clearly. Given a digital image or a real person ahead of camera, the primary goal of face detection is to determine whether or not there are any faces in front of the camera [2]. This technique has become the most promising technique for detecting individuals in recent years since, instead of allowing the people to access the physical and virtual domains based on passwords, PINs, smart cards, plastic cards, tokens, keys and so, this method examine an individual's behavioral characteristics in order to detect his/her face. Passwords and PINs are difficult to remember and can be stolen; cards, tokens, keys and the like can be misplaced, forgotten, or duplicated. However, an individual's behaviors cannot be misplaced, forgotten, stolen [3]. Comparing between a real person in front of the camera and an image of that person on a phone to identify the difference using FDS is discussed in this paper. MATLAB R2020a is used for comparing the real person in front of the camera and an image of that person on a phone which can detect the face and then show it in a graphical manner.

II. Literature view

Ashu Kumar · Amandeep Kaur · Munish Kumar(2018) has presented survey on different techniques for face detection in digial images and also explaining various challenges and application for face detection and at the last various databases are given with their features for face detection [2].

Divyarajsinh N. Parmar, Brijesh B. Mehta(2013) has described the common methods like holistic matching method, feature extraction method and hybrid methods, also had discussed about the application, examples and the future directions of face detection [3].

Gurpreet Kaur, Manbir Sandhu, Purnima(2016) has reviewed and provided a survey of still based face recognition that provide some internal study of machine recognizing the faces and has also discussed the issues, techniques and application [4].

Sandeep Kumar, Sukhwinder Singh, Jagdish Kumar(2017) has provided a survey of different face recognition techniques for knowing the age and gender of a person it has also highlighted on the future directions [5].

III. Process framework:

This is the actual implementation of the system in MATLAB which detects the face in front of the camera and plots the graph which can then be easily used for analysis.

Creating the graph of the face after detecting the face from an image using MATLAB R2020a Application:

Step 1: Open MATLAB R2020a Application.

Step 2: Get the list of webcam connected to the system and store it in variable C

Step 3: Create a webcam object "cam" where device number is a numeric scalar value that identifies a particular webcam by its index number. When you use the webcam function with an index as the input argument, it creates the object corresponding to that index and connects it to that camera.

Step 4: Create a preview window that displays live video data for the webcam object "cam".

Step 5: Create a face detector object. The CascadeObjectDetector uses the Viola-Jones algorithm to detect people's faces.

Step 6: Use method snapshot() to acquire frame as a single image from the camera "cam" and assign it to the variable I.

Step 7: Use method step() which contains the faceDetector object and the variable I.

Step 8: Use method imwrite() to save the image to the disk.

Step 9: Use method imhist() to display the histogram of the image.

Step 10: Use method closePreview() which stops the image acquisition object cam from previewing, and use method clear() to remove all the variables from the current workspace, releasing them from system memory.

Step 11: Use method insertObjectAnnotation() which contains the image that inserts rectangles and labels at the location indicated by the position matrix and assign it to the variable IFaces.

Step 12: Use methods figure, imshow() and title to create a new image window and to show the image with the title Detected faces.

 ISBN: 978-81-946375-7-8

Editor - C:\Users\dell\Desktop\ResearchPaper\Test\FDS.m

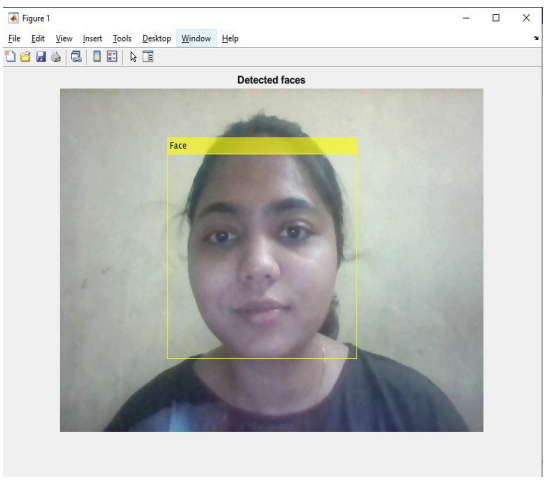
FDS.m

```
C = webcamlist;
cam=webcam(C{1});
preview(cam);
NotYet = false;
faceDetector = vision.CascadeObjectDetector;
while ~NotYet
pause(2);
I = snapshot(cam);
disp('took a snapshot. checking to find a face ....')
bboxes = step(faceDetector, I);
if ~isempty(bboxes)
NotYet = true;
disp('face found!');
% Save this image to disk.
fullImageFileName = fullfile(pwd, 'myfirstimage.jpg');
imwrite(I,fullImageFileName);
img1 = imread('myfirstimage.jpg');
% Display Histogram of Face
imhist(img1)
disp('Histogram of Face');
break;
end
disp('no face detected :(, repeating...');
end
closePreview(cam);
clear('cam');
IFaces = insertObjectAnnotation(I, 'rectangle', bboxes, 'Face');
figure, imshow(IFaces), title('Detected faces');
```

Figure no 1: Steps of detecting a face from an image and plotting a graph

Graphical Representation

1) Photo of a real person in front of a camera and its graph



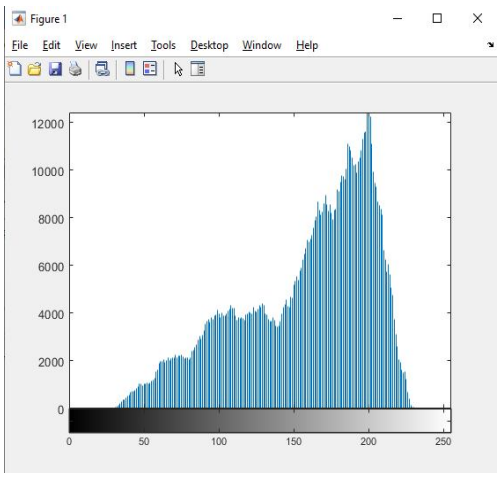


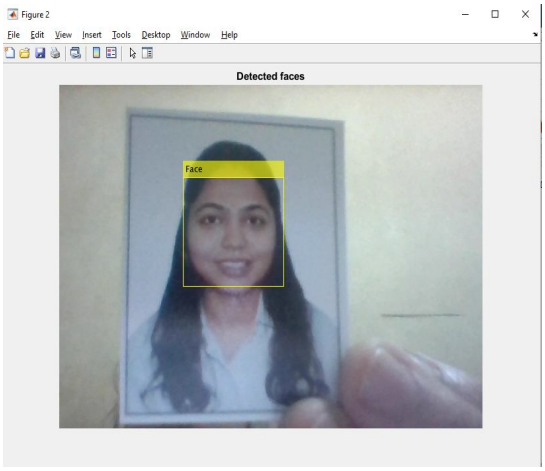
Figure no 2: Photo of a real person in front of a camera

Figure no 3: Graph of a real person in front of a camera

Figure 2 shows that the person is present in front of the camera and Figure 3 shows that on detecting a face the frequency changes as there is a change in the pixel of the image.

2) Image of a person kept in front of a camera and its graph:

 ISBN: 978-81-946375-7-8



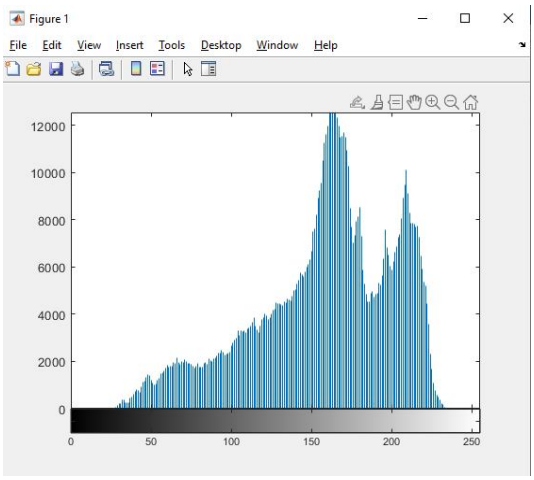


Figure no 4: Photo of a person in front of a camera

Figure no 5: Graph of a person's photo in front of a camera

Figure 4 shows that there is an image of the person present in front of the camera and Figure 5 shows that on detecting a face the frequency depending on the pixel is been changing from time to time.

So, from Figure 3 and 5 it is clear that the system finds different frequency due the pixel and due to that the graph is different when the person is present in front of the camera and an image of that person is present in front on the camera.

For Examples: To enhance town center surveillance in Newham Borough of London, it has 300 cameras linked to the CCTV controller room. The city council said that "the technology has helped to achieve a 34% drop in crime" since the system was linked . Similarly in different countries this systems are in place like in Birmingham, England [3].

To eliminate duplicates in a voter registration system because there are cases where the same person was assigned more than one identification number. The face detection system compares the face images of the voters to differentiate one from the others. If two faces are highly similar face image, then manual review will be required as to avoid duplicity and to make sure they are different persons [3].

In 2014, Facebook launched its DeepFace program, which can detect whether two photographed faces belong to the same person, with a 97.25% accuracy rate [6].

In June 2015, Google launced its FaceNet going one step ahead of facebook. By using the more useful Labeled Faces in the Wild (LFW) dataset, With a new record FaceNet achieved 99.63% accuracy [6].

