Volume 7, Issue 1 (XIV)
January – March 2020

ISSN 2394 - 7780

International Journal of Advance and Innovative Research

Indian Academicians and Researchers Association
www.iaraedu.com

International Journal of Advance and Innovative Research

Volume 7, Issue 1 (XIV): January - March 2020

The views expressed in the articles are those of the contributors and not necessarily of the Editorial Board or the IARA. Although every care has been taken to avoid errors or omissions, this publication is being published on the condition and understanding that information given in this journal is merely for reference and must not be taken as having authority of or binding in any way on the authors, editors and publishers, who do not owe any responsibility for any damage or loss to any person, for the result of any action taken on the basis of this work. All disputes are subject to Guwahati jurisdiction only.

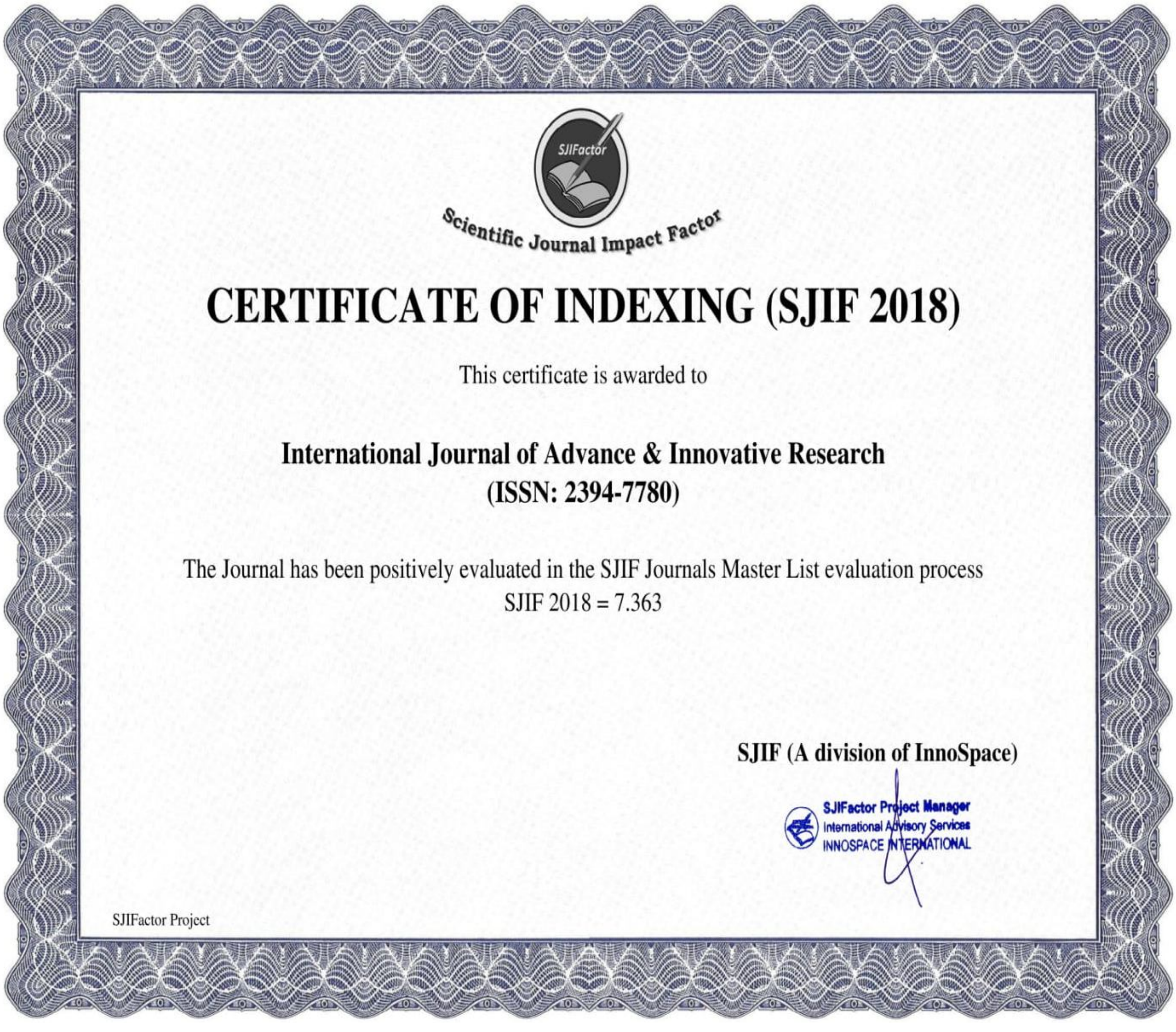
International Journal of Advance and Innovative Research

IJAIR

IMPACT FACTOR

7.36

ISSN : 2394-7780



Scientific Journal Impact Factor

CERTIFICATE OF INDEXING (SJIF 2018)

This certificate is awarded to

International Journal of Advance & Innovative Research
(ISSN: 2394-7780)

The Journal has been positively evaluated in the SJIF Journals Master List evaluation process
SJIF 2018 = 7.363

SJIF (A division of InnoSpace)

SJIFactor Project Manager
International Advisory Services
INNOSPACE INTERNATIONAL

SJIFactor Project

International Journal of Advance and Innovative Research

Volume 7, Issue 1 (XIV) : January – March 2020

CONTENTS

Research Papers

ISSN 2394 - 7780

MOBILE CLOUD COMPUTING– ADVANTAGES, CHALLENGES, AND NEEDS

Punit Harpalani
Chikitsak Samuha's S. S. & L. S. Patkar College of Arts & Science, and V. P. Varde College of Commerce & Economics

ABSTRACT

As cell computing and wi-fi networks grow; mobile subscriptions also increase. This has brought about a boom in cellular programs and services for mobile users. This brings a big studies and business opportunities in mobile computing. In this paper we will talk what tendencies are more inside the marketplace and possibilities and related business using forces Then an outline of cell computing, features, scope and advantages. In addition, its troubles and challenges in cell cloud computing.

Keywords: Cloud computing, cell cloud computing, mobile-cloud services, cell-cloud applications, and mobile computing.

INTRODUCTION

- McCarthy stated that "machines as simple as thermostats can be called trust, and having confidence is a characteristic of most machines capable of problem-solving performance."

- As Increase in progress of cell computing and wireless networks, mobile membership has also increased. According to TechNavio analyst forecasts, Mobile Cloud Computing in North America will grow at 18.12% CAGR in 2011-2015. (Chang, et al., 2013) One of the major factors contributing to the growth is the increasing demand for enterprise mobility. Major primary vendors here include Amazon, IBM, Salesforce.com and others. This advancement brings a significant benefit for users today with cloud infrastructure and platform supply as well as scalability and high resource utilization and sharing.

What is Mobile Cloud Computing?

- Mail and Grace stated that "cloud computing is a version to enable ubiquitous, convenient, on-demand network get right of entry to to a shared pool of configurable resources (eg, networks, servers, storage, applications, and services) that is faster. Can be provisioned from. And launched with minimal management effort or carrier provider interaction. This cloud version on five essential features, three provider fashions and 4 Is made up of planning fashions.

- Mobile cloud computing is a technology thru which cell programs are built, operated and hosted the usage of cloud technology. A cellular cloud computing approach enables developers to expand a program which might be in particular designed for cell users with out being bound by the cellular operating gadget and the computing or memory potential of the smartphone. MCC is generally accessed thru a mobile browser from a faraway web-server, generally with out the putting in a consumer application on the recipient phone.

BENEFIT

1. Flexibility

Mobile cloud computing lets in customers to retrieve and store records from any region from anywhere as lengthy as it's far linked to the Internet. This allows get entry to to easy change of data each time necessary.

2. Multiple platform support

The person can use cellular cloud computing via any platform because cloud computing supports many distinctive styles of platforms to run distinctive applications.

3. Data availability at all times

Through the use of cell cloud applications the user can receive real-time information each time needed. This permits the person to get admission to their facts whenever they want and also saves information on the cloud if the consumer desires to surf offline.

4. Cost Efficiency

This provider of cellular computing may be very pocket-pleasant as there may be no heavy duty because these days this carrier is based totally on pay according to use only.

5. Data Backup

As a continuous generation of new data on cell devices, the cellular cloud application enables the person to backup their statistics at the cloud when it desires to be covered or whilst the facts aren't always in use.

6. Data Recovery

In case of a disaster user losing their important information, the cloud application permits you to get better user records from the cloud by always following a certain process. Recovery of user records from any region is viable if the user is connected to the Internet and has sufficient storage area on their device.

LOSS

1. Data Privacy

Many instances the consumer has touchy content on the cloud and at some stage in the facts flow there can be a breach in the network that can cause statistics loss. It is extremely critical to pick out the right provider issuer that will make certain that consumer facts remains secure always and below any circumstances.

2. Connectivity

When the provider the person desires to use depends completely at the Internet connection, it is important to look that the relationship is up all of the time so that customers do now not should face the cloud connection, which will affect the transfer of their statistics.

Generations of MCC

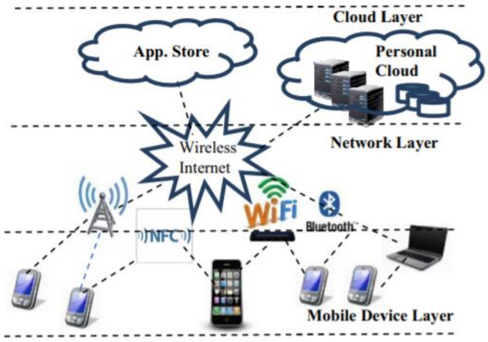
Gen One- Personal Mobile Cloud

In today's terms, the number of mobile cloud for mobile users is. All of them are known as personal clouds that provide mobile data, storage, content, services such as audio, video, photos, calendar and others to the mobile user, it provides the first generation of mobile cloud services to mobile users.

Features of the first-generation mobile cloud:

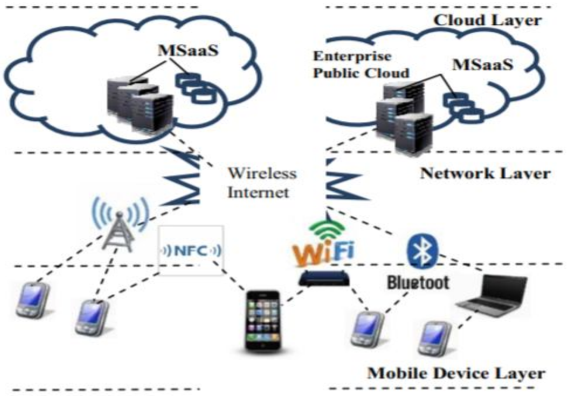
- Mobile application service
- Network communication
- Scalability
- Resource
- Synchronization

O mobility



(First generation – Personal mobile cloud)

Second generation – Cloud-Based Mobile Cloud Infrastructure.



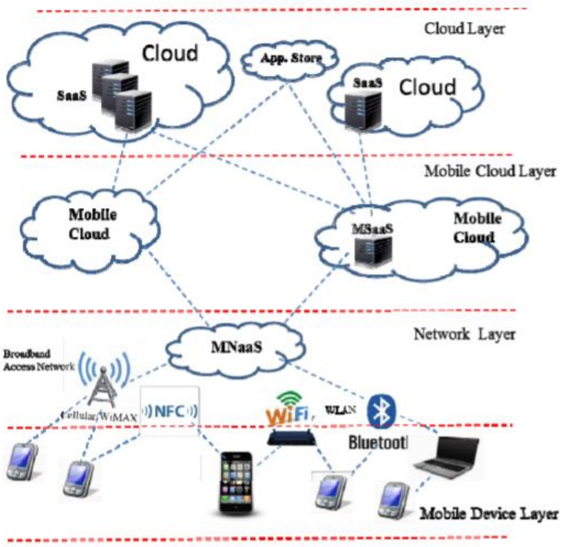
(Second generation – Cloud-Based Mobile Cloud Infrastructure)

Features of Second-Gen Cloud-Based Mobile Cloud Infrastructure

- Mobile application service
- Network communication
- Scalability
- Resources
- Synchronization
- Mobility
- Multi-tenancy
- On-demand service

Gen Third - Services of Mobile cloud

- As pointed out by means of (CEO Mark Heraghty of Virgin media Business)there's an explosive growth in cell information utilization which has led to dramatic shifts in how mobility is utilized in enterprise, to replacing plastic for payments, to emerging technology like SDNs and network virtualization.
- As pointed with the aid of (Lee Chooking) this is great innovation and disruptive time and he trust that today's ICT operation will look drastically exceptional a decade from now. Due to giant boom of cell access site visitors there are sure limitations on this Also, network scalability and restricted scalability in site visitors support
- Career-oriented infrastructure based on network
- Limited portability and connectivity between numerous wireless networks hosted and operated by manner of wireless carrier vendors, network beneficial useful resource sharing, green use of community resources
- Extreme dependence on configured pre-configured bodily network components and devices.
- Reduced Green Computing in Networking



(Third Generation - Mobile cloud services)

Challenges and Solution of Mobile Cloud Computing.

Restricted battery life and computer potential

When it involves battery ability, mobile devices are exceptionally limited as compared to stationary gadgets, where there may be lots room to enforce strong, long-lasting strength sources. The mobile features a confined computational potential that limits its many functions. For instance, the usage of location services like GPS consumes loads of energy because of the sizable use of sensors in it. Similarly, some apps that need an massive processing capability, like photo processing, speech synthesis, tongue processing, augmented reality and wearable computing for video games. Such services constitute a computational mission for software developers, as they are now not able to implement packages that meet the shape of requirement. Given the very fact that this problem is constant by manner of limited battery capability and therefore the small duration of cell devices, hardware is much more likely to get to the bottom of this hassle using software development than hardware.

Fix's

According to (R. Change, 2013) Regarding energy limitations, some researchers confirmed that offloading component interfaces on cloud servers in cell video games can save 27% of battery energy utilization for laptop video games and 45% for chess games. The next solution can be a short-range get admission to point. To perform under immoderate speed connectivity for lengthy distances for cellular devices, the battery is compromised. Therefore, making smaller batteries can make access unrestricted while preserving long battery life (low power consumption).

Connectivity

Maintaining touch with numerous connection mechanisms applied in a project for MCC. In the case of 3G connectivity, there is increased data cost and latency. As clients move, there's variability in sign energy that impedes running processes. This may additionally be due to variability in signal reception or the presence of blind spots that have no connectivity. Compared to forced out networks, mobile gadgets have limited network bandwidth. Signal quality (QoS) delivered to the individual is suffering from non-pressing delays inside the execution of applications, rejecting the same old connectivity and overuse of limited cell resources. Not pleasant does this issue in resources have an effect on the signal, but it furthermore has its impact on the strategies of applications.

Fix's

According to (Satyanarayanan, Gahl, Caceres, & Davis, 2009) , As a manner to this problem, the usage of virtual machine (VM) generation is proposed "to short boost up custom designed issuer software program on close by cloudlets". They advocate that the company be used over a wi-fi LAN. Therefore, for the cause that Cloudlet is a resource-rich reliable laptop with extraordinary connectivity to the Internet, the cellular phone acts as a thin consumer and the purchaser can assure correct connectivity even on the go. In addition, a new generation that offers an answer for facts caching for use for cellular devices is HTML5. This era allows cloud packages to retain performing, notwithstanding the reality that there are a few interruptions to the connection.

Security and Privacy Issues related to data

One of the major troubles going through MCC is records security. (M. Prasad) centered on a few protection troubles related to MCC, collectively with lack of bodily protection, handling encryption and decryption keys, safety and auditing problems of digital machines, low necessities for records integrity, and company platform from various agencies Includes inconsistencies. In addition, while records is loaded onto the cloud, there are problems about information protection and privacy. The offloading approach places the client's data at danger of facts breach and invasion of private data. There is also issue about motive violations, that are seen as hacking a particular tool to sabotage a particular character or steal vital data. Although exceptional from hackers who breach random clients' statistics, such pre-determined attacks can be even greater bad and negatively effect consumer privacy. Any man or woman can get right of get entry to to sensitive data if the sufferer fails to protect the purchaser's statistics. Social media and online charge appears to be the weakest beneficial beneficial useful resource for hacking as it carries essential and important data for clients and customers seem like tons masses less careful specifically at the same time as sharing statistics with social media

Fix's

One of the measures that may be taken to treatment the piracy problem is the deliver of encryption and decryption keys to get entry to the ones materials. The statistics integrity issue can be solved with the resource of authenticating each get right of entry to example via users. A right authentication process needs to be in area to solid mobile facts get right of access to. Regarding community security, some researcher suggested the use of SDN's features, which includes its centralized control, which permits a dynamic authentication of valid hosts primarily based on records approximately the host received all through registration.

The Act of malicious software in Security

The huge range of mobile packages used by cellular packages to use other cell gadgets by cell telephones is an appealing medium for malware creators. This trouble isn't any longer because of unauthorized users gaining access to unauthorized facts. Rather, it's far taking place because people are agreeing to put in malware on cellular gadgets that may transfer and leak personal information to malware creators.

Fix's

There is extra than one way to clear up this problem. First, mobile users want to use antispyware and anti-malware applications and recognize what they count on when the use of a cloud-connected cell device. Second, the cloud can layout and put into effect its very personal apps. In this manner, in the case of a safety thread, the cloud can repair lost records from dependable backup within the cloud itself. Constant improvement within the cloud infrastructure enhances safety in greater strategies than one and offers an average secure experience for all cellular users.

- **Future scope**

Mobile computing technology is one thing that makes it possible for people to use network services wherever possible and at whatever time they want, with proper internet connection. (Helal) Mobile computing has grown significantly and it is said that in future, mobile computing will control almost all technological activities in the world.

- **CONCLUSION**

This paper focuses on mobile cloud computing, mobile cloud services and mobile cloud computing.

Then, the paper discusses the research scope related to mobile cloud computing.

In particular, it presents three generations of cellular cloud provider infrastructure through evaluating their key capabilities and limitations.

In addition, the paper also discusses the future scope of mobile cloud computing.

REFERENCES

1. CEO Mark Heraghty of Virgin media Business. (n.d.).
2. Chang, R.-S., Gao, J., Gruhn, V., He, J., Roussos, G., & Tsai, W.-T. (2013). Mobile Cloud Computing Research – Issues, Challenges, and Needs. IEEE Seventh International Symposium on Service-Oriented System Engineering.
3. Dijiang Huang, Z. Z. (n.d.). Secure Data Processing Framework for Mobile.
4. Helal, A. (. (n.d.). Anytime, Anywhere Computing : Mobile Computing Concepts and Technology. Kluwer Academic.
5. Lee Chooking, i. (. (n.d.).
6. M. Alzadeh, W. H. (2013). challenges and opportunities of mobile cloud computing.
7. M. Alzadeh, W. H. (Sardinia, 2013). Challenges and opportunities of Mobile Cloud Computing.
8. M. Prasad, J. G. (n.d.). Mobile Cloud Computing. International Journal of Computer Application.
9. McCarthy, J. (1979). Ascribing Mental Qualities to Machines.
10. Mell, P., & Grance, T. (2011, September). The NIST Definition of Cloud Computing. The NIST , pp. 800-145.
11. R. Change, J. G. (2013). Mobile Cloud Computing Research Issues, Challenges, and Needs. in 7th International Symposium on Service-Oriented System Engineering,.
12. Satyanarayanan, M., Gahl, P., Caceres, R., & Davis, N. (2009). Mobile cloud computing as future for mobile applications - Implementation methods and challenging issues. IEEE International Conference on Cloud Computing and Intelligence Systems, Beijing, 2011.

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

TO DETECT AND UNDERSTAND BRAIN TUMOR USING A SOFT-COMPUTING ASSISTED TOOL

Amar Dashrath Dhuri
Department of M.Sc. IT, S.S & L.S. Patkar Varde College, University of Mumbai, Mumbai

ABSTRACT

Brain cancer is one of the infection diseases in the human community that are increasing nowadays. Early identification is broadly wanted to offer better treatment so as to spare the patient. In the literature, considerable procedures are proposed and implemented by the researchers. Mainly the brain malignancy is due to the abnormal lump in the brain neural system called a tumor and it can be detected using procedures, such as Magnetic Resonance (MR) imaging and Computed Tomography (CT). Normally, the MR imaging is widely considered initial imaging procedure due to its multi-modality nature and it is also efficient in providing superior information compared with CT. Premature recognition may reduce the severity of sickness and also diminish the death rate. In this work, heuristic approach supported automated tool is proposed to mine and analyse the tumor part from the Magnetic resonance imaging (MRI) of brain. Both of the Bat Algorithm and Kapur's Entropy is chosen as the beginning process section to improve the visibility of tumors region and the Active Contour (AC) is then chosen as the last processing section to extract the MRI image of tumor of brain. The proposed tool is experimentally investigated using various bench mark brain MRI dataset. The results confirm that, proposed tool is efficient in extracting the abnormal region from the considered brain MR test images. This tool defines the value of Jaccard, Dice, sensitivity, specificity and accuracy. From the after effects of the trial work, it very well may be noticed that, the proposed methodology offers better outcome on the dim scaled mind MRI picture. At last, the clinical hugeness by contributing unrivaled estimations of picture likeness and factual measure esteems.

Keywords: Tumor, Entropy, Active contour, Analysis, Brain MRI, Bat Algorithm

I. INTRODUCTION

Lately years, processed helped approaches are broadly embraced to mechanize the modern and medical industry.[1-4] The PC helped techniques are considered because of its exactness, straightforwardness and versatility. In medicinal field, these methodologies are ordinarily considered to help the specialists in illness assessment, treatment arranging and medical procedure. In restorative field, imaging method is generally considered to break down the inner body parts and its diseases. Malignancy is dangerous malady broadly influences the interior body organs. For example, cerebrum, bone marrow, lungs, and so on.

The principle point of this paper is to understand a mechanized device to inspect the tumor areas from the mind Magnetic Resonance Imaging (MRI) database. The picture preparing writing affirms that, heuristic calculation approach is proficient in breaking down the customary and therapeutic pictures[5-10]. Past works additionally affirms that, various picture preparing methods are proposed by the specialists to research the malignant growth with the assistance of clinical pictures.[11]

In the mention work, a best heuristic methodology, known as the Bat Algorithm (BA) is considered to investigate the cerebrum MRI dataset. The heuristic methodology based picture preparing is generally examined because of its effortlessness and tolerance in usage[12-16]. Kapur's entropy is one of the methodology, broadly utilized as a result of its predominance and versatility. Kapur's entropy based multi-thresholding was to a great extent talked about by the specialists to extricate the significant data from the RGB/dim scale test pictures.[17-21] The capacity of proposed division task is then affirmed by methods for a similar appraisal among the sectioned tumor area and the ground truth picture offered by the master part. The experimental results confirm that, proposed approach is effective in acquiring the better picture quality and comparability list qualities. The ongoing outcome additionally affirms the prower of the proposed methodology. Consequently, the proposed strategy is clinically noteworthy and in future, this cost can be utilized to help the specialist to investigate the mind tumor pictures recorded with Magnetic Resonance procedure.