IV. Applications

Application Areas	Specific applications
Entertainment	Virtual reality, training programs, robot training.[4]
Smart cards	E-commerce, driving license , passports, voter cards, aadhar cards, pan cards[4]
Information security	Parental control, mobile device login, Application security, Encrypted file Internet access, medical records, Secure trading.[4]

Surveillance and Law enforcement	Access to restricted areas like banks, military sites, airports and law enforcement premises. CCTV control and surveillance. Suspect tracking and investigation.[4]
Education	Biometric Attendance System, Examination System

V. Future enhancement

In the near future, detecting the face in night can be done by improving the system and also detecting a particular person from a group of person to know the accurate result even if the person alone is not there in front of the camera. The straightforward future for face detection is to further improve it in presence of some problems like face occlusion and non-uniform illumination [2]. Facial biometric are provided by many recent companies in mobile phone for purpose of access. In near future it can be used in different fields like payments, security, healthcare, advertising, etc. [2]. The face detection system can also be modified to identify various properties from the image like gender, age, beard, moustache, scars, hair, skin color, glasses, facial marks, tattoos etc. From these different attributes the system can extract data and compare this data with the image already stored in database to determine a proper identity [5].

VI. Conclusion

In the fast few years face detection system has developed an attention from different kind of groups and also it has improved in couple of past years. From the above analysis in this research paper, it can be stated that the face detection system (FDS) can differentiate between a real person and an image of that person. But there is a possibility that sometime it gives access to the wrong person by getting fooled. So it will need a more proper system so that it can detect and give proper and accurate result.

VII. References

1. Maneesh Ayi, Ajay Kamal Ganti, Maheswari Adimulam, Badiganti Karthik, Manisha Banam, G Vimala Kumari(2017): Face Tracking and Recognition Using MATLAB and Arduino, IJRASET | Volume 5 Issue XII | ISSN: 2321-9653

2. Ashu Kumar · Amandeep Kaur · Munish Kumar(2018): Face detection techniques: a review, DOI: 10.1007/s10462-018-9650-2

3. Divyarajsinh N. Parmar, Brijesh B. Mehta(2013): Face Recognition Methods & Applications, IJCTA | Vol 4 (1) | ISSN:2229-6093

4. Gurpreet Kaur, Manbir Sandhu, Purnima(2016): Facial Recognition: Issues, Techniques and Applications | IJARCSSE | Volume 6, Issue 2 | ISSN: 2277 128X

5. Sandeep Kumar, Sukhwinder Singh, Jagdish Kumar(2017): A Study on Face Recognition Techniques with Age and Gender Classification | DOI: 10.1109/CCAA.2017.8229960

6. Facebook and Google launch's DeepFace program (in 2014) and FaceNet (in 2015) (16 February 2020) https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/ biometrics/facial-recognition

 ISBN: 978-81-946375-7-8

Review on Security in Bigdata

Shazia Moinuddin Shaikh Topiwala
MCA, DES's NMITD, Mumbai, India

Swapnali D. Mahadik
Assistant Professor, MCA Department, DES's NMITD, Mumbai, (Affiliated to University of Mumbai)

Abstract
Big Data originated new issues related not only to the amount or the range of the information, but also to data security and privacy. This paper provides a literature review on the necessity of security and privacy problems with various big data applications. the first section gives a fast description of massive data. The second section reviews the numerous big data applications and, hence, explains the importance of privacy and security of massive Data.

Keywords: Big Data, Volume, Security, Privacy.

I. Introduction

The provision of huge data and their use for several confidentiality - sensitive and privacy - sensitive tasks make, however data security and privacy both are critical requirement. as an example, the provision of multiple datasets, which will be easily combined and analyzed, makes very easy to assemble sensitive information. Pervasive data gathering from multiple data sources and devices, such smart electric meter, smart phones etc., further exacerbates the matter of knowledge privacy. Pervasive data collection often aims at collecting privacy sensitive information like information on personal habits. Also as data often encode belongings (IP) and other information on high value to organizations, attacks increasingly specialize in data theft and infiltration.

It also important to focus that as data are often used for critical decision making, data trustworthiness could be a crucial requirement. Data have to be protected against unauthorized modification. Comprehensive data trustworthiness solutions are difficult to attain as they have to mix different techniques, like semantic integrity, digital signatures, data quality techniques. To assuming data trustworthiness may require a wise control on data management process which successively has privacy implication.

In this article we first discuss the concept of big data then reviewed the data security and privacy for giant data we then outline relevant challenges and research directions in data security and privacy for giant data afterward we conclude

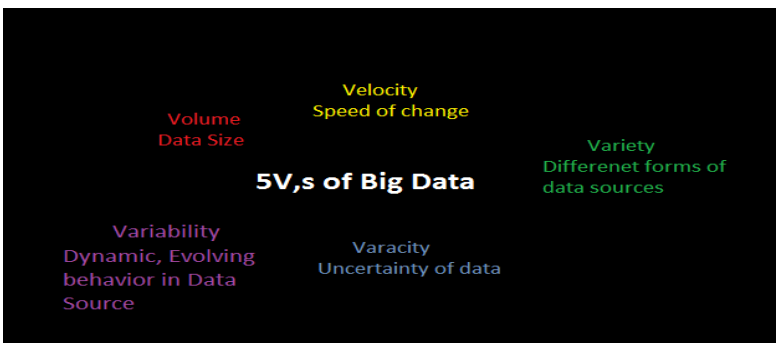


Figure 1 5V's Big Data

 ISBN: 978-81-946375-7-8

II. Literature Review

Ninny Bhogal and Shaveta Jain(2017)[1] presents a memory-and-time efficient probabilistic method for viably handling the huge facts to fulfil the worth of records development. Also, the massive data protection exam is classed against the BF and its variations. By directing a discover one among a form avenues concerning a large extent of records, the provided machine along a calculation for removing the information is tried. one among our destiny works is to actualize the Dynamic Bloom filter (DBF) and examine its execution towards the Counting Bloom filter out (CouBF) with reference to the substantial records ordering and recovery. During this paper they proposed a talented and top quality grained fact get to govern plot for giant facts where they get to method which won't launch any protection statistics. Not the identical because the contemporary techniques which simply mostly hide the characteristic values inside the get to processes, their method can shroud the complete property (as against simply its features) within the get to preparations.

ALGORITHM/TECHNIQUES

Application Software Security, Maintenance, Monitoring, and Analysis of Audit Logs, Secure Configurations for Hardware and Software, Account Monitoring and Control.

Prof. Amar Nath Singh, Er. Anurag Pattanayak, Er. Gyanachanda Samantaray(2016)[2] observed that, now each day we want a high speed processing environment. Hence we want a correct concrete solution for processing of such data. in order that they have opt for the analysis of huge data first before the processing. The technological advances in storage, processing, and analysis of huge Data include

(a) the flexibleness and cost-effectiveness of information centres and cloud computing for elastic computation and storage

(b) the event of recent frameworks like Hadoop, which permit users to require advantage of those, distributed computing systems storing large quantities of information through flexible multiprocessing.

Hence, by using this approach, the normal approach is now a day’s not used.

Big Data analytics used for security purposes: Network Security, Enterprise Events Analytics, Advanced Persistent Threats Detection.

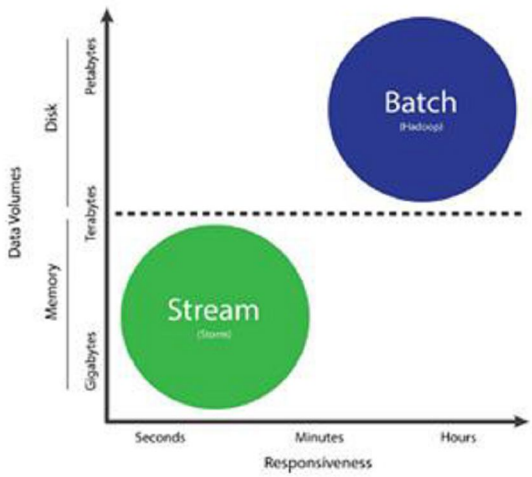


Figure 2 Batch and Stream Processing [2].

ALGORITHM/TECHNIQUES

Fraud detection is one all told the foremost visible uses for big data analytics mastercard companies have conducted fraud detection for several years.

Minit Arora and Dr Himanshu Bahuguna(2016)[3] presents a survey that organizations used various methods of deidentification to verify security and privacy. the foremost common solution to verify security and privacy is additionally oral and written pledges. Passwords, controlled access, and two factor authentication is low-level, but routinely used, technical solution to enforce security and privacy when sharing and aggregating data across dynamic, distributed data systems. Access permissions like these can potentially be broken by both the intentional sharing of permissions and also the continuation of permissions after they're not required or permitted. More advanced technological solution is cryptography. The famous encryption schemes have Advanced Encryption Standard (AES) and Rivest–Shamir–Adleman (RSA). Recent revelations show that the National Security Administration (NSA) may have already found ways to interrupt or circumvent existing internet encryption schemes. Virtual barriers like firewalls, secure sockets layer and transport layer security are designed to limit access to data. Each of these technologies are broken, however, and thus should be constantly monitored, with fixes applied pro re nata. Tracking, monitoring or auditing software is developed to produce a history of knowledge flow and network access by a private user so on confirm compliance with security related. The limitation of this technology is that it's difficult and expensive to implement on an outsized scale or with distributed data systems and users because it requires dedicated staff to read and interpret the findings, and also the software are exploited to watch individual behaviour rather than protecting data. Thus the conventional de-identification techniques don't seem to be applicable within the age of giant Data since the de-identification technique widespread uses. The tasks of ensuring Big Data security and privacy become harder as information is increased. Computer scientists have repeatedly shown that even anonymized data can often be reidentified and attributed to specific individuals.

Naveen Rishishwar, Vartika and Mr. Kapil Tomar(2017)[4] presents an outline about the large Data security issues and challenges. Big data handles a petabyte of knowledge or more. it's distributed redundant data storage. Can leverage parallel task processing, provide processing (MapReduce or equivalent) capabilities and has extremely fast data insertion. Has central management and orchestration. Is hardware agnostic. Is extensible where its basic capabilities may be augmented and altered. Nothing is ideal each and each thing have their own merit and demerit (pros and cons) so big data even have their own. a number of them are given below with their possible solution.

- Storage issues,
- Security,
- Processing issue in Big Data,
- Privacy in Big Data,
- Redundancy.