II. MATERIALS AND METHOD

The materials used in the detection process are MRI images of brain tumor and different soft computing tools. These tools provide for automation of brain tumor detection. The materials have been described as follows:

a) MRI Images

MRI is a non-invasive technology used by the neurologist to detect the presence of any abnormality in the human brain. As compared to other imaging technologies like CT Scan, PET Scan etc. MRI Imaging has preferred in tumor detection because of its various advantages. Brain tumor MRI picture as below in Figure 1, have been used here to analyse the presence of tumor.

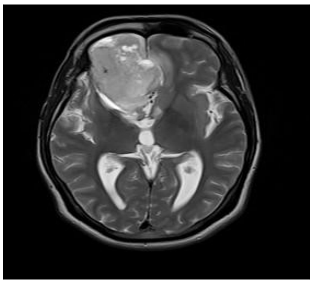


Figure-1: A Brain Tumor MRI Image

b) Soft Computing Tools: Soft computing are a group of methodologies that is meant to handle real life problems by using analytical and reasoning capacity. These tools are emerging as a promising help in different areas of science and technology like image processing, pattern recognition, in data clustering, in medical diagnosis etc. and many more. Soft Computing has the ability to handle uncertainty, imprecision, partial truth and provide better solution in comparable to a human mind. There are different soft computing tools like Neural Networks, Fuzzy Logic, Evolutionary Algorithm, Support Vector Machines etc.

III. METHODS OF USE

The procedure utilized in mind tumor identification starts with the progression of most reduced degree of data extraction known as preprocessing pursued by highlight extraction and arrangement as appeared in fig

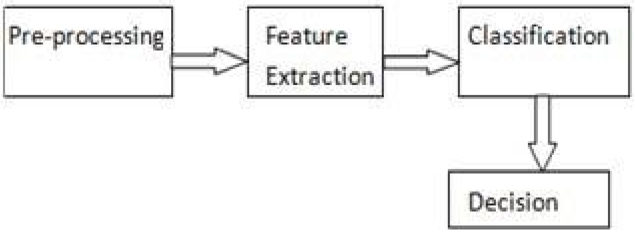


Figure-2: Methodology of Brain Tumor Detection

The information MRI picture needs to initially experience a progression of preprocessing steps like change of the picture to dim scale, histogram balance, binarization, edge recognition and so forth to get least measure of data of the info picture. Pre-preparing is then trailed by highlight extraction, where significant highlights are removed from the info information which contains important data about the info picture. Dim Level Co-event lattice has been utilized by numerous specialists for highlight extraction alongside wavelet based pre-handling. Head Component Analysis have additionally been utilized for include extraction. At long last grouping is performed which orders the tumor mind and ordinary cerebrum pictures. Numerous strategies have been proposed for order utilizing neural system calculations The upside of utilizing ANN is the necessity of less time utilization in identification of enormous measure of MRI pictures. These delicate registering apparatuses in reality demonstrates to be a productive technique for location of mind tumor.[20,21]

A. Bat algorithm

Isolation of tumor area in mind MR picture is a conspicuous undertaking that immediately gives simpler tumor conclusion, which prompts powerful radiotherapy arranging. For a considerable length of time together, a few division techniques for a mind tumor have been displayed and as of not long ago, improved tumor division method will in general be a difficult undertaking since, MR pictures are for the most part innate with shifted tumor measurements of disproportioned limits. To address this issue, we build up an improved mind picture division strategy called BAT based Interval Type-2 Fuzzy C-Means (BAT-IT2FCM) grouping. The BAT calculation is used to discover the ideal group area from which the bunching activity by Interval Type-2 Fuzzy C[12-16].

B. Kapur’s entropy

Image division is a fundamental piece of picture investigation, which directly affects the nature of picture examination results. Thresholding is straightforward and broadly utilized techniques for picture division.

Thresholding can be either bi-level, which includes parceling of a picture into two sections, or staggered, which segments a picture into different fragments utilizing numerous limits esteems. This paper centers around staggered thresholding. A decent division plot through staggered thresholding recognizes appropriate limit esteems to upgrade between-class fluctuation or entropy foundation. For such enhancements, nature propelled metaheuristic calculations are normally utilized. This paper displays a Kapur's entropy based Crow Search Algorithm (CSA) to appraise ideal estimations of staggered limits. Crow Search Algorithm depends on the smart conduct of crow rush. Crow Search Algorithm have indicated better outcomes on account of less number of parameters, no untimely assembly, and better investigation abuse balance in the inquiry procedure. Kapur's entropy is utilized as a target work during the advancement procedure. The analyses have been performed on benchmarked pictures for various edge esteems (for example 2, 4, 8, 16, 32 edges). The proposed technique has been surveyed and execution is contrasted and surely understood metaheuristic streamlining strategies like Particle Swarm Optimization (PSO), Differential Evolution (DE), Gray Wolf Optimizer (GWO), Moth-Flame Optimization (MFO) and Cuckoo Search (CS). Test results have been assessed subjectively and quantitatively by utilizing great performed assessment techniques specifically PSNR, SSIM, and FSIM. Computational time and Wilcoxon p-type esteem additionally thought about. Exploratory outcomes show that proposed calculation performed superior to PSO, DE, GWO, MFO and CS regarding quality and consistency.[10,19]

IV. Brain Tumor recognition utilizing diverse Soft Computing Tools

1)Fuzzy Logic: The utilization of Fuzzy Logic is to deal with imprecision and incomplete truth. Its utilized has been increased quickly in picture handling system. Fuzzy Cmeans is picking up significance as information grouping system. Changed fluffy c-implies is additionally being utilized by numerous specialists in legitimate tumor division. Fluffy intellectual maps are likewise being utilized for poor quality tumor characterization. [20]

2)Artificial Neural Network (ANN): ANN is an organically propelled figuring calculation which is a model of the human focal sensory system. ANN has two methods of activity: the preparation mode and the testing mode. The ANN is utilized in characterization of cerebrum tumor . MLP is likewise proposed by numerous analysts for order of typical and tumor cerebrum pictures. A MPL is a feed forward neural system comprising of numerous layers of hubs as appeared in Figure3.

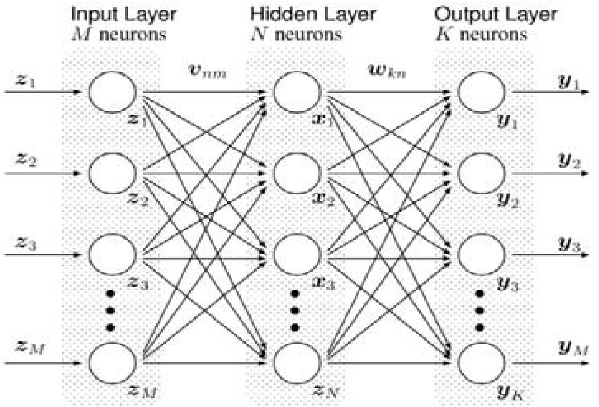


Figure 3: Multilayer Perceptron(MLP)

In the above referred to figure of MLP M, N and K speaks to include, covered up and yield layers separately. "z" indicates a lot of sign which is contribution to the neurons in the information layers of the system. In the shrouded layers every neuron takes the weighted total of the yield signals from the info layer. The (association) weight from the n-th neuron in the info layer to the m-th neuron in the concealed layer is signified by "vnm". The yield of the neuron in the concealed layer, signified by xn, is controlled by the condition[20,22]

$$x_n = g\left(\sum_{m=1} V_{nm}\ z_m\right)$$

Where g is a quaternionic actuation work presenting no n-linearity between the activity potential and yield in the neuron. What's more, the yield of the neuron in the yield layer is shown by the condition

$$y_k = h(\sum_{n-1}^{N} w_{kn}\ x_n)$$

Where is the association weight between the neuron in the shrouded layer and the neuron in the yield layer.

3) Genetic Algorithm: Genetic calculation is a bio-roused search calculation created by the analysts for take care of enhancement issues and other hunt issues. Hereditary calculation is likewise being favored as a proficient device for mind tumor identification. Scientists have likewise utilized techniques like watershed division alongside hereditary calculation for better exactness of tumor recognition. [24-25]

4). Current Status of Research

Brain Mind Image Segmentation is one of the most productive methods utilized now-a-days in identification of cerebrum tumor. Different new techniques have been proposed for location of cerebrum tumor. Division systems like Edge based division, Histogram Thresholding strategies, utilization of Morphological and Water Shed techniques and so forth have been utilized for tumor identification. In any case, it was discovered that manual division of mind needs giving legitimate and the necessary data and is time destroying, for which now the consideration is coordinated towards robotization of cerebrum picture division which is relied upon to give preferable outcomes over prior division strategies in less calculation time. Additionally, in the present current reality where countless MRI pieces of information are taken for a solitary patient, the utilization of non-canny strategies are an excess of work broad, so the analysts are being constrained to discover options utilizing delicate registering devices. Apparatuses like ANN, Fuzzy Logic, Genetic Algorithm are picking up significance right now.[24-25]

V.CONCLUSION

Legitimate location of mind tumor is essential for the nervous system specialist to complete further analysis and treatment. Right now, data is acquired utilizing shrewd techniques that includes the utilization of delicate registering apparatuses. Mind Image Segmentation is one of the major testing task in the present restorative imaging. Confinements of manual division compelled the scientists to coordinate their consideration towards mechanization of mind tumor location. In this way started the time of utilizing delicate figuring instruments in tumor discovery. Delicate figuring instruments utilizing fluffy c-implies, neural system, hereditary calculation have picked up significance in this control. The key commitment of these devices are that they can deal with genuine issues, dissect them and give astute arrangements when contrasted with human personality. The utilization of these apparatuses decreases the time utilization just as it limits the human endeavors required

REFERENCES

[1] R.Yogamangalam, and B. Karthikeyan, “Segmentation Techniques Comparison in Image Processing ,” International Journal of Engineering and Technology (IJET), vol.5, no.1, pp.307-313, 2013

[2] N. G. Thangamma, and S. Prasanna, “Design And Development Of Medical Image Processing Techniques and to Study their Applications Using Graphical System Design in Ovarian Cancer,” International Journal of Engineering and Technology (IJET), vol.8, no.2, pp.1252-1255, 2016

[3] Arif Widyatama , Okto Dinaryanto, Indarto, and Deendarlianto, “The Application of Digital Image Processing to Study Slug Flow Characteristics in A Horizontal Pipe,” International Journal of Engineering and Technology (IJET), vol.8, no.6, pp.2654-2663, 2017.

[4] G.E. Hemamalini, and J Prakash, “Medical Image Analysis of Image Segmentation and Registration Techniques,” International Journal of Engineering and Technology (IJET), vol.8, no.5, pp.2234-2241, 2016.

[5] S. Samanta, S. Acharjee, A. Mukherjee , D. Das and N. Dey, “Ant Weight Lifting Algorithm for Image Segmentation,” IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Madurai, Dec 26-28, pp. 1 - 5, 2013. DOI: 10.1109/ICCIC.2013.6724160.

[6] S. Bose, A. Mukherjee, Madhulika, S. Chakraborty, S. Samanta , and N. Dey, “Parallel image segmentation using multi-threading and k-means algorithm,” IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Madurai, Dec 26-28, 2013. DOI: 10.1109/ICCIC.2013.6724171.

[7] M. Tuba, “Multilevel image thresholding by nature-inspired algorithms: A short review,” Computer Science Journal of Moldova, vol. 22, pp.318-38, 2014.

[8] R. Kumar, F.A. Talukdar, N. Dey, A. S. Ashour, V. Santhi, V. E. Balas, and F. Shi, “Histogram thresholding in image segmentation: a joint level set method and lattice boltzmann method based approach,” Advances in Intelligent Systems and Computing, vol.455, pp 529-539, 2016.

[9] P. Roy, S. Goswami, S. Chakraborty, A. Taher Azar, and N. Dey, “Image segmentation using rough set theory: A review,” International Journal of Rough Sets and Data Analysis (IJRSDA), vol.1, no.2, pp.62-74, 2014.