Bhavani Thuraisingham(2014)[5] presents an summary about the large data and privacy together with its security. Many privacy enhancing techniques are proposed over the last fifteen years, ranging from cryptographic techniques like oblivious data structures that hide data access patterns to data anonymization techniques thattransform the information to create tougher to link specific data records to specific individuals. The Privacy-Enhancing Symposium (PET) series, and journals, like Transactions on Data Privacy. However, many such techniques either

don't scale to very large data sets and/or don't specifically address the matter of reconciling security with privacy. At the identical time, there are some approaches that target efficiently reconciling security with privacy and that they discuss them in what follows.

1. Privacy-preserving data matching,
2. Privacy-preserving collaborative data processing,
3. Privacy-preserving identity verification

The computational, storage and communication costs of given protocols have to be considered. These costs can be especially significant for privacy-preserving protocols that involve cryptography. Given these three dimensions, one can imagine a multi-objective framework where different dimensions can be emphasized:

1. Maximize utility, given risk and costs constraints,
2. Minimize privacy risks, given the utility and cost constraints,
3. Minimize cost, given the utility and risk constraints.

Comprehensive solutions to the matter of security with privacy for giant data require addressing many research challenges and multidisciplinary approaches. They outline significant directions in what follows:

Both techniques are widely investigated. However, for access control systems for giant data we want approaches for:

Merging large numbers of access control policies,

Automatically administering authorizations for large data and particularly for granting permissions,

Enforcing access control policies on heterogeneous multi-media data,

Enforcing access control policies in big data stores,

Automatically designing, evolving, and managing access control policies

Trupti V. Pathrabe(2017)[6] presents a comprehensive survey on security problems with growing technology which is said to Big Data. The analysis of huge Data involves multiple distinct phases which include data acquisition and recording, information extraction and cleaning, data integration, aggregation and representation, query processing, data modelling and analysis and interpretation. Each of those phases introduces challenges. Heterogeneity, scale, timeliness, complexity and privacy are certain challenges of huge data.

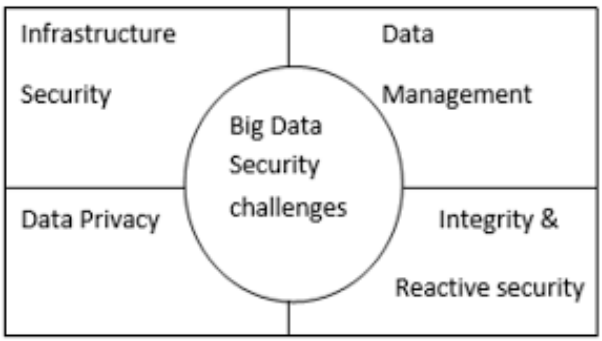


Figure 3: Challenges in Big Data Security [6].

There are four different aspects of huge Data security:

Infrastructure Security-Security for Hadoop, Availability, Architecture Security, Group Communication, Communication Security, Authentication.

Data Privacy- Cryptography, Access Control, Confidentiality, Privacy-Preserving Queries, Privacy in Social Networks, Anonymization, Differential Privacy.

Data Management- Security at Collection or Storage, Policies, Laws, or Government Sharing Algorithms,

Integrity and Reactive Security-Integrity, Attack Detection, Recovery.

Vinod B. Bharat, Pramod B. Deshmukh, Laxmikant S. Malphedwar, P. Malathi and Nilesh N. Wani(2017)[7] presents a thought about Big Data and Database Security. during this paper they targeting the large information security and protection challenges. They targeting survival security professional oriental exchange diaries to centre an underlying rundown of high- need security and protection issues and landed at the accompanying main ten difficulties.

1. Secure calculations in disseminated programming structures
2. Security best practices for non-social information store
3. Secure information stockpiling and exchanges logs
4. End-point info acceptance/sifting
5. Ongoing security observing
6. Adaptable and compostable security saving information mining and examination
7. Cryptographically upheld information driven security
8. Granular access control
9. Granular reviews
10. Information provenance

ALGORITHM/TECHNIQUES

This paper has uncovered the $64000 security issues that ought to be cared-for in Big Data handling and capacity. some specialists have realized the employment of encryption along with Kerberos convention keeping in mind the tip goal to form the knowledge safer. In any case, these security and protection issues are available in various structures such Kerberos won't not be sufficient to completely secure the knowledge. Amid Map-Reduce system in Hadoop, mapper hubs handle a given arrangement of knowledge and recovery the center person information inside their nearby documents. The reducer hubs will then duplicate this information from the mapper hubs and afterward total it to form the overall result. we would want to present an additional focal hub which interfaces with both the mapper and therefore the reducer hubs. The delegate information will then be put away during this hub instead of the mapper hubs' nearby record framework. a position safeguard component will then be utilized to screen all the movement going into and out of the hub to secure the knowledge.

Tilwani Mashook, Patel Malay and Pooja Mehta(2017)[8] presents a case study on Tracking and Monitoring System supported Security and Privacy in Big Data. the most goals of Employee Tracking Systems are to observe the workers or the sector labourers and help them analyse themselves also to let the organizations to analyse their performance. The information obtained

from the servers is processed online or offline for detailed analysis at the remote server per the appliance requirements.

Most distributed systems' computations have only one level of protection, which isn't recommended.

Non-relational databases (NoSQL) are actively evolving, making it difficult for security solutions to stay up with demand.

When a system receives an outsized amount of data, it should be validated to stay trustworthy and accurate; this practice doesn't always occur, however.

Access control encryption and connections security can become dated and inaccessible to the IT specialists who depend upon it.

Some organizations cannot – or don't – institute access controls to divide the amount of confidentiality within the corporate.

Recommended detailed audits don't seem to be routinely performed on Big Data because of the massive amount of data involved.

Because of the scale of huge Data, its origins don't seem to be consistently monitored and tracked.

Some of the foremost recent challenges observed within the massive Data by the Tracking and Monitoring Application's Organization: user data privacy, granular access, monitoring in real-time, granular audits, preserve the privacy in data processing and analytics, encrypted data-centric security, data provenance and verification, integrity and reactive security.

ALGORITHM/TECHNIQUES

During the transition phase, the HER vendor must work closely with the healthcare provider for a smooth and secure transition. the corporate should provide some sort of comprehensive user guide for the users within the provider's practice for implementing a number of the software's or applications within the device of the human who is meant to be tracked or monitored.

III. BIG DATA SECURITY: FUTURE DIRECTIONS

The following are a number of the longer term enhancements which i've got found while referring these papers. to bolster big data security- target software protection, in location of tool safety. Isolate gadgets and servers containing important facts. Introduce real-time security data and event control. Provide reactive and proactive protection[1]. Another major thing are privacy requirements in big data collection, storage and processing[3]. Major big data security challenges are: In Big Data most distributed systems computations have only one level of protection, which isn't recommended. Non-relational databases (NoSQL) are actively evolving, making it difficult for security solutions to remain up with demand. When a system receives an oversized amount of data, it should be validated to stay trustworthy and accurate[4].

CONCLUSION

The study of assorted methodologies by many researchers are making the info secured and provide privacy which made clear about the assorted methods, its merits and demerits and inabilities for providing security and privacy in Big Data. With this, we are able to come to conclude that we required some new technologies or the considerable modifications within the available technology.

REFERENCES

[1] Ninny Bhogal, Shaveta Jain," A Review on Big Data Security and Handling", International Research Based Journal, Vol(6)-Issue(1) , ISSN 2348-1943, March,11, 2017.

[2] Prof. Amar Nath Singh, Er. Anurag Pattanayak, Er. Gyanachanda Samantaray,"Data Analytics Application used in the field of Big Data for Security Intelligence" (IJETTCS), Volume 6, Issue 1, ISSN 2278-6856, January - February 2017.

[3] Minit Arora, Dr Himanshu Bahuguna, "Big Data Security – The Big Challenge", International Journal of Scientific & Engineering Research, Volume 7, Issue 12, ISSN 2229-5518, December-2016.

[4] Naveen Rishishwar, Vartika, Mr. Kapil Tomar," Big Data: Security Issues and Challenges", International Journal of Technical Research and Applications, e-ISSN: 2320-8163, Special Issue 42 (AMBALIKA), PP. 21-25, March 2017.

[5] Bhavani Thuraisingham, "Big Data Security with Privacy", NSF Workshop,September 16-17, 2014.

[6] Trupti V. Pathrabe, "Survey on Security Issues of Growing Technology: Big Data", IJIRST, National Conference on Latest Trends in Networking and Cyber Security, March 2017.

[7] Vinod B. Bharat, Pramod B. Deshmukh, Laxmikant S. Malphedwar, P. Malathi and Nilesh N. Wani, "Big Data and Database Security", IJCTA, 10(8), pp. 517-528 ISSN: 0974-5572, International Science Press, 2017.

[8] Tilwani Mashook, Patel Malay, Pooja Mehta," Security and Privacy – A Big Concern in Big Data a case study on Tracking and Monitoring System", IJIRST, National Conference on Latest Trends in Networking and Cyber Security, March2017.