[10] B. Akay, “A study on particle swarm optimization and artificial bee colony algorithms for multilevel thresholding,” Applied Soft Computing Journal, vol. 13, no. 6, pp. 3066–3091, 2013.

[11] Palani T Krishnan, Parvathavarthini Balasubramanian, Chitra Krishnan, “Segmentation of brain regions by integrating meta heuristic multilevel threshold with markov random field,” Current Medical Imaging Reviews, vol.12, no.1, pp. 4-12, 2016.

[12] X. S. Yang, Nature-Inspired Metaheuristic Algorithms, Luniver Press, Frome, UK, 2nd edition, 2011.

[13] B. Joyce Preethi, R. Angel Sujitha, and V. Rajinikanth, “Otsu based Multi-level Image Segmentation using Brownian Bat Algorithm,” International Journal of Computer Applications, ICCCMIT 2014, no.3, pp. 10-16, 2015.

[14] V. Rajinikanth, J.P. Aashiha, and A .Atchaya, “Gray-level histogram based multilevel threshold selection with bat algorithm,” International Journal of Computer Applications, vol.93, no.16, pp.1-8, 2014.

[15] Suresh Chandra Satapathy, N. Sri Madhava Raja, V. Rajinikanth, Amira. S.. Ashour, and Nilanjan Dey, “Multi-level image thresholding using Otsu and chaotic bat algorithm,” Neural Computing and Applications, pp.1-23, 2016. DOI: 10.1007/s00521-0162645-5.

[16] X-S. Yang, ‘‘Bat algorithm: Literature review and applications,’’ International Journal of Bio-Inspired Computation, vol. 5, no.3, pp.141–149. 2013.

[17] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, “A new method for gray-level picture thresholding using the entropy of the histogram,” Computer Vision, Graphics, and Image Processing, vol. 29, no. 3, pp. 273–285, 1985.

[18] A.K. Bhandari, A.Kumar, and G.K Singh, “Modified artificial bee colony based computationally efficient multilevel thresholding for satellite image segmentation using Kapur’s, Otsu and Tsallis functions,” Expert Systems with Applications, vol. 42, pp.1573–1601, 2015.

[19] V. S. Lakshmi, S.G. Tebby, D. Shriranjani, and V. Rajinikanth, “Chaotic cuckoo search and Kapur/Tsallis approach in segmentation of t.cruzi from blood smear images,” International Journal of Computer Science and Information Security (IJCSIS), vol.14, CIC 2016, pp. 51-56, 2016.

[20] Roshan G. Selkar, Prof. M. N. Thakare “ Brain tumor detection and segmentation using thresholding and watershed segmentation,” IJAICT Vol1, Issue 3, July 2014

[21] Swapnali Sawakare and Dimple Chaudhari, “Classification of Brain Tumor Using Discrete Wavelet Transform, Principal Component Analysis and Probabilistic Neural Network,” International journal for research and emerging science, vol-3 November-2014

[22] Indah Soesanti, Adhi Susanto1, Thomas Sri Widodo1, Maesadji Tjokronagoro, Optimized fuzzy logic application for MRI brain image segmentation,” International Journal of Computer Science & Information Technology (IJCSIT) Vol 3, No 5, Oct 2011.

[23] E.I. Papageorgiou ,P.P. Spyridonos, D. Th. GlotsosC.D. Stylios, P. Ravazoula,G.N. Nikiforidis, P.P. Groumpos, “Brain tumor characterization using the soft computing technique of fuzzy cognitive maps,” Applied Soft Computing 8 (2008)

[24] Kailash Sinha1,, G.R.Sinha, “Segmentation of Brain MRI Images for Tumor extraction by combining Cmeans clustering and Watershed algorithm with Genetic Algorithm,” International Journal of Digital Application & Contemporary Research, Volume 1, Issue 1, August 2012 [16] Amanpreet Kaur, Gagan Jindal, “Tumor Detection Using Genetic Algorithm,” Ijstc Vol. 4, Issue 1, Jan - March 2013

[25] Minakshi Sharma, Dr. Saurabh Mukherjee, “Fuzzy CMeans, ANFIS and Genetic Algorithm for Segmenting Astrocytoma –A Type of Brain Tumor” International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue 6, June-2013, pp. 852-858.

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

CLOUD CRYPTOGRAPHY

Ayesha Ansari
Chikitsak Samuha’s S.S. & L.S. Patkar College of Arts & Science, and V.P. Varde College of Commerce & Economics

ABSTRACT

Cloud computing is the most in demand technology nowadays and has become the technology for the coming generation. This technology has changed the old computing technologies into new and advanced technologies. Many advantages are being utilized in the field of IT companies by this technology, even though it has to overcome many challenges to satisfy its maturity level. This paper discusses and explains about cloud computing and cryptographic encryption methods and algorithms used to enhance the security of the cloud.

Keywords: cloud storage, encrypted data, searchable encryption, searchable symmetric encryption, public key encryption, Cryptography

I. INTRODUCTION

A. What is a Cloud?

"The cloud" alludes to servers that are gotten to over the Internet, and the product and databases that sudden spike in demand for those servers. Cloud servers are placed in server farms everywhere throughout the world. Running applications or checking physicals servers from time to time is not required by clients and organizations as everything is virtually stored in cloud.

B. What is Cloud computing?

Cloud computing is the collection of networks. The modalities of cloud computing can be used by the user whenever demanded or required. Instead of fixing their own physical infrastructure, the users ordinarily prefer a mediator provider for the service of the web in cloud computing. This technology uses pay per use policy. The workload can be shifted to reduce the workload in cloud computing. A load of service is handled by the networks which forms the cloud that's why the load on local computers isn't heavy while running an application. To use cloud computing a web browser is sufficient

C. Characteristics of Cloud Computing:

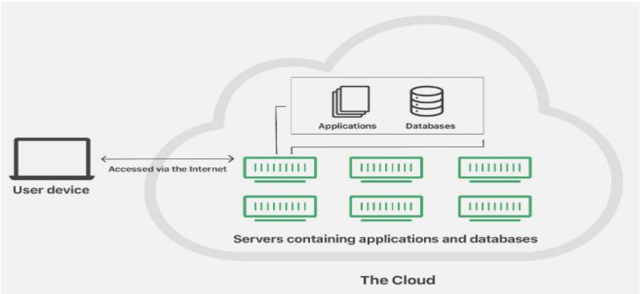
1.**On-request self-administration**: A user can make changes in the arrangement of computing capabilities, for example, server time and system storage, varying consequently without contacting and system administration for their use as per need.

2.**Broad Network access**: Users can use and access the cloud as per their wish on any device in the network. (e.g., cell phones, tablets, PCs, and workstations).

3.**Resource pooling**: The service provider’s assets and tools are designed to serve different consumers by using a multi-tasking model having various virtual and physical resources. These resources can be reused once the other user has finished using it.

4.**Rapid versatility:** Rapid Versatility is used to describe the scalability of cloud or the capability by which cloud provides scalability. Cloud resources are easily acquired by the consumers in any quantity with the help of highly scalable cloud infrastructure. The cloud infrastructure is designed in such a way that it can easily provide services as per user's changing demands. If the resources are not fully utilized by the user the cloud resources are scaled down to avoid resource wastage.

5.**Measured assistance:** The use of cloud resources is generally controlled by cloud frameworks. This is done to measure and monitor the services for a number of reasons including planning of future use, effective utilization of resources and billing as per use.


Applications
Databases
User device
Accessed via the Internet
Servers containing applications and databases
The Cloud

II. SERVICE MODELS

Despite the fact that cloud administrators gives "Everything as a Service" (with the abbreviations EaaS or XaaS, or just aas), Cloud computing suppliers offer their "services" as per various models, of which the three standard models according to NIST are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Increasing abstractions are offered by these models which are portrayed as stack layers. For example, a provider can provide PaaS or IaaS layers without where programs can be run without using SaaS.

1. Software-as-a-Service (SaaS)

SaaS applications are facilitated on cloud servers, rather than clients introducing an application on their gadget, and clients get to use them over the Internet. SaaS resembles leasing a house: the owner keeps up the house, yet the occupant utilizes the house by paying rent for it and possess it for the decided period of time.

Examples - Office 365, Google Apps, Salesforce, Citrix GoToMeeting, Cisco WebEx and Netflix**.**

2. Platform-as-a-Service (PaaS)

In this model, for building own applications organizations pay for resources required to make those applications. The resources required for building an application, including advancement instruments, framework, and working frameworks, over the Internet are offered by PaaS providers. PaaS can be contrasted as the whole house is not leased rather all of the apparatuses and hardware fundamental for building a house are leased to make desired house

Examples - Heroku and Microsoft Azure.

3. Infrastructure-as-a-Service (IaaS)

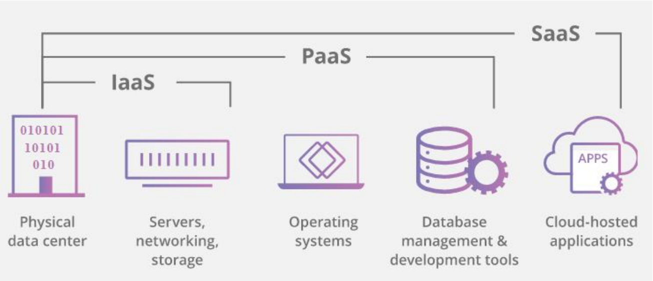
In this model, the cloud supplier gives the servers and capacity on lease to organizations for their utilization. The organizations at that point utilize that cloud foundation to fabricate their applications or programming projects. IaaS resembles as an organization can rent a plot of land on which they can construct anything but they need to give their structure and materials to make the desired thing.

Examples - DigitalOcean, Google Compute Engine, and OpenStack.

Some time ago, SaaS, PaaS, and IaaS were the three primary models of cloud computing, and basically all cloud administrations got fit into one of these classifications. Recently a fourth model has come up:

4. Function-as-a-Service (FaaS)

FaaS, otherwise called server less processing, separates cloud applications into much littler parts that are possibly used when they're required. Suppose it was conceivable to lease a house each smidgen in turn: for example, the inhabitant pays for the lounge area at supper time, the room while they're dozing, the parlor while they're sitting in front of the TV, and when they aren't utilizing those rooms, they don't need to pay their rent.


SaaS
PaaS
IaaS
010101
10101
010
APPS
Physical data center
Servers, networking, storage
Operating systems
Database management & development tools
Cloud-hosted applications

III. CLOUD DEPLOYMENT MODELS

The most common cloud deployments are:

1. **Private cloud:** A server, data center or distributed network which is only accessible with one organization for use is a private cloud.
2. **Public cloud:** A service run by an external vendors which may include servers in one or multiple data centers is known as public cloud. Multiple organizations share public clouds as per their requirements.
3. **Hybrid cloud:** The combination of public and private clouds is known as Hybrid clouds. They may also include on-premises legacy servers. An organization may utilize their private cloud for some accommodations and their public cloud for others or they may utilize the public cloud as backup for their private cloud.
4. **Multicloud**: The use of multiple public clouds is known as Multicloud. In other words, an organization with a multicloud deployment can rent virtual servers and services from several external vendors – to

continue the analogy used above, this is like leasing many adjacent plots of land from different owners. Multicloud deployments can also be hybrid cloud, and vice versa.

IV. CRYPTOGRAPHY

A. What is Cryptography?

Cryptography is the craft of ensuring data by changing it (scrambling it) into a mixed-up group, called figure content. Just the individuals who have a mystery key can interpret (or unscramble) the message into plain content. Scrambled messages can here and there be broken by cryptanalysis, additionally called codebreaking, although current cryptography strategies are for all intents and purposes unbreakable.