[9] A Research on Big Data Analytics Security and Privacy in Cloud, Data Mining, Hadoop and Mapreduce

 ISBN: 978-81-946375-7-8

Fog Computing and Internet of Things

Vaijayanti Sambhaji Mondkar
MCA Student, YMT College of Management, Navi Mumbai

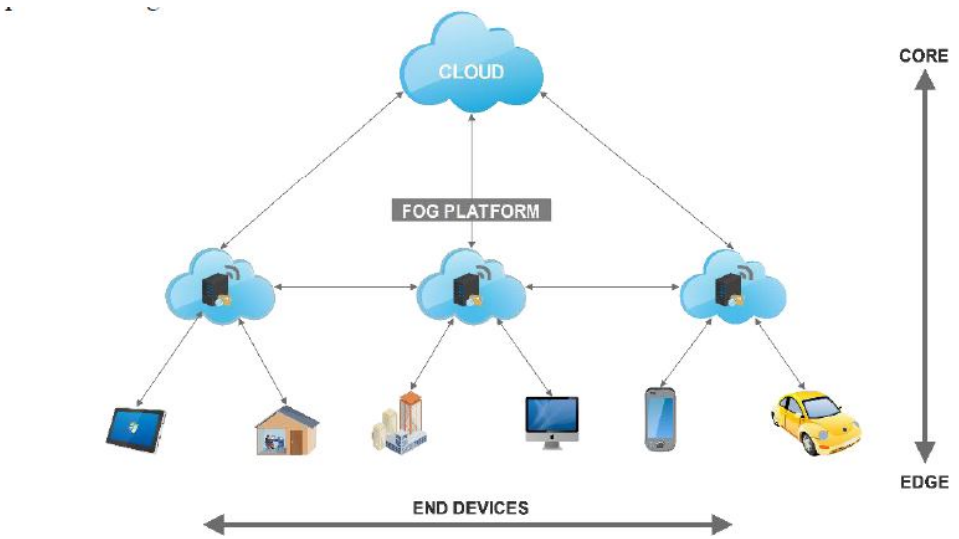
Abstract
Fog Computing spread out the Cloud Computing to the authority of the set-up, thus allowing a new applications and services. Fog computing make available data, stowage, and application services to users. fog computing is regarding for security incidents. fog computing handier to IoT devices. fog computing has appeal to ample attention due to the enormous potential. Fog cloud devices carry only specific information that is often used in it and usually the scope of the data in fog cloud device is very less in contrast with cloud server. The IoT development providing through cloud computing.

Keywords: Fog computing, Security Problems, Internet of Things, IoT with fog computing, MD5.

I. Introduction

Fog computing, is a brand novel and developing computing technology in the current computing world. Fog computing is established as a technology to link between remote data hubs and IoT devices. Fog computing surroundings, fog nodes make available service based on the information composed from the IoT node. Fog Computing phenemenon innovates CISCO. It as works on previously connected devices in IoT to scamper at fringe networks.

Figure.1[1] Basic structural design of fog computing



Fog computing is based on given that data processing abilities and storage nearby to fog devices as an alternative of transfer to the cloud. The resolution of fog computing in the IoT is to increase efficiency, performance and decrease the volume of data moved to the cloud for processing, analysis and storage. The grouping of fog computing with the IoT generates a new chance for services, which is called fog as a service (FaaS), where a service provider builds an array of fog nodes across its geographic footprint. Local estimate, networking and storage abilities presented by each fog node. FaaS will permit new business models to provide services to customers.

Table.1[1,2]Contrast of Fog and Cloud Computing Concepts

Service	Fog Computing	Cloud Computing
Model	Distributed	Centralized
Mobility	Supported	Limited
Security	Defined	Undefined
Connectivity	Wireless	Leased line
User	Mobile users	Internet users

Characteristics of Fog Computing Architecture

In order to take advantage of the elasticity of cloud computing architecture as well as to solve the above difficulties encountered in the IoT applications, Cisco Inc. proposed the concept of fog computing in 2012. Fog computing is theoretically the extension of cloud computing. Fog computing was named after cloud computing due to the idea that "cloud at ground level is known as fog. "Compared with cloud computing, the architecture used for fog calculation is more distributed and closer to network edges. Fog computing organizes data, data processing, and applications at network edges, which is unlike cloud computing that keeps them nearly entirely in the cloud. The storage and processing of data are more consistent on local facilities than servers in fog computing. Therefore, cloud computing is a new generation of centralized computing, while fog computing is a new generation of distributed computing, which is in line with the "decentralization" feature of Internet [3,4].

Fog computing is mainly based on small cloud such as personal, private, and enterprise cloud, while cloud computing is mainly based on IT services, public cloud. Fog calculation is powerful in large quantity and emphasizes the quantity, where the single computing node plays an important role, while cloud computing emphasizes the overall computing power, which is typically calculated by a bunch of concentrated high-performance computing facilities.

Fog calculation expands the network-computing model of cloud computing and extends the network computing from the center to the edges of the network, as shown in Figure1, which is therefore more widely used in a variety of applications [5–8]. The noticeable characteristics of fog calculation are as follows.

(1) Low latency and position sensing characteristics as fog computing locates at edges: it is of significance to the current informatics IoT in water conservancy engineering; for example, low latency is required in sudden floods, water pollution, and personal safety.

(2) Dense geographical distribution: this coincides with the wide distribution of water management of the SNWTP.

(3) Enormous amounts of nodes: a large-scale sensor network with numerous network nodes can be used to monitor the environment.

(4) Adaptability to access to mobile devices: mobile devices in fog computing can communicate directly with each other without transferring to the cloud or station, so that fog computing is highly adaptable to mobile devices, for example, mobile water quality monitoring and water level monitoring [9].

(5) High real-time feature: fog computing supports computation and processing of data at network edges with a low latency [10].

II. LITERATURE REVIEW
RELATED WORK
A. OVERVIEW OF FOG
Fog computing is definite in words to encounter the objectives of equal notions. Fog computing is a software platform reflected as non-trivial supplement of menace access where summons and interventions will relocate the fog having some illustrious possessions.

B. NEED FOR FOG
IoT distributes with big data investigative, Mobile and cloud computing where IoT is trending automation. Internet of things assurances to implement high-range functions related to homes, automation of vehicles as smart works. To increase the options and requirements evaluation, IoT needs a maturing creativity to make an application reliable at gateway functionalities.

C. FOG NODE
Fog nodes are the hardware appliances or physical elements as gateways, switches, cloud lets[3], routers etc.., to provide the essential resources for IoT applications through these apparatus. The fog nodes can be in between the cloud layer and IoT application layer in suburbanized manner to organize the mechanism at geo-locations dispense over the scalability to federate along to form cluster of nodes/servers.

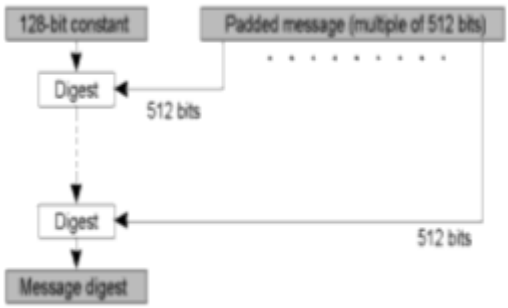
Fog computing used for the internet of things[4]: safety and privacy problems, author Arwa Alwaris explored and talk over safety and privacy tasks of presenting fog computing in Iot surroundings [4]. Fog computing will carry cloud networking, computing and storage abilities down to the upper hand of the network, which will address the actual problem of IoT devices and afford secure and ef cient IoT applications.The fog computing paradigm: situations and safety Problems, author sheng wen attending on fog computing benefits for services in a number of domains and make available the examination of the state-of-art and safety Problems in current paradigm[5].The Fog computing paradigm is well maintenance heavily distributed data collection points, and distributes benefits in performing, personal computing and other applications. Fog computing to defend actual and understanding data in cloud, author Ashwini and Mrs. Anuradha held fog computing is a model which aids the presentation of the user and providing safety to the user data[6].

Fog collectors at the edge ingest the data generated by grid sensors and devices. Some of this data relates to protection and control loops that require real-time processing (from milliseconds to sub seconds). This first tier of the Fog, designed for machine-to-machine (M2M) interaction, collects, process the data, and issues control commands to the actuators. It also filters the data to be consumed locally, and sends the rest to the higher tiers. The second and third tier deal with visualization and reporting (human-tomachine [HMI] interactions), as well as systems and processes (M2M). The time scales of these interactions, all part of the Fog, range from seconds to minutes (real-time analytics), and even days (transactional analytics). As a result of this the Fog must support several types of storage, from ephemeral at the lowest tier to semi-permanent at the highest tier.[6]

III. Research Methdology
Cryptographic hash function(CHF):
A cryptographic hash function(CHF),it is a hash function which takings input (or message) and earnings a fixed-size alphanumeric string[7]. The string is called message digest(MD). Hash functions are routinely used to check integrity or for error finding of transmitted messages. A cryptographic hash function must be one-way. The MD5 digests have been broadly cast-off in the software world to make accessible assurance about honesty of transferred file. Ron Rivest is developed MD5 algorithm.MD5 algorithm can be cast-off as a digital signature mechanism.

 ISBN: 978-81-946375-7-8

Figure.2[7] :MD5 algorithm structure



MD5 is quite fast, and Produces 128-bit message digests. It takes input as a "fingerprint" or "message". The input text is handled in 512-bit blocks. The output of the algorithm is a set of four 32-bit blocks, which create 128-bit MD.

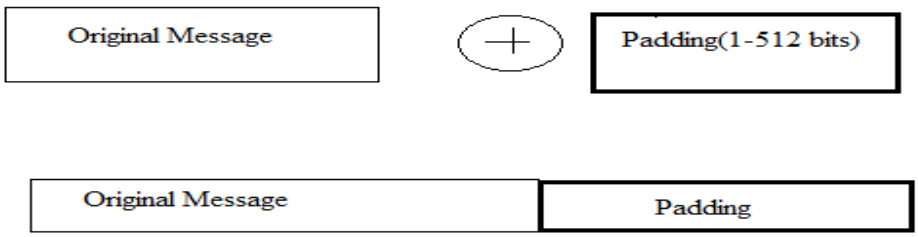
Steps in MD5 algorithm

STEP1: Append padded bits .

The input message is padded so that its length is consistent to 448, modulo 512[7].Means lengthy to just 64 bits shy of being of 512 bits long[7]. A single "1" bit is appended to the message, and then "0" bits are attached so that the length in bits equals 448 modulo 512[7].