B. Cloud Cryptography:

Cryptography in the cloud utilizes encryption systems to verify information that will be utilized or put away in the cloud. It sanctions clients to helpfully and safely get to shared cloud administrations, as any information that is facilitated by cloud suppliers is secured with encryption. Cryptography in the cloud ensures touchy information without postponing data trade.

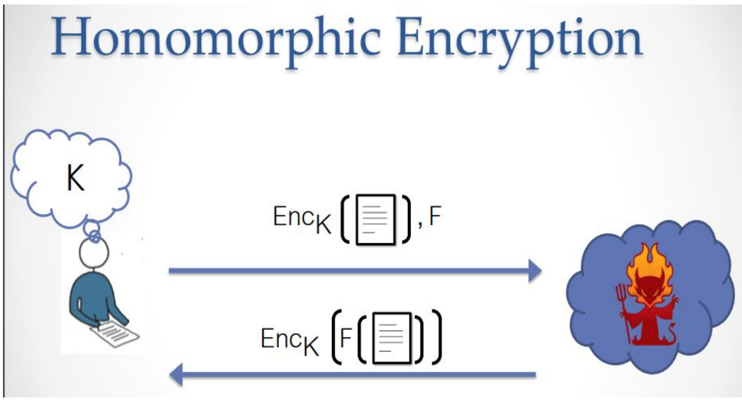
Cryptography in the cloud takes into consideration verifying basic information past your corporate IT condition, where that information is never again heavily influenced by you. Cryptography master **Ralph Spencer Poore** explains that "*information in motion and information at rest are best protected by cryptographic security measures. In the cloud, we don't have the luxury of having actual, physical control over the storage of information, so the only way we can ensure that the information is protected is for it to be stored cryptographically, with us maintaining control of the cryptographic key*."

V. TYPES OF ENCRYPTION

1. Homomorphic Encryption:

It is an encryption calculation that give exceptional calculation facility over encoded information (figure content) and return scrambled outcome. This calculation can comprehend numerous issues identified with security and privacy issues. In this calculation encryption and unscrambling occurring in customer site and supplier site works upon encoded information. This can understand danger while moving information among customer and specialist co-op, it conceals the plaintext from specialist organization, supplier works upon figure message as it were.

Homomorphic encryption enables complex numerical activities to be performed on encoded information without utilizing the first information. For plaintexts X1 and X2 and comparing ciphertext Y1 and Y2, a Homomorphic encryption plot allows the calculation of X1 Θ X2 from Y1 and Y2 without utilizing P1 Θ P2. The cryptosystem is multiplicative or added substance Homomorphic relying on the activity Θ which can be augmentation or option.



2. Searchable encryption

SE (Searchable Encryption) is a positive method to ensure clients delicate information, while saving hunt capacity on the server side. ... The two fundamental parts of SE are:

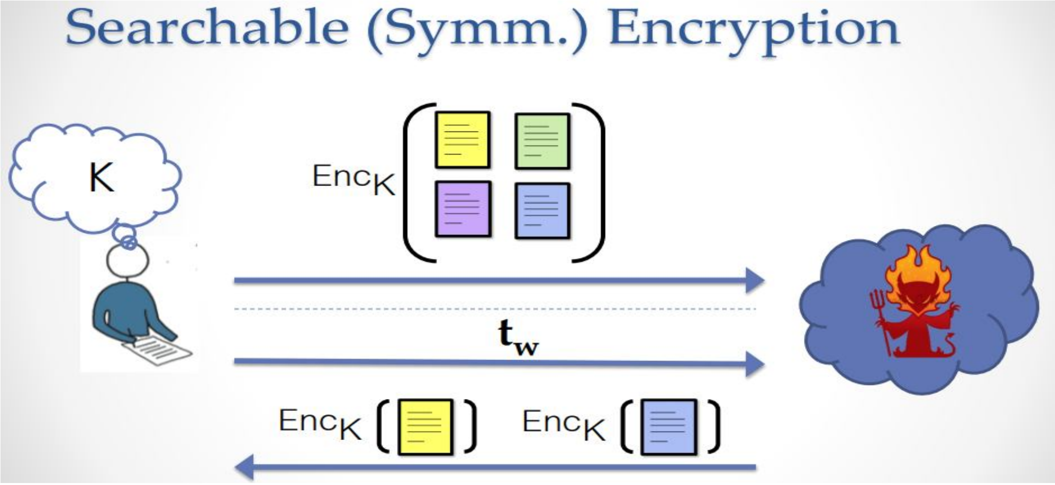
•SSE (Searchable Symmetric Encryption)

•PEKS (Public key Encryption with Keyword Search).

It ensured that cloud never observes and archive and search watchwords.

Its Symmetric variation is productive however uncovers access and search designs.

[GO96] tells the best way to shroud this yet it is costly.



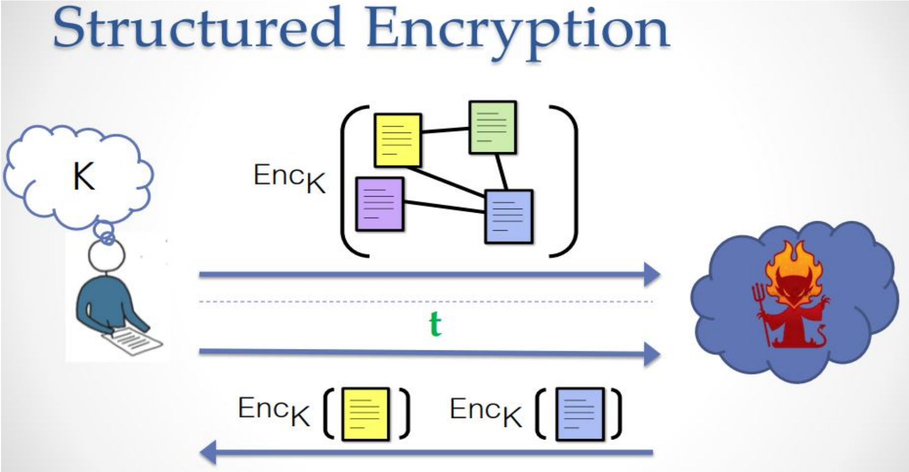
3. Structured encryption

An organized encryption conspire encodes organized information so that it very well may be questioned using an inquiry explicit token that must be produced with information on the mystery key.

The utilizations of organized Encryption are:

- Private questions on scrambled information
- Controlled revelation for nearby calculations
- Query over encoded web charts

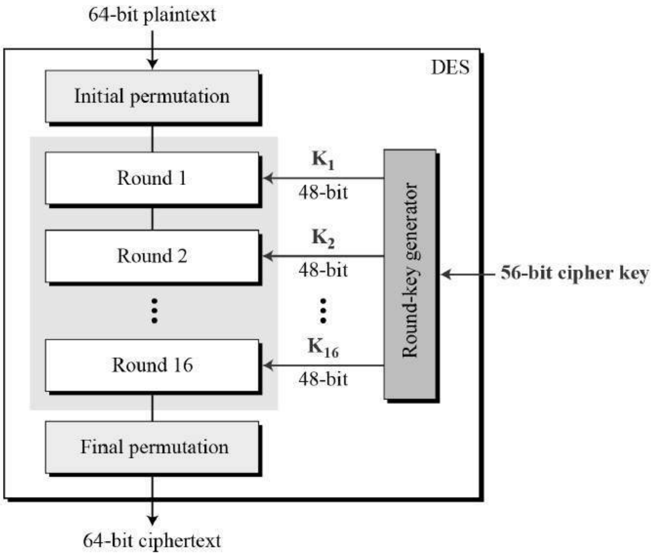
It ensured that cloud never observes information and questions. One burden is that it can Reveals access and search designs



4. Data Encryption Standard (DES):

The Data Encryption Standard (DES) is a symmetric-key square figure distributed by the National Institute of Standards and Technology (NIST). It utilizes single key (mystery key) for both encryption and decoding. It works on 64-piece squares of information with 56 bits key. The round key size is 48 bits. Whole plaintext is isolated into squares of 64bit size; last square is cushioned if vital. Numerous stages and substitutions are utilized all through so as to build the trouble of playing out a cryptanalysis on the figure. DES calculation comprises of two changes (P-boxes) and sixteen Feistel adjusts. Whole activity can isolate into three stage. First stage is Initial change and last stage is the last changes.

1. Introductory stage adjusts the bits of 64-piece plaintext. It isn't utilizing any keys, working in a predefined structure.
2. There are 16 fiestel adjusts in second stage. Each round uses an alternate 48-piece round key applies to the plaintext bits to deliver a 64-piece yield, produced by a predefined calculation. The round-key generator creates sixteen 48-piece keys out of a 56-piece figure key.
3. At long last stage perform Final change, invert activity of starting stage and the yield is 64-piece figure content.



5. Advanced Encryption Standard (AES)

AES is a symmetric-key square figure distributed by the National Institute of Standards and Technology (NIST). Most embraced symmetric encryption is AES. It works calculation on bytes as opposed to bits, treats 128 bits of plaintext obstruct as 16 bytes. These 16 bytes are organized in four sections and four lines for preparing as a lattice. It works on whole information hinder by utilizing substitutions and stages. The key size utilized for an AES figure determines the quantity of change adjusts utilized in the encryption procedure.

Potential keys and number of rounds are as following:

- 12 adjusts for 128-piece keys.
- 12 adjusts for 192-piece keys.
- 14 adjusts for 256-piece keys.

Significant focal points of AES over DES are

1. Information square size is 128 bits.

2. Key size 128/192/256 bits relying upon rendition.

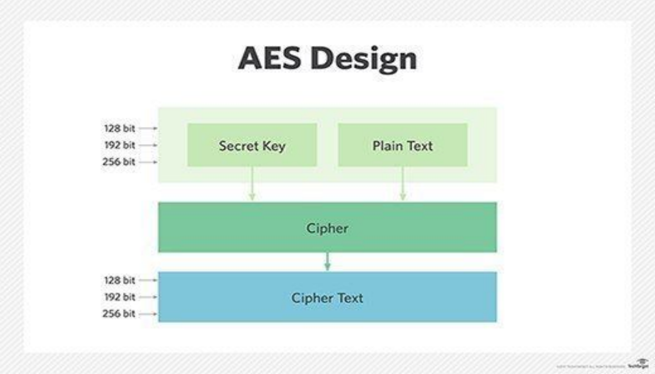
3. Most CPUs presently incorporate equipment AES bolster making it exceptionally quick.

4. It utilizes substitution and changes.

5. Potential keys are 2128, 2192 and 2256

6. More secure than DES.

7. Most embraced symmetric encryption calculation.



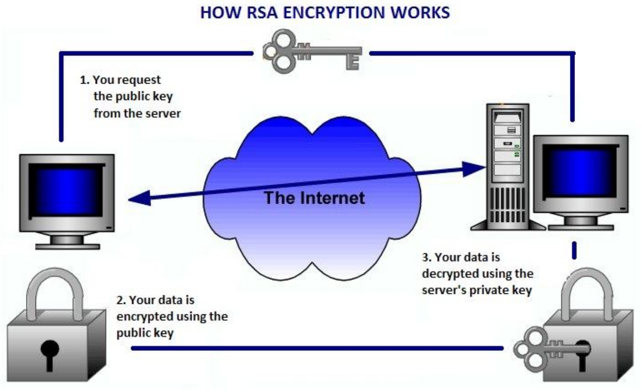
6. Rivest-Shamir-Adleman (RSA)

RSA is an open key figure created by Ron Rivest, Adi Shamir and Len Adlemen in 1977. It is most well-known deviated key cryptographic calculation. This calculation utilizes different information square size and different size keys. It has unbalanced keys for both encryption and unscrambling. It utilizes two prime numbers to create general society and private keys. These two unique keys are utilized for encryption and decoding reason.

This calculation can be comprehensively ordered in to three phases; Key, age by utilizing two prime numbers, encryption and unscrambling.