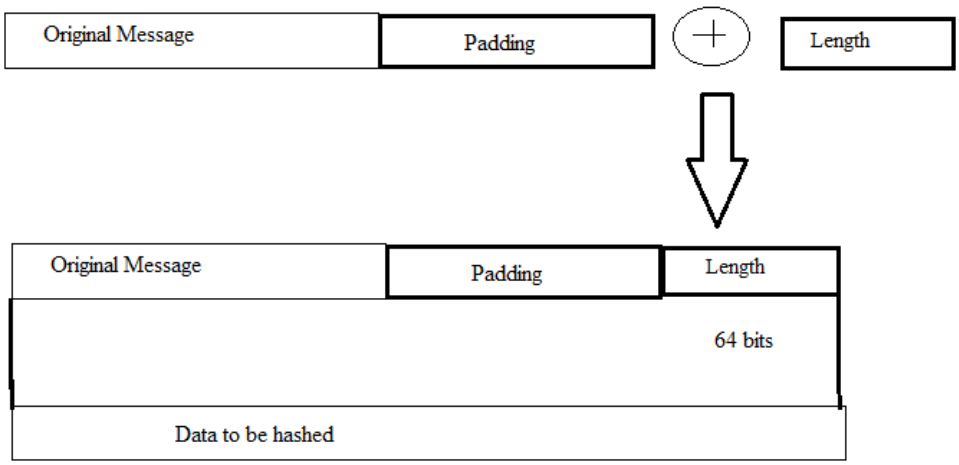
The padding process is show in fig.

Figure.3 Padding Process



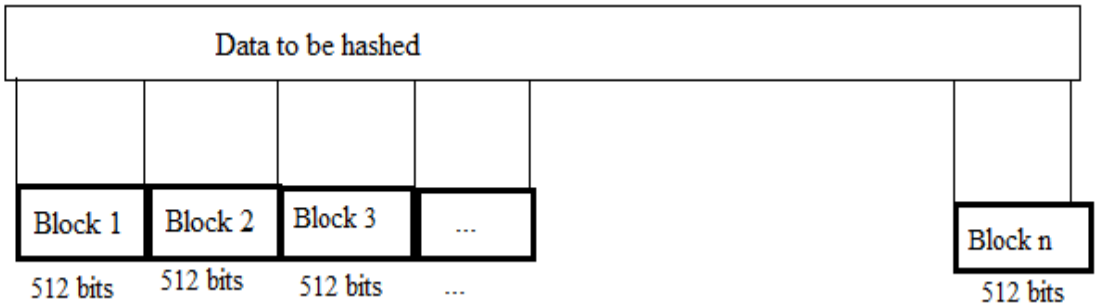
STEP2: appended length A 64 bit illustration of b is appended to the result of the preceding step[7]. The resulting message has a length that is an precise multiple of 512 bits[7].

Figure.4 Append length



STEP3: prepare MD buffer A four-word buffer (A, B , C, D) is cast-off to compute the message digest[7]. Here each of A, B, C, D is a 32 bit register[7].

Fig.5 Data divided into 512-bit blocks



These registers are prepared to the following standards in hexadecimal:

Fig.6 Chaining Variables

A	Hex	01	23	45	67
B	Hex	89	AB	CD	EF
C	Hex	FE	DC	BA	98
D	Hex	76	54	32	10

STEP 4:Process message in 16-word blocks Four supplementary functions that takes as input three 32-bit words and create as output one 32-bit word[7]. F(X,Y,Z)= XY v not(X)Z G(Y,Y, Z)=XZ v Y not(Z) H(X, Y, Z)=X xor Y xor Z I(X, Y, Z)=Y xor(X v not(Z) If the bits of X,Y,Z are free and impartial, the each bit of F(X,Y,Z), G(X,Y,Z), H(X,Y,Z) and I(X,Y,Z) will be free and impartial[7].

STEP 5: The message digest produced as output is A,B,C,D, i.e., output begins with the low-order bits of A and end with the high-order byte of D[7].

IV. Conclusion

Data theft attack come to be serious problem for cloud service suppliers. Fog computing is a model which helps in noticing the performance of the user and given that safety to operator data. In fog computing we extant a new style for solving the teething troubles of insider data theft occurrences in a cloud using enthusiastically produced decoy files .So by with decoy technology with user presentation profiling we can decrease insider occurrence in cloud. Cryptographic hash functions(CHF) are cooperative tool in safety of integrity.MD5 algorithm safeguards a huge file and is easy to contrivance and simple. The stimulating is when two messages are of similar message digest.

V. Feature Enhancement

Data encryption is necessary to need to be secured from the source to the destination during entire transmission in all layers of fog computing based IoT. Cache management is possible by using the appropriate cryptographic key used but should be confidential. Fog system makes use of innovative cache techniques which leads to the disclosure of confidential data. An authentic licensed hardware or software based tool should be used for network analyzing or management in order to monitor the full network with the support of appropriate ACL. But this list may be large cause of thousands of devices or objects. The required software and prototype which the developer has embedded in the electronics circuits should be accordingly in order to make sensors and controllers work accordingly with the system and periodically updated.

References

1. Abdullah Al-Noman Patwary1, Nishat Hossain2, Mohammad Aslam Sami3 “A Detection Approach for Finding Rogue Fog Node in Fog Computing Environments.” American Journal of Engineering Research (AJER) 2019
2. Sourav Kunal Arijit Saha Ruhul Amin “An overview of cloud-fog computing: Architectures, applications with security challenges” Published on: 16 May 2019
3. Wikipedia, Fog computing[EB/OL],2016,https://en.Wikipedia .org/wild/Fog computing.
4. W. Steiner and S. Poledna, “Fog computing as enabler for the Industrial Internet of Things,” Elektrotechnik und Informationstechnik,vol.133,no.7,pp.310–314,2016.
5. W.Fang, “A paradigm shift to fog computing from cloud computing and edge computing,” Journal of Nanjing University of Information Science and Technology: Natural Science Edition , vol.8,no.5,pp.404–414,2016.
6. F.Bonomi,R.Milito,J.Zhu,andS.Addepalli,“Fog computing and its role in the internet of things,” in Proceedings of the 1st edition of the ACM Mobile Cloud Computing Workshop on Mobile cloud Computing(MCC’12),pp.13–15,Helsinki,Finland, August2012.
7. S.Tomovic,K.Yoshigoe,I.Maljevic,andI.Radusinovic,“Software-defined fog network architecture for IoT,” Wireless PersonalCommunications,vol.92,no.1,pp.181–196,2017.
8. M.Chiang and T.Zhang, “Fog and IoT: an overview of research opportunities,”IEEEInternetofThingsJournal,vol.3,no.6,pp. 854–864,2016.
9. T. V. N. Rao, M. A. Khan, M. Maschendra et al., “A paradigm shift from cloud to fog computing,” IJCSET, vol. 5, no. 11, pp. 385–389,2015.
10. N.Peter, “Fog computing and its real time applications,” International Journal of Emerging Technology and Advanced Engineering,vol.5,no.6,pp.266–269,2015.
11. Lingjuan Lyu, Student Member, IEEE, Karthik Nandakumar, Senior Member, IEEE, Ben Rubinstein, Jiong Jin, Member, IEEE, Justin Be do and Marimuthu Palaniswami, Fellow, IEEE, “PPFA: Privacy Preserving Fog-enabled Aggregation in Smart Grid”, DOI 10.1109/TII.2018.2803782, IEEE Transactions on Industrial Informatics.
12. Alrawais, Arwa, et al. "Fog Computing for the Internet of Things: Security and Privacy Problems." IEEE Internet Computing 21.2 (2017): 34-42.
13. Stojmenovic, Ivan, and Sheng Wen. "The fog computing paradigm: Scenarios and security Problems." Computer Science and Information Systems (FedCSIS), 2014 Federated Conference on. IEEE, 2014.
14. Ashwini, Thogaricheti, and Mrs. Anuradha SG. "Fog Computing to protect real and sensitivity information in Cloud."
15. Sai Keerthi.Tulluru, Dr.Anuradha S.G Dept of Computer Science, RYMEC, Ballari, India Associate Professor, Dept of Computer Science RYMEC Ballari, India “Fog Computing a Paradigm: Scenarios And Security Problems” International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 5, Issue 4, April 2018

 ISBN: 978-81-946375-7-8

Performance of Machine Learning Recommendation System Algorithm

Miss. Vaishnavi Lonkar
Student, Navinchandra Mehta Institute of Technology and Development, Dadar (W)

Dr. Rasika Mallya
Associate Professor, MCA Department, DES's NMITD, University of Mumbai, Mumbai

Abstract
This paper shows the performance of machine learning recommendation system algorithms. Recommendation systems are software equipment and procedures that provides recommendations for items to the user. It gives description about machine learning techniques recommendation system also types of recommendation system. Also it will give description about Collaborative filtering and how it works, their two types Item-based collaborative filtering and User based collaborative filtering. So as result, one will get to know performance of machine learnings recommendation system.

Keywords: Recommender System, Collaborative Filtering, Item-Based Collaborative, User-Based Collaborative Filtering, Bayesian network, Horting.

I. Introduction

Machine learning is an application of artificial intelligence that gives systems the power to consequently take in and improve as a matter of fact without being explicitly customized[8]. Recommendation system in machine learning uses Big data to recommend additional products to customers based on a variety of reasons. Recommendation is one of the techniques of machine learning that one can used to solve some common business problems. Recommendation uses your information along with community information to determine the likelihood of your preferring a product or not. Netflix uses recommendations to suggest movies, and uses it to build radio stations any harmony uses it to match people. Machine learning programs the system in such a way that they automatically recognize and understand the input data and make wise decision based on provided data. There are three important steps in recommendation system, the first one is collection of information of items and representation then finding the similarity in decisions and then recommendation computations.