RSA today is utilized in several product items and can be utilized for key trade, computerized marks, or encryption of little squares of information. This calculation is for the most part utilized for secure correspondence and confirmation upon an open correspondence channel.

While contrasting the presentation of RSA calculation and DES and DES. At the point when we utilize little estimations of p and q (prime numbers) are chosen for the structuring of key, at that point the encryption procedure turns out to be too frail and one can have the option to decode the information by utilizing arbitrary likelihood hypothesis and side channel assaults. Then again on the off chance that huge p and q lengths are chosen, at that point it expends additional time and the presentation gets debased in correlation with DES. Activity speed of RSA Encryption calculations is moderate contrast with symmetric calculations, in addition it isn't verify than DES.


HOW RSA ENCRYPTION WORKS
1. You request the public key from the server
The Internet
2. Your data is encrypted using the public key
3. Your data is decrypted using the server's private key

VI. CONCLUSION

Cloud computing is the widely emerging in the world. It is very useful for coming generations in the world. Although whichever new technology comes in the market has some challenges and some benefits. Security is the most difficult factor in any technology. In this paper, I have mentioned a number of encryption algorithms to overcome this issue which also deals with benefits and advantages og these algorithms. Here we conclude that AES algorithm is the most opportune algorithm in cloud computing surroundings to impenetrable their treasured facts in a network. The competency of AES algorithm to operate actions on encrypted facts permits high protection than different algorithms such as DES, RSA, etc.

REFERENCES

[1] https://digitalguardian.com/blog/cryptography-cloud-securing-cloud-data-encryption

[2] https://www.cloudflare.com/learning/cloud/what-is-the-cloud/

[3] https://cs.brown.edu/~seny/slides/cc-RCLE11.pdf

[4] https://en.wikipedia.org/wiki/Cloud_computing

[5] Secure searchable encryption: a survey published Journal of Communications and Information Networks by WANG Yunling, WANG Jianfen, CHEN Xiaofeng. https://link.springer.com/content/pdf/10.1007%2FBF03391580.pdf

[6] Analysis of Various Encryption Algorithms in Cloud Computing by Nasarul Islam.K.V, Mohamed Riyas.K.V published in International Journal of Computer Science and Mobile Computing

[7] https://www.tutorialspoint.com/cryptography/data_encryption_stan

[8] https://don.p4ge.me/rsa-explained-simply/programming

[9] https://searchsecurity.techtarget.com/definition/Advanced-Encryption-Standard

[10] https://crises-deim.urv.cat/web/docs/publications/journals/948.pdf

[11] https://ijcsmc.com/docs/papers/July2017/V6I7201729.pdf

[12] https://eprint.iacr.org/2011/010.pdf

TESTING INTERNET OF THINGS

Maulik Chandan
Masters in Information Technology, SVKM’s Usha Pravin Gandhi College, Mumbai

ABSTRACT

The Internet of Things (IOT) refers to small embedded components combined together into a network chain of Internet which can be controlled by a centralized technology. With Internet of Things coming into presence the world is increasingly connected to such gadgets with ease of use making it a necessity in human’s day to day life.

IOT technology is all about providing precise, time saving and effortless solution.

Testing becomes a broader region when complex systems like IOT comes into picture, as IOT has a big connected chain of technologies in working. As and when users will get accustomed to this IOT enabled devices connected to the gadgets, their expectations for the new technologies will be to work flawlessly [3].

With increasing market in IOT only the best quality products will be the brand that wins

I. INTRODUCTION TO INTERNET OF THINGS

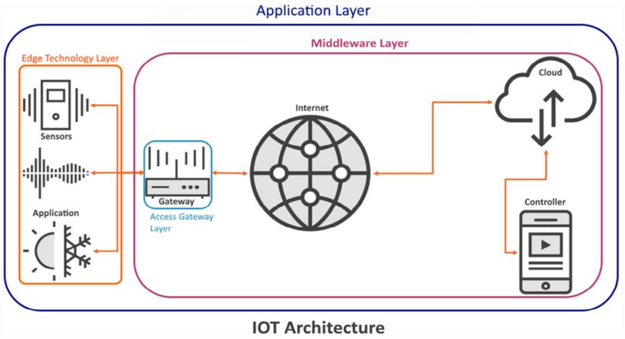


Fig.1: IOT Architecture

The four broad regions of IOT Architecture can be classified as follows: [Fig: 1]

Things (Edge Layer) consists of sensors to collect data as input and actuators to perform certain actions based on those inputs as per commands received from the cloud.[5]

Network gateways are used for data filtering, preprocessing and transferring the same to the cloud to and fro into from the controller and edge layer [5]

- Cloud gateways to ensure transmission between different channels and gateways or centralized servers
- Streaming data processors to ensure that data received from the Input is distributed to all the required solution components that needs to process it.

Data Controller – Database to store all the relevant data with defined and undefined value. [5]

- Data Warehouse to store and process valuable data for Big Data analytics.
- Machine learning algorithms to generate modules that can be used further to control applications.
- Data analytics for manual data processing

Application Controller to send commands to actuators.[5]

- User application with embedded controller to remotely access and monitor with their IOT connected devices

II. NECESSITY TO TEST IOT

With recent trends in the IOT –enabled devices, there is an upsurge in the smart devices in the fields of Consumer utilities, Industrial development, Healthcare, etc. on a Year on year basis.

Current Digital devices not only works with inputs given via keystrokes or clicks, but also operates on sensory, emotional, physical interactions, which makes real world “Human Experience” testing as a very important aspect of testing.[2]

As most of the IOT devices remains independent, Testing must involve all possible permutations and combinations with Operating Systems along with various software's supported. Simulators and Emulators along with stubs and drivers will be required at large to support end-to-end testing [2]. Each layer of IOT will require individual encompassing test labs.

Also, Companies over the world are rapidly switching, producing, and mounting out latest IOT –enabled devices into the market, testing IOT remains to be a big challenge if the approach is not properly determined.

As per the Press Release by Gartner "Following will be the business-critical areas where most organizations will need to pay a very strong attention to in the coming years to sustain in the High quality Low cost driven market."[1] [2] [Fig: 2]

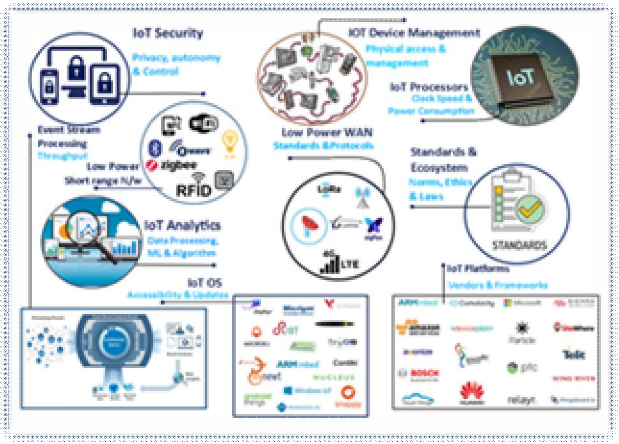


Fig-2: IOT Testing Areas

III. APPROACH

The comprehensive Test Strategy approach should be followed end-to-end to accommodate maximum coverage of IOT testing.

Testing IOT devices along the Business critical areas is broadly covered in following Testing types.

Testing Types in IOT

A. Usability Testing

A usability test determines the device built in such a way that it matches the users expectation upon performing actions which proves the effectiveness and efficiency is satisfying manner. It fairly deals about benefits of product and not its features, and how does it add values to their lives.

- **IOT Device Management –**

Management of devices can be determined by –

> The place where devices are going to be used.

> The target User who is going to use.

> The purpose of using the device.

> The knowledge of what all devices are to be used.

- **IOT Operating System –**

Operating system and UI basically makes possible the usability of the device.

> User Friendly with Ease to access.

> Simple usage interface to familiarize with the system.

> User satisfaction on using the system with ease on performing operations.

B. Compatibility Testing

There are various devices connected into an IOT together for single or multiple applications. These devices have distinct software and hardware and software configuration. Therefore looking at the complex architecture it is necessity to keep testing compatibility of each permutation and combination of this configuration designs. As a result various communities have tried building generic platforms which are highly compatible and reliable.

- **IOT Platforms**

There are various open source and proprietary platforms which together works to bring IOT compatibility to test following aspects.

> Multiple Operating system (respective versions)

> Browser types (respective versions)

> Generation of devices (processors)

> Communication modes (networks),etc.

C. Security Testing

In IOT from security perspective the first thing that needs to be secured is data, than system security, physical security. However each security has interrelation to the data as data is the most critical information that has high stakes which can lead to action. As the market of IOT grows, the devices that are connected directly to enterprises and households to make lives easier and more comfortable at the same time will increase the risks of possible attacks. Data security should be monitored right from its creation, transmission, storage and retrieval. [6]

- **Physical Device Security (Privacy, autonomy, control) –**

> To install Tamper resistant hardware which has strong port and boot level locks.

> Regular Patches and updates should be provided to fill the security holes & stability improvements.

> Data Protection should be tested like if physical security is exploited, data should be disabled or destroyed.

> Performance requirements should ensure less power more processing with capabilities to function even during connectivity disruptions.

> Doing Penetration testing & dynamic code analysis by an ethical hacking perspective.

- **Networks –**

> Network channels should have encrypted data along with network standards from IEEE

> Strong Authentication should be kept at every network nodes.

> Divide large networks into segments to implement firewall using VPN’s

> Limit the network traffic by minimizing the hardware and kernel level bandwidth.

- **Securing Data –**

> Set unique default passwords for new devices along with reset option to delete stored information.

> Data leakage and consumer privacy should be kept by only collecting of necessary data.

> Custom network communications for ensuring that device is visible only to known networks.

D. Performance Testing

Performance testing of IOT is bit different from that of traditional Performance Testing as, IOT Produces data continuously which is saved and analyzed for future decisions using Business Intelligence Analytics. IOT requires operational performance along with cost optimization. Performance can be measured by simulating real-workload models using business requirements and historic data combining with geographic spread, usage patterns and peak usage, normal usage statistics. [7]

- **Protocols and Tools -**

> IOT uses various protocols Hyper Text Transfer Protocol (HTTP), AllJoyn, IOTivity, MQ Telemetry Transport (MQTT), The Constrained Application Protocol (CoAP), Advanced Message Queuing Protocol (AMQP) & more.

> Some performance testing tools support some of this standard protocols

- **Load Conditions -**

> IOT load tests should be done using real world simulations.

> IOT has large geographical spread scope hence needs testing complex patterns.

- **Real Time Decision making -**

> IOT requires fast processing and quick decision making as it has a long chain of connected of dependent devices.

> Request and response time monitoring should be done to test timely performance.

- **Monitoring**

> Monitoring of device performance to check bottlenecks while testing Stress and Endurance.

> Monitoring breakpoints and alternatives of connected device chain.

E. Functional testing

Functional testing of IOT is different from testing basic applications as IOT has thousands of interconnected components together. Hence the goal should be to test most important features and the breakpoints where the integration is done.

- **Power**

> Testing the components behaviour with and without power supply for backup and restore.

> Testing with excess power load to presume hardware efficiency.

- **Embedded computing**

> Processors, sensors and devices should be tested in different environments

> Ensuring physically safe to use device with compactness.

- **Real Time Operating System**

> Timely responsiveness of devices should be tested for safety critical applications.

> Testing real time for machine to machine request, response & act

F. Data Integrity Testing

Data Integration testing is done by ETL (Extract Transform Load) analytics, where huge gathered and processed data is randomly tested with different test conditions along the heterogeneous data sources.