II. Literature Review

2.1 Recommendation system

Throughout the years, different methodologies for building recommender frameworks have been made[11]; collaborative filtering has been an effective methodology in both research furthermore, practice, and in data sifting and web based business applications[12]. Recommendation system utilize efficient prediction algorithm which is used to provide the accurate recommendation to users. The system finds and recommend items that the users are interested in. When one visits the e-commerce site such as Apple, Netflix, Amazon etc., the host of that site recommend some products and it is derived from historical data. The system saves the user's time by suggesting them the best of their choice and increase the potential of the sale. Recommendation system made up of some prominent algorithms which will suit to their needs.

a. **Random Prediction algorithm:** This algorithm picks an item randomly from large set of items to recommend it to the user. This algorithm is not reliable because it completely depends on the luck, so it considered as a failure one.

b. **Frequent Sequence algorithm:** This algorithm is based upon how frequent a user rates an item, on this basic host recommend an item to the users. It uses the frequent pattern for suggesting items to the users. This algorithm is useful when user makes minimum purchases.

c. **Collaborative filtering:** In collaborative filtering user rate the items. These ratings able to be seen as an almost picture of the user's interest. The system matches one user rating against the another users and discovers the people with similar tastes[2]. In the event that an individual A has a similar point of view as an individual B on an issue, A is more prone to have B's conclusion on an unexpected issue in comparison to that of a randomly picked individual[7].

d. **Content based:** Content based recommender utilize the data that user provides, either by rating or by clicking on a link. User profile is generated based on the provided data, which is then used to make suggestions to the user. The aim of this function is to recommend items that are similar to items the user wanted once. Items that are picked up for recommendation are items whose content match up the most with the user's preference.

e. **Hybrid Recommender System:** This system is the combination of both collaborative filtering and content based filtering.

f. **Demographic Recommender system:** On the basis of analytical or measurable information such as age, gender and income this recommendation system categorizes the users or items.

Thus, collaborative filtering can filter the products based on difficult and hard to illustrate hypothesis such as taste and nature of a product. However, collaborative filtering is the best approach. While there is a ton of work in the field of collaborative filtering calculations, as of late a few papers have focused on developing shilling attack models[13,15] and on benchmarking the heartiness of recommender frameworks against shilling trigger attacks [16, 14].

2.2 Types of recommendation System:

The Collaborative filtering was first coined by Goldberg for email filtering system called Tapestry. Tapestry used to be an electronic text messaging system that let users to rate messages. Tapestry used to provide good suggestion, but it has drawback: the user was needed to write complex queries.

The GroupLens generated automated recommendation system. The GroupLens application provided users with recommendation on Usenet6 postings[3]. It used to recommend articles to the users.

There is Ringo recommender system which was developed by Sharadanand and Maes and it is used as recommendations for music albums and artists. In Ringo recommendations are done using emails. Also Video recommender system are there.

Bayesian network based recommendation system is represented by decision tree. Details of uses is represented by nodes and edges. It is fast to deploy because the capacity of the trained model is very small. It does not give the exact prediction for the constant changing situation.

Horting is a graph based technique. In this the node represent user and the similarity measure between two users. In this, the recommendation is formed by finding for the nearest neighbor nodes and then integrating the scores of the neighbors.

2.3 Collaborative Filtering

Collaborative filtering is the most common recommendation technique used for recommending products or services to the user. It learns the past user-item relationships from a community of user who share the same preferences and tastes [4]. Using the opinions of the other people it filters and evaluate the items. Collaborative filtering has taken its roots from something humans have been doing for centuries-sharing opinions with others. It finds the nearest neighbor based on similarity measures. This similarity calculation depends on the rating that user gave to the different items.

Collaborative filtering works in three steps:

1. In the first step user expresses preferences by rating items or services. These ratings can be viewed as an approximate representation of the clients enthusiasm for the relating area.
2. In the second step system identifies the people with most similar tastes by matching the user ratings against the other user.
3. With similar user, the system suggest items that the similar users have rated exceptionally however not yet being evaluated by this client.

Collaborative filtering has two methods User-based collaborative filtering and the Item-based collaborative filtering.

2.3.1 User-Based Collaborative Filtering:

User-based collaborative filtering recommend items by finding similar users. This is regularly harder to scale in view of the dynamic idea. This is also called as nearest-neighbor based collaborative filtering. To generate predictions it makes use of the entire user-item database. To find consumers nearest-neighbors such kind of systems use statistical techniques. After finding nearest neighbor of users, these systems perform different algorithms to combine the likings of neighbors to generate predictions for the target user. Similarities are compared using row-wise. This user-based algorithms does not scale well and are not appropriate for larger databases.

2.3.2 Item-Based Collaborative Filtering:

To overcome the problem of user-based Item-Based collaborative system were developed, Item-Based collaborative filtering calculate similarity between items and make recommendations. Items usually don't change much, so this often can be computed off line. The similitudes are evaluated using column wise. This algorithm makes use of item-user rating matrix. User-item matrix is described as an m, n ratings matrix $R_{m.n}$ where row shows m users and column represents n items[3]. The element of matrix $r_{i.j,}$ means score rated to the user I on the item j, which commonly is acquired with rate of user interest.

This algorithm has two steps

1. The algorithms scan past information of the users, the ratings they gave to items are gathered during this step. From these ratings, similarities between items are built and inserted into an item-to- item matrix[1]. The elements of the matrix represents the similarity between the items in row i and the item in column j.

2. This algorithm select items that are most similar to the particular item a user is rating. Values are calculated using different measures. Then the next step is to find the target item neighbors, this is calculated using the threshold-based selection and top-n technique. The final step is the prediction from the top-n results.

III. Implementation:

Performed implementation on Amazon instant 5 core file. The 5 core implies that each video has at-least 5 ratings and each user has rated at-least 5 videos.

User collaborative filtering is meant for speed and not for caching and memory usage. The basic idea behind user-user collaborative filtering is to predict a customer, users rating of a product, by finding customers that are similar to user and use their rating of a product to estimate how customer would rate it. For finding similar customers Cosine similarity function is used[5]. The cosine similarity between user *u* and user *v* is the normalized dot product of their row vectors (r_u and r_v respectively) in the rating matrix R.

$$s(u, v) = \frac{\mathbf{r}_u \cdot \mathbf{r}_v}{\| \mathbf{r}_u \| \ \| \mathbf{r}_v \|}$$

Unknown ratings are assumed to be 0, this will effect in drop out of numerator. Cosine similarity function generate predictions for users on product then used a weighted average of set of similar users(*N*) and customer ratings of product.

$$p_{u,i} = \bar{r}_u + \frac{\sum_{v \in N} s(u, v)(R_{v,i} - \bar{r}_v)}{\sum_{v \in N} |s(u, v)|}$$

When pre-computing of all the cosine similarities between users finishes then similarity matrix is formed. In user-user collaborative filtering algorithm, I compared each users similarities to all the other users and made suggestion. In this it loops over all the users and all the products and treat each product for each user as missing value and then predict its value. Then will get the error between the predicted rating and the actual rating.

If one reduce the number of customers that one have used to generate the recommendation it will help to improve the runtime of algorithm. It doesn't include those customers in computation whose choice is dissimilar to other customer of interest and this will not effect the accuracy of prediction.

Following are the examples of sites that achieved growth in their business using recommendation system.

- According to McKinsey, 35% of the purchases on **Amazon** are the outcome of their recommender system.
- According to Alizila, during the Chinese global shopping festival of November 11, 2016, **Alibaba** achieved gain of up to 20% of their conversion rate using personalized landing pages**[10]**..
- Recommendations are responsible for 70% of the time people spend watching videos on YouTube**[10]**.
- According to McKinsey, 75% of what people are watching on **Netflix** comes from recommendations**[10]**..

Following are some examples of best recommender system that come from Netflix, Best Buy, LinkedIn, Amazon and Google.

a. **LinkedIn:** LinkedIn is the world's largest professional network. In the past year and half 50% of the total job applications and job views by members are direct result of recommendation which was increased from 6% to 50%.

b. **Netflix:** Netflix built very accurate recommender system. Netflix has a secret system which collects information about how users interact with the resource like what user watch, when user watch, how user rate movies. Netflix recommendation system combines collaborative

filtering with content-based filtering and it makes use of machine learning algorithms to generate recommendations so it is an hybrid recommendation system.

c. **Amazon:** Amazon has an integrated suggestion into various phase. They recommend the popular and desired product on the main page. For returning customers amazon offer the personalized recommendations. On a product's page there's a number of places where they recommend related items. When user add new item to cart, new suggestion appear. When amazon sends an email to its customers the, it also include recommendations in it. Amazon has been reported that they saw 29% of increase in sales after implementing a recommendation system.

d. **Google**: Google influences users searches with autocomplete. It predicts what users are going to type in the search box. For prediction google uses machine learning, examination, and normal language preparing to recommend pertinent phrases. Google has their adds.

e. **Best Buy:** From 2015 this company started using recommender system. This company decided to focus on their online sales and in 2016 they reported 23.7% gain in their business.

IV. Conclusion

Recommendation systems are natural fit for analytics platforms. They involve large amount of customer data for processing which they collected online. In today's world, recommendation systems are everywhere and for many online platforms their recommendation engines are the actual business. Building a recommendation system results in saving customers time and money and also help the companies like YouTube, Netflix, Spotify or other Shopping sites in increasing their business worldwide. So basically, it proves beneficial for companies.

References

1. Almazro, Dhoha & Shahatah, Ghadeer & Albdulkarim, Lamia & Kherees, Mona & Martinez, Romy & Nzoukou, William. (2010). A Survey Paper on Recommender Systems.

2. Dilek Tapucu, Seda Kasap and Fatih Tekbacak, "Performance Comparison of Combined Collaborative Filtering Algorithms for Recommender Systems", Proc of IEEE 36th International Conference on Computer Software and Applications Workshops, 2012.

3. N. Susan Thampi, Dr. Senthil Kumar T., and Johnpaul, C. I., "Performance Analysis of Various Recommendation Algorithms Using Apache Hadoop and Mahout", International Journal of Scientific & Engineering Research, vol. 4, pp. 279-287, 2013.

4. Tran, Thi Ngoc Trang & LE, Viet Man & Nguyen, Minh Duc. (2015). REACTIVE MULTI-AGENT MODEL FOR COLLABORATIVE FILTERING - BASED RECOMMENDER SYSTEMS. THE UNIVERSITY OF DANANG, JOURNAL OF SCIENCE AND TECHNOLOGY. 1. 80-84.

5. Xiwei Wang, Erik von der Osten, Xuzi Zhou, Hui Lin and Jinze Liu, "Case Study of Recommendation Algorithms", Proc. of International Conference on Computational and Information Sciences, 2011

6. Deshpande, M., and Karypis, G. Item-based top-n recommendation algorithms. ACM Trans. Inf. Syst. 22, 1 (2004), 143–177

7. van Meteren, R., and van Someren, M. Using content-based filtering for recommendation.

8. M. P. O'Mahony, N. Hurley, and G. C. M. Silvestre.Promoting recommendations: An attack on collaborative filtering. In DEXA '02: Proceedings of the 13th International Conference on Database and Expert Systems Applications, pages 494–503, London,UK

9. M. O'Mahony, N. Hurley, N. Kushmerick, and G. Silvestre. Collaborative recommendation: A robustness analysis. ACM Trans. Inter. Tech.,4(4):344–377
10. R. Z. Robin Burke, Bamshad Mobasher and R. Bhaumik. Identifying attack models for secure recommendation. In Beyond Personalization.
11. S. K. Lam and J. Riedl. Shilling recommender systems for fun and profit. In WWW '04: Proceedings of the 13th international conference on World Wide Web,pages 393–402, New York, NY, USA,
12. https://github.com/mdh266/RecommenderSystems/blob/master/Recommender.ipynb
13. https://www.quora.com/Which-companies-use-recommender-recommendation-systems
14. https://expertsystem.com/machine-learning-definition/
15. https://sigmoidal.io/recommender-systems-recommendation-engine/
16. https://thenextweb.com/contributors/2018/11/24/googles-underrated-discover-feature-can-change-how-we-surf-the-web

 ISBN: 978-81-946375-7-8

Review on Women Health Medical Devices

Mr. Janmejay Patil and Ms. Tejal Wasankar
MCA, Deccan Education Society's Navinchandra Mehta Institute of Technology and Development, University of Mumbai

Abstract
This is a review paper representing an idea about the devices which can be used for monitoring the critical health issues in women. These health issues can be like ovulation, cervical cancer, breast cancer. The attempt has to be made to make a public awareness about these devices so that the rural parts of many countries and also the urban areas can make use of it to anticipate their health and take actions if any critical condition is perceived. These devices contain various components like sensors, auto strainers, and test channels. With the help of these devices women can monitor their health and take precautionary measures and consult the physician if arrives the necessity.These devices aids in determining ovulation and detecting cervical cancer and breast cancer at its earliest, and makes in convenient for women's to initiate their cure if found malignant.

I. Introduction

Oblivious health care systems are available but there arrives a need to augment a bridge between making it aware to every part of the world to make the most of it in controlling the death ratios due to cancer, ovulation, etc. Machine learning is a branch of artificial intelligence science i.e. the systems that can learn data[4]. Algorithms are casted-off using an Artificial Intelligence based computational pathology systems which are totally driven by a set of algorithms. These algorithms are rendered using sample data sets so that the machine learns with time being. Machine learning aids in forming accurate patterns as the input evolves. Accuracy and fidelity are major aspects in every device in order to scale the market for better productivity. Image processing aids in instigating the patterns for recognizing the health and making comparisons with the previously recognized patterns which are captured and processed using the algorithms. Deep Learning is used to build the core technology of the system and for clinical reasoning from images. Deep learning is a branch of machine learning based on a set of algorithms that attempt to model high level abstractions in data[5].

II. Literature Review

Youti Kuo presented research paper based on "Image processing system for predicting ovulation". In this paper an idea is presented regarding the device called "Electrical toothbrush" which is developed for predicting the ovulation based on sensors and camera and an important factor here is that the prediction is done on the basis of saliva sample. With the help of image processing body fluid like saliva is used for detection.

Dr Geetha Manjunath along with Nidhi Mathur discovered that Artificial Intelligence accompanying Machine Learning can be aided to detect lumps in women's breast, this device is named as "Niramai". In this research the team started with examining thermal images of the cancer patients and experimented on women with zero symptoms. This paper is solely dependent on the device which uses Artificial Intelligence algorithms for early detection of cancer.

Adarsh Nataranjan CEO and Founder of AIndra Systems developed the cervical cancer detection system which is constructed on the concepts in the likes of Artificial Intelligence along with Deep Learning called "CervAstra". It was this vision that drove us to use deep

technology as an enabler to create products that can create a huge impact on the society[8]. CervAstra provides the reports at the point of care, which involves faster screening process and sample collection along with the economical aspect of affordability.

III. Study Analysis Device 1: Electrical Toothbrush

As we know, that across the world there is a need of home devices which can predict the ovulation in women. There are many ways for detecting the pregnancy in woman's body. This device named Electrical Toothbrush was invented to detect the same but with the help of saliva. Saliva is the compound body liquid which contains various electrolytes.

This device makes use of concepts like image processing to detect the ovulation days. The test channel is used for sample collection of saliva which is the main part of this device. It has very small camera so that it can capture the image of the saliva. Drive-head is attached to the test channel which has one element named bristle and one notch which traverse around drive-head. In order to move back and forth of the bristle element the switch is turned on. After turning on the drive-head vibrates and it leads to the opening of the test channel. When the device comes in contact with the tongue these vibrating channel starts to accumulate the saliva from the tongue or from anywhere inside the mouth. A conductivity sensor is placed inside the internal location of the open channel. The test channel should be fully filled and it is sensed by the instant drop of the saliva which generates the electrical resistance which is measured by the sensor. Also this sensor detects the dryness of the saliva because the operation we are going to perform is on dried saliva only. The small camera makes use of an algorithm and captures the image of the dried saliva. A fern like structure is computed on the basis of the connection of the dark like pixels of the saliva dots which leads to the analyzing of the image pattern based on the line segments. The algorithm determines the length of the line of the saliva's crystalline pattern. The line length and the ferning index are plotted as drift curves and are displayed in the display unit which is attached to the handle. And these drift curves provides the prediction of ovulation.

This device does not contain any harmful rays which can penetrate in woman's body.

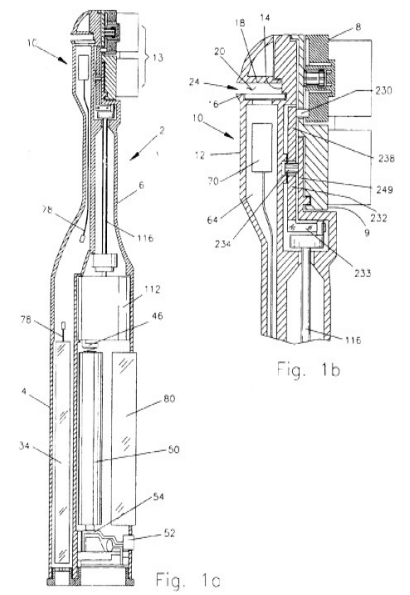


Figure: Electrical Toothbrush with enlarged view [1]

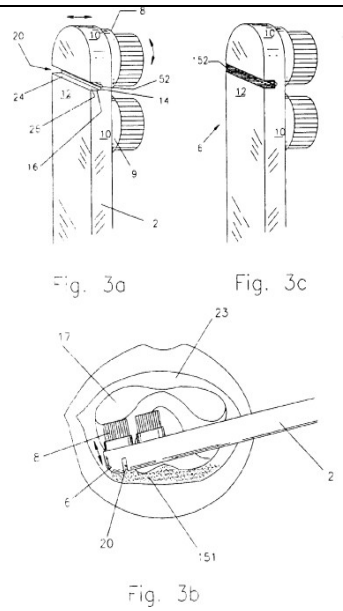


Figure: Toothbrush inside the tongue for collecting saliva [1]

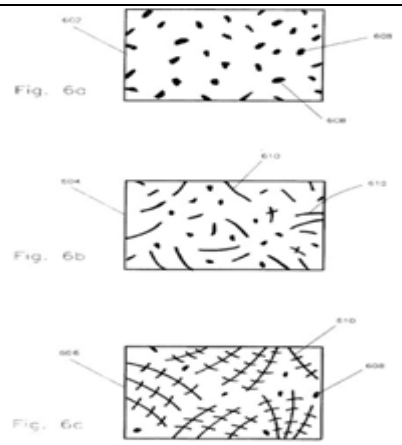


Figure: Saliva dots connection [1]

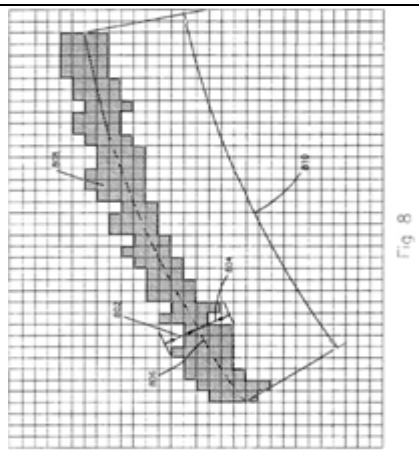


Figure: Finding the length and width for detecting [1]

Device 2: CervAstra

Cancer, most prominently ranks higher among the most chronic and deadly disease ever, which is the leading cause of deaths worldwide, which ranks 2nd in the number of deaths caused in a year. Innovative and modern technology is being implemented in the goal of providing proper diagnosis, prediction and in some cases treatments for cancer[3]. Among the numerous number of Cancers, Cervical cancer is one which is developed in Women and is the 4th most prevalent cancer in Women. Cervical cancer starts when the cells that line cervix begins to grow abnormal changes, these cells either turns back to being normal or they may develop into cancer causing cells. Cervical carcinoma is one of the most common and dreaded diseases of women[2]. Cervical cancer is basically divided into two types Carcinoma and adenocarcinoma, among which the former one (Carcinoma) which is the cause of more deaths as compared to the latter one (adenocarcinoma) i.e., it accounts for nearly about 80% of the cervical cancer as compared to the 20% of cervical cancer which is caused by adenocarcinoma.