- **Extract data from source systems**

> Only the data that contains information which can be processed as per BI should be extracted

- **Apply transformation logic**

> The machine learning algorithms use this data to transform into logical information

- **Load data into target warehouse**

> Summarized data is loaded which has informative knowledge to improve the performance of end user

G. User Acceptance Testing

The final IOT products developed can be used for either commercial, Industrial or Healthcare where the final testing is done in close to real environment. The UAT is done in two basic type

- **ALPHA testing**

> Testing the product in the development environment itself before production.

> White Box and Black box testing should be done

- **BETA Testing**

> Final product should be released in the real time environment to some exclusive users, for feedback and mass testing.

> In Beta test the data collected is real time and are near proximate.

> Beta Testing should be done at pre-release of every up gradation.

IV. CHALLENGES IN IOT TESTING

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

i. End to End testing:

End to End testing in IOT remains a challenge as IOT is a blend of various technologies which has does not have any common tool for E2E testing.

However there are various independent tools used to test each testing processes which are evolving as with the technology [4].

- Wireshark – Network Testing
- Shodan – Connectivity Testing
- Tcpdump – Command Line network Testing
- SOASTA Cloud Test – Performance Testing

ii. Technology Skill Gap

As there is a blend of various technologies which leaves small gap between Hardware and software as a technology. Testing of hardware is different from that of testing a software. However when tested in integration the expected results are compared with the actual results which leaves the gap of knowledge and also chances of QE miss.

iii. Upgradation

Deploying of upgradations and changes to all the parts of IOT becomes difficult.

As the IOT is connected to thousands of devices together, lot of time and efforts are required every time there is a change or update.

iv. Restricted usage of wireless technology

Smart home automation and other general IOT devices are moreover restricted to the use of Wi-Fi and Bluetooth due to higher efficiency and throughput, regulated standards and traditional usage approach. There are other cost efficient technologies like RFID, Z- Wave, NFC, LORA, Sigfox and NB_IOT as substitutes some of which have Government approvals while others are yet to get. And also QE needs to have knowledge about security and other aspects of testing for the new wireless technologies[4]

v. Security

Security remains the biggest challenge as for IOT devices needs both physical as well as network security. The outreach of IOT connected devices over geographic area makes it difficult to secure remotely.

V. IMPLEMENTATION

- Current Implementation
- Amazon web service individually offers IOT services from the edge to the cloud.
  - AWS provide device compatible software, control services along with security and Data services for analytics.
  - AWS has solutions to 3 of the 6 challenges by managing security, up gradation and end to end testing.
  - Amazon web services has open source development program for IOT enabled devices using Alexa.
  - The programs developed uses Device shadow (Lambda function) which can run code virtually on any type of backend services or applications – that too with zero administration.
- There are other vendors too that provide end to end solutions like Azure IOT Hub, IBM Watson IOT etc.
- Suggested Implementation
- Bridging the Gap between technology understandings for integrated testing.

Any hardware expert should have the same knowledge as of Software expert of the IOT device that is built by learning how hardware interacts with software for better QE understanding and test requirement design.

- Making use of most efficient Wireless technologies as per requirement in the IOT system that is built. Increasing trend of different technologies will enforce the market to get used to migrate to such technology, creating different standards and protocols of its use and enhancing security.
- Maximizing usage of edge gateways to communicate with the end point devices for interconnectivity, data transportation and up gradation.

As when the upgraded for hardware devices is sent from the controller via cloud. Edge gateways can monitor the updates and performance of the end point devices by itself.

VI. CONCLUSION

The goal of this white paper is to cover the broad approach of Testing in IOT and its areas. By 2020, Gartner estimates “Internet-connected things will outnumber humans 4-to-1, creating new dynamics for marketing, sales and customer service.” This will open the market and scope for testing at large. With increasing number of automation technology the approach towards testing will be broad. Also security will be the major challenge due to exposure of technological devices all around.

Gartner predicts “By 2020, more than 25% of identified attacks in enterprises will involve the IOT, although the IOT will account for less than 10% of IT security budgets”[1]. Hence to sustain with IOT enabled products consumers and business have to keep key focus on enhancing existing security domains and introducing new security standards in the market.

A. Authors Biography

Maulik Chandan, Functional Tester, Capgemini

2.5 years of experience in Functional testing

□ ISTQB Foundation Tester Certified

□ Experience in Duck Creek Claims.

□ Expertise in Insurance domain for Auto and General Liability

REFERENCES

[1] Gartner - Global research and advisory firm providing information, “Gartner Says Worldwide IOT Security Spending Will Reach $1.5 Billion in 2018 – Press Release”STAMFORD, Conn., March 21, 2018

[2] Cigniti Technologies - Global Leader in Independent Quality Engineering & Software Testing Services. “The Need for Testing the Internet of Things, – Blog 29th April, 2018.

[3] AFour Technologies - Software Product Engineering Services, “IoT Testing Services” – Article

[4] Subhasis- QE “Internet Of Things (IoT) Testing: Challenges, Tools And Testing Approach,”- November 10, 2019

[5] Alex Grizhnevich - Process automation and IoT consultant, ScienceSoft, “IOT architecture: building blocks and how they work,” – Blog Apr 1, 2018

[6] Anna Bryk - Market Research Specialist,”Internet of Things Security: Challenges and Best Practices,” – Article Jan 16, 2018

[7] Yakub Reddy Gurijala – Senior Technology Architect, “Performance Testing IOT” Article 2018

Reference URL

1. http://www.gartner.com/newsroom/id/3165317

2. https://www.cigniti.com/blog/the-need-for-testing-the-internet-of-things/

3. https://afourtech.com/iot-testing-services/

4. https://www.softwaretestinghelp.com/internet-of-things-iot-testing/

5. https://www.scnsoft.com/blog/iot-architecture-in-a-nutshell-and-how-it-works

6. https://www.apriorit.com/dev-blog/513-iot-security

7. https://www.infosys.com/IT-services/validation-solutions/features-opinions/Documents/performance-testing-iot.pdf

ISSN 2394 - 7780

CAN CYBER INSURANCE SAFEGUARD COMPANIES FROM DATA PRIVACY BREACHES?

Mihir Patel
Master of Science in Information and Technology from SVKM's Usha Pravin Gandhi of Arts, Commerce and Science, Affiliated with Mumbai University, Mumbai

ABSTRACT

With data privacy gaining a lot of traction in the recent few years, especially after the enactment of the General Data Protection Regulation (GDPR) and the proposed Personal Data Protection Bill (PDPB) – India, organizations are facing a daunting task to get compliant to the applicable regulations. The primary area of investigation was to analyze the current market of cyber insurance and to understand the level of protection that the insurances provide.

Keywords: Data Privacy, GDPR (General Data Protection Rule), Data Security, Cyber Insurance, Data Breach

I. INTRODUCTION

A. What is Data Privacy and what is the General Data Protection Regulation (GDPR) all about?

Data privacy can be defined as the various rights and obligations that an organization and a natural person have against the use, retention and disclosure of personal information. On the 25th of May 2018, the European Union passed a new regulation called the GDPR which revolutionized the way the world looks at privacy. Failure to adhere to its norms could lead to companies facing tough penalties which could be within the bracket of 20 Million or 4% of the organization's global annual revenue, whichever is higher. New rights have been introduced which include: Right to be forgotten, Right to portability, etc. It states that consent should be obtained by organizations to process personal data of data subjects. The consent obtained should be free and against an unambiguous privacy policy. Data breaches, if any, should be reported by the company to the supervisory authority within 72 hours. Wherever needed dedicated personnel called Data Privacy Officer should be appointed to take care of privacy practises of the organization. For projects where privacy risks are high, Privacy Impact Assessments will be required. Apart from the above, GDPR also instructs companies to embed "Privacy" into all processes of the organization by default. Apart from the regulatory fines, data subjects can also claim material and non-material losses.

B. What is a cyber incident?

"Cyber incident is a breach of a system's security policy in order to affect its integrity or availability and/or the unauthorized access or attempted access to the system." [1] The key threat agent here is to understand that breaches can be caused by internal as well as external agents. Over the course of this research we will determine how does a privacy breach get influenced by the threat agent.

C. What is the history of the cyber insurance market?

There are two types of insurances available – one for individuals and other one is for organizations and conglomerates. For the context of this research paper, we will assess cyber insurances pertaining only the organizations. In the late 1990s, the cyber insurance industry had just started to gain some traction. Incidentally, one of the first cyber insurances policies were said to be drafted in order to cover liabilities arising out of a third-party negligence. In the early 2000s, the cyber insurers observed certain other factors which needed to be addressed like rouge employees and regulatory penalties. However, by the late 2000s, insurers observed that each organization has a different appetite, they have different issues and also have different views and goals to achieve on Cyber Security. With the 2010s bringing in bigger breaches, insurers had to update their offerings as well. There are now more customized offers.

As of 2018, USD 4.3 billion is the amount at which the cyber insurance market was valued at. This value is only expected to grow at roughly 25.6% over the next few years. [2]. In 2018, it was reported that on an average the totoal cost of a breach was close to USD 3.86 million. This figure was higher than that of 2017 by 6.4% [3].

In all the above metrics, the key factor to note is that the amounts estimated for the breaches include all types of breaches ranging from identity theft to large scale hacks.

II. LITERATURE REVIEW

A. Challenges for the Insurers and Insured in Quantifying and Mitigating Cyber Risk [4]

The approach adopted by the author was to understand the cyber risk domain and current options for mitigating the same. It touches upon the cyber insurance industry from the perspective of both the insurers and the insured. The author very beautifully pointed out the factors that go on to determine how does an insurer determine whether a risk can be insured or not. The author primarily tried to understand all the possible risks. Based on his understanding, we learnt that there are various forms of coverages offered by insurers such as: Coverage for destruction of data, blackmail, theft, denial of service, error, failure to protect data and defamation. These coverages typically cover costs for regular audits, aftermath relationship management and investigative expenses under the Liability Coverage.

Also based on the criteria prepared by Berliner in 1982, insurers used the following techniques to determine insurability: How random is it for a loss to occur? What is the maximum possible loss? On every occurrence, what is the Average loss? Is there any information asymmetry? Is the premium amount easily insurable? Does the insurance coverage raise issues from a public and morals standpoint? Will the insurance have any legal restrictions?

The conclusion of this research was to determine the types of risks that the insurers have to determine. ‘Silent’ exposures need to be carefully managed. Based on the history of the cyber insurance industry, it can be seen that the industry has been a bit slow in its initial strides and going forward insurers need to have a better picture to define a better policy. This research does not however deep dive into what issues the insured face from a perspective of data privacy. With one of the stringiest laws and regulatory fines in place, data privacy laws should be given a more important look.

B. The Importance of Well-Prepared Cyber Risk Insurance and Open Source Intelligence (OSINT) [5]

The author states that the need for risk transfer is very essential in cyber security. Industries should undertake risk evaluations, implement cyber risk insurances and prepare management processes to help mitigate or prepare better for cyber incidents in future. The study was primary conducted from the perspective of cyber risk insurance’s importance.

The author first tries to understand the cyber risk domain. Systematically the second step involves understanding the approaches that can be taken to remove /reduce risks. One of the approaches suggested by the author involve making the risk measurable by providing it metrics. This can be achieved by performed a periodic risk assessment for an organization.

A ‘risk score’ should be assigned to each organization and results from previous assessments done should be compared to evaluated whether the organization can be insured or not and if yes, until what extent. The author also suggests the use of OSINT tool for preparing the risk score. Apart from these keys suggestions, the author also suggests other solutions such as the insurer should provide: a coverage for consultants who will help the impacted company during or after a breach, free of charge Intrusion Detection system (IDS) and firewall protection, guidance on measures to be taken before a breach, etc.

The above-mentioned solutions again similar to the previous case do not point out the measures that an organization would need to take from the perspective of newly introduced data privacy regulations. It also does not touch on factors that would not be typically covered under a cyber insurance or be excluded from an average cyber insurance policy. Hence, the reader may not have a wholesome view of all aspects of the insurance.