This cervical cancer detection system which is powered by AI has proven itself as an influential device which reduces the cancer screening time considerably very less than usual. It has been seen that the women residing in rural areas will take anywhere between 5-6 weeks to receive the report after the collection of her screening samples. This process of collecting samples until the final step of arriving at a conclusion is so prolonging that it takes more processing time and assistance from various pathologist who had examined the samples than usual. Also the cost involved in such screening is pricey than expected.

A pap test is carried out, within which a pap smear (pap sample) is collected from the surface of the cervix i.e., the abnormally grown cells are collected and which are in turn processed through the first initial component of CervAstra, which is nothing but an autostainer which is called as Intellistain. The sample collected is stained in return to induce better results. This stained sample is then scanned using the very next i.e., the second component called as VisionX. VisionX is a whole slide imager which is proficient in generating a digital image out of the stained sample. The third and the final component is the Astra, the digitized image is either passed to this component or is used for Telepathology, telepathology employ pathology at a distance. Astra is mechanized with powerful artificial intelligence and machine learning algorithms which classifies whether the sample is “Normal” or “Cancerous”. A report of the same is generated and then shared with the sufferer at the point of care.

This AI based computational pathology system platform predominantly comprise of three components as follows:

- The first being the Intellistain, the autostainer that stains the pap smear collected from the diseased person.
- The second being the VisionX, which is chiefly a Whole Slide Imager which converts the stained slide of the pap smear into a digitized image.
- The third and the last one being the Astra, which is highly constructed on the concepts of Deep Learning and Deep Neural Networks. The algorithms of Deep Learning are being deployed in this component which is trained by the use of sample data sets which in successions gets more reliable due to the concepts of Deep Learning. In an analogous way the algorithms are trained and set up to descry cancerous cells from normal ones.

Figure: Device for testing the cervical cancer[7]

Device 3: Niramai

Breast cancer is a prominent form of cancer in women. Diagnosis of breast cancer is an important issue in Medical Science. This diagnosis can be related to the cause-effect graph or cause-effect ratio of software terminologies i.e., the more time it takes to find the root of the disease or criticality the more it costs to cure it and on the contrary the more effects it can cause to the human health. A radical method or test for detecting breast cancer early is Image Processing. Image processing comes with a quandary which is generated by different reasons namely Sampling, digitalization, intensity, etc.

Thermalytix a technique developed by Niramai is an innovative computer-aided diagnosis solution which helps in identifying cancerous and non cancerous patients. This technique comes with the primacy of being affordable for the middle class people, accessible and most importantly it is effective. The solution is contactless, is non intercoursable, involves low cost and also is portable which is used for detecting breast cancer. Prior to this Thermalytix technique is Mammography which has certain limitations in it and that is, it has low sensitivity for women under the age of 45 years while Thermalytix proves to be effective for all the age levels. The core system software is built on the novel computer algorithms based on the concepts in the likes of artificial intelligence as well as machine learning.

The whole process is very simple and it takes just 5-10 minutes to get it done at a normal room temperature. Niramai solution has its roots in the core concept of Thermography. Thermography is an approach wherein the heat radiated from the human body is captured. As the cancer cells tend to have high metabolic activity due to which heat is emitted from that particular region, using thermography these thermo graphic images can be obtained to visualize the abnormality. The Niramai thermal device machine is placed 3 feet away from the patient. This machine captures 5 images in all, each from various angles and of various regions. This innovation uses Artificial Intelligence over thermal images. These thermal images not only checks the rise in temperature but also checks the region in a way that the doctor can figure out whether the rise in

 ISBN: 978-81-946375-7-8

temperature is because of cancer being present or is it because of some other benign reason. The core innovation in embedded in the machine software which makes it easy to analyze the thermal images and which in turn returns the reports with the aid of Artificial Intelligence and Machine Learning which determines whether the lady is malignant or not.

What Niramai looks for is the abnormal tissue activity which is about five to eight years even before the lump is formed, as it detects these abnormal activities at very early stage, it has immense potential to save many lives. As the machine is kept away from the body and no human activities involved in any screening activities, this process of screening is contactless and also privacy aware, hence many women can undergo this screening activity. Another substitute technique to Niramai is Mammography which uses radiation, hence many women hesitate to undergo mammography screening primarily, on the contrary Niramai uses absolute no radiation. Another key benefit of the latter (Niramai) is that it works seamlessly for all age groups while on the contrary Mammography is more suited for women above 40-45 years.

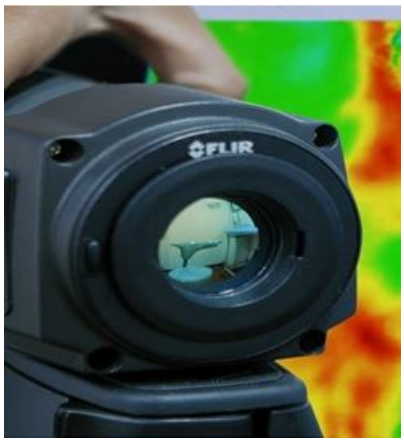


Figure: Device (Camera capturing the photos)[6]

Figure: Analyzing captured photo[6]

IV. Conclusion

Testing for ovulation requires women to get their urine samples and at a particular time of the day, so Electric Toothbrush proves to be a commending device as it comes over the shortcomings of normal devices used for detecting ovulation. CervAstra has it's advantage of providing results in a short span of time which provides leverage in starting the cure as soon as possible. Niramai anticipates in providing contactless treatment and laudatory results as it assesses the lump forming tissues very early than usual.

References

[1] Youti Kuo; “Image processing system for predicting ovulation”;Intellectual Property Consultant at Pharmanova Inc, Rochester, New York

[2] Anand Srivastava, Jata Shankar Mishra, Shruti Srivastava, Bhudav C Das, Shilpi Gupta; “Cervical cancer screening in rural India”; Lucknow Volume: 148 Issue: 6

[3] Shyma M, Al-Ani, Maysam Abbod; “Prediction of Cancer Behavior Based on Artifical Intelligence”

[4] Ali Heydarzadegan, Yaser Nemati, Mohsen Moradi; “Evaluation of Machine Learning Algorithms in Artificial Intelligence”; Volume: 4 Issue: 5

[5] Usha Devi N, Dr R. Mohan, Dr P.Kiran Sree; “An Extensive Survey on Deep Learning Applications”; International Journal of Innovations & Advancement in Computer Science IJIACS; Volume: 6 Issue: 2 ISSN 2347 - 8616

[6]. https://inc42.com/startups/how-niramai-is-helping-india-women-battle-deadly-breast-cancer-with-ai- driven-early-detection/

[7]. https://yourstory.com/socialstory/2019/07/aindra-device-cervical-cancer-artificial-intelligence

[8]. https://analyticsindiamag.com/meet-ai-powered-medtech-startup-making-cervical-cancer-screening- affordable-reliable-accessible/

ABOUT EDITORS

Dr. Samadhan K. Khamkar is an alumni of University of Mumbai belonging to Master of Management Studies (MMS). He has completed his Doctorate from University of Mumbai in "Business Policy and Administration" and also hold a Post-Graduation degree in Pharmaceutical Management. He has corporate experience of 6 years in pharmaceutical marketing and marketing of dairy products.

He has teaching experience of 9 years. He has been a resource person at various institutes, which includes SIES, N.L. College and SIBM. He has chaired various National and International Conferences. For over a decade, he has organized various Seminars, Conferences, and Mega cultural events. He has attended various FDPs and Workshop organized by IIT, ASCs, AICTE and Sydenham Institute for continuous improvement of quality of teaching and mentoring students.

He has also been a part of various NGOs who work for agri-business development programmes in rural area. He was also a channel partner for AMUL. His student friendly approach makes him a favorite faculty. His area of interest includes Marketing Application and Sales Management. He has published 22 research papers in National and International Journals of repute.

Dr. (Mrs.) Rasika Shailesh Mallya, HOD, MCA, DES NMITD, She has completed her Ph.D. in Requirements engineering for Wearable technologies. She has completed MCA and M.Sc. CS. She has more than 15 years of teaching experience. She is a passionate teacher who likes to teach technical subjects like Artificial Intelligence, IoT, Soft Computing, Multiagent Systems, Object Oriented Design using UML, Software Engineering for IoT, Data

Structures, Project Management, Software Engineering, C++. She has her excellence in guiding PG students for projects as well as writing research papers. Research Papers Published:

She has authored more than 20 research papers out of which 12 research papers are published in IEEE journals, 4 in Scopus indexed journal and some of them are published and presented in UGC-CARE and UGC recognized journals. She has also received Award as "Distinguished Researcher" awarded by, "International Journal for Research Under Literal

Access" Accredited by, "IJRULA Learning Center, Malaysia", for research article " System Requirements Model for Multi-Agent Autonomic Fetus Monitoring System ".

Empyreal Publishing House
India | UAE | Nigeria | Uzbekistan | Montenegro
ISBN 978-81-946375-7-8
9 788194 637578