III. METHODOLOGY

The research will be conducted using the Qualitative Approach. Based on primary interactions with key Information Technology stakeholders in various conglomerates, two pain points regarding data privacy emerged. The following were the primary set of questions asked during interviews:

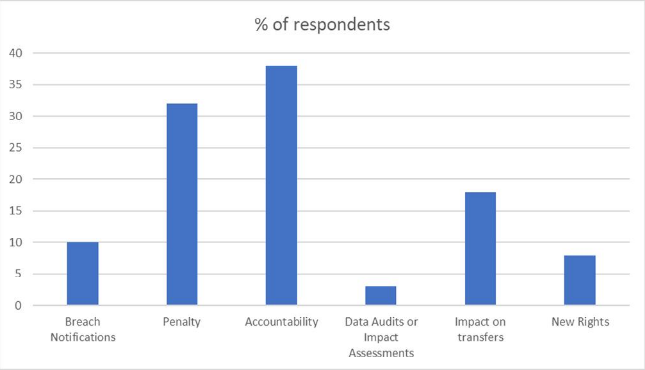
- Is your organization subject to any data privacy regulations?
- How well aware are you with the GDPR?
- Does your organization place privacy compliance as a top priority?
- Is there a process set in your organization for handling data privacy incidents?
- Does your organization have and allocated budget for data privacy or information security? If yes, please specify the bracket?
- Do you plan on organizing a dedicated team to take care of data privacy in your organization?
- How well aware are you with the Personal Data Protection Bill (PDPB) put forth for discussion?
- On a scale of 1-5, kindly rate the below key concerns of data privacy where 1 represents the parameter that concerns you the least and 5 which is your major concern in data privacy.

- Breach Notifications
- Penalty
- Accountability
- Data Audits or Impact Assessments
- Impact on transfers
- New Rights

The questionnaire above was designed to understand key pain points of the organization. Apart from this secondary data was analysed from various sources over the internet, interactions with cyber insurance companies, interactions with regulatory bodies.

IV. RESULT

As per my survey, IT professionals from few of India's biggest conglomerates pointed out the following:



Based on the above results, we can see that the biggest concern of data privacy is that is Accountability. Organizations are concerned about liabilities arising out of breaches and with multiple parties involved in a single data processing operation, this is a major concern. The second biggest concerns is the penalties arising out of the breach which is again linked back to who takes accountability and in-turn who pays the penalties for the breach.

Now based on my interactions with leading cyber insurance providers, I have learnt that a very huge chunk of costs involved in a data privacy breach are covered by insurance. For e.g. when Target's systems were breached, and a major chunk of their customer data was exposed, the total cost of the breach reached USD 252 Million. But about USD 90 Million was covered by cyber insurances. This is because the costs only covered the data breach and all that which was covered under the cyber liability insurance. This included: Forensic/investigative costs, restoration costs, business interruptions and hiring of PR firm to repair and contain the reputational damage of the company. So, Target still had to pay USD 160 Million but a sizeable portion was saved when the insurance cover kicked in.

The costs that can be recovered from insurances is based on the coverage which in turn is determined by the factors stated in II.A. There will always be a residual factor which will be difficult to cover. One of the two factors which has come as a key outcome of this research is the fact the personal claims i.e. claim for material and non-material damages as prescribed under the GDPR Article 82- Right to compensation, are generally not covered under this section since there are no caps on this amount.

V. CONCLUSION

Even after taking the best of the precautions, risk factor is never mitigated completely. No system / process can be 100 % safe in the ever-emerging cyber world. In order to allow organizations to breathe easier, insurances are the best bet.

We know that regulatory fines can under general conditions be up to 20 Million or 4% of the global annual revenue. However, with no caps on the personal claims (based on Article 82 of GDPR), there could very well arise occasions that could run an organization into bankruptcy. This can be countered with a customer protection liability policy that could probably insure each customer against a certain sum of amount depending on the likeliness of a breach and probable impact of the same on a customer. Customer with higher risks can be insured for a larger amount.

The second key factor that also was a derivation from my research was that GDPR is not insurable in many jurisdictions just within the EU. A clear picture of this is not available for review as of now, however it is very much a concern that needs further investigation. This also raise concerns that probably other upcoming data privacy regulations also may not be insurable under certain conditions. As of now, the provision of allowing companies to take insurances for GDPR is purely in the hands of a particular Member state. One Member state within the EU may accept while another may disallow an insurance cover for data privacy. However, as much as data privacy is not insurable in some jurisdictions, it can very likely be covered through other liability insurances. However, this should be planned carefully by the organization to ensure maximum coverage by policies.

VI. REFERNCES

[1] National Cyber Security Center (UK), "Cyber Incidents".

[2] Grand View Research Inc., "www.grandviewresearch.com," [Online]. Available: www.grandviewresearch.com.

[3] European Network and Information Security Agency, "Cyber Security Breaches Survey 2018," May 2018. [Online]. Available: https://www.enisa.europa.eu/news/member-states/cyber-security-breaches-survey-2018.

[4] M. Payne, "An Overview of the Cyber Insurance Industry: Challenges for Insurers and Insureds inQuantifying and Mitigating Cyber Risks," 2016.

[5] M. S. T. a. N. U. Baris CELIKTAK, "The Importance of Well-Prepared Cyber Risk Insurance and Open Source Intelligence (OSINT)," International Journal of Recent Scientific Research , vol. 9, no. S(I), pp.27101-27107, 2018.

ANATOMY OF VIRTUALIZATION

Tejas Prakash Tamboskar
Department of Science and Information Technology, Patkar Varde College, Mumbai

ABSTRACT

Virtualization is a science that lets in developing a virtual mannequin of laptop resources, such as hardware architecture, strolling system, storage, network, etc. With this a single computer can act like many machines working independently. The crucial issue in a virtualization is referred to as hypervisor, having extra-privileges, which makes it successful to play quintessential role of dealing the input of records and has many benefits involving the costing, the minimalism of performance and the ease of use. But in the different hand, virtualization makes it the ideal target for representing users focused on to attack the nearest on hand machine.

Keywords: Computing in cloud, safety in clouds, security in networks, Virtualization, Effects of virtualization

I. INTRODUCTION

Cloud computing is becoming very popular as virtualization supremacy, as virtualization offers dispensed computing with server cluster and increasing in the availability of web. The IT trade's effort on virtualization skill has become large in the previous few years. However, the concept has been around in a very slow way. Virtualization also offers a high level view of the technology and procedures that occurs in our day to day life, and number of reasons why virtualization. The world of IT is searching ahead for the offerings provided by means of cloud computing for that reason expansion of the cloud computing. Cloud computing is the effect of grid computing, utility computing and computerized computing.

Cloud is a parallel and disbursed computing device which consists a set of inter connected and virtualized workstations which gives one or greater unified computing resources primarily based on the requirements between carrier carriers and service consumers.

Cloud computing is on demand pay-as-use i.e billing is completed primarily based on the utilization of the client which downs the operational and capital cost. Users can get right of entry to purposes which are existing outdoor the working site which can get admission to far flung purposes through internet connection devices. By this, computer assets can be effectively used and consume much less computing electricity and resources are shared cooperatively.

The principal strength of cloud computing lies in the way facts is stored, how it is transmitted and accessed. A virtualized platform with administration competencies like availability, computerized load balancing and fault tolerance reduces infrastructure cost and upkeep fee

II. CLOUD COMPUTING COMPONENTS

Computing in cloud has users, data centers and disbursed servers as the components.

Clients: Users like computers, laptops, tablets computer systems cellular telephones or PDA's.

Data Centers: These are a series of servers where the application is hosted. Virtualization is done where a couple of situations of digital servers are created.

Distributed Server: Servers which stay none regionally which are geographically far.

The major goal of cloud computing is to provide computing power, storage and software "as a service". By the help of offerings in cloud computing presents usability, high-quality grained components. Services provide scalability, multi tenancy and device independence. There are three types of cloud services

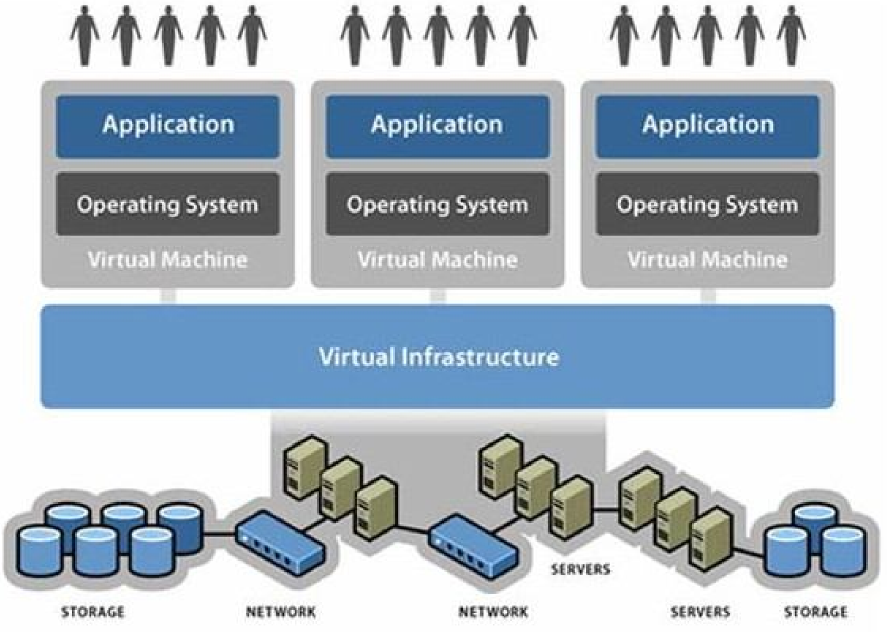
SaaS: Software as a Service is the model in which an application is hosted as a service to customers who access through internet. Users can access their application anywhere if they are connected to internet.

PaaS: Platform as a Service. This is another application delivery model which provides resources required to build application and services completely from internet without purchasing.

IaaS: Infrastructure as a service. This presents the required hardware so that customers can put something they required. IaaS allows renting of resources like server space, processing unit cycles, community equipment's, reminiscence and storage area

III. ANATOMY OF VIRTUALIZATION

Virtualization in computing is layout of no longer actual something such as hardware, software, platform or working gadget or a loading or a network. In a virtualized scenario IT creativity has to reap many modifications as the changes appear e in virtual ecosystem than in a physical environment. Virtualization technology makes cloud computing surroundings without difficulty to control the resources. It abstracts and isolates the underlying hardware, and networking resources in a single web hosting environment



Advantages of Virtualization in cloud computing is that Virtualization technological know-how makes cloud computing surroundings without difficulty to control the resources. It abstracts and isolates the underlying hardware, and networking resources in a single hosting environment

It will increase the protection of cloud computing by way of defending both the integrity on guest digital desktop and cloud factors virtualized machines can be scaled up or down on demand and can furnish reliability. It gives resource sharing, high utilization of pooled resources, fast provisioning, and workload isolation.

Hypervisor: A hypervisor is a software, hardware or a firmware that gives virtual partitioning competencies which runs directly on hardware. It is described as the virtual machine supervisor which permits more than one working structures to run on a system at a time supplying sources to every OS barring any interaction. Hypervisor controls all the guest systems. As the operating machine quantity will increase managing is hard these leads to protection issues. If a hacker gets manage over the hypervisor he can control the visitor systems with the aid of understanding the conduct of the machine which causes records processing damage. Advanced safety machine is to be developed to monitor the things to do of the visitor Virtual machine

IV. VIRTUAL MACHINE

If we choose to define digital a computing device simply as there are remote packing containers which shared hardware between them. These bins are carefully separated to each different and act like one of a kind physical computers, which can be linked through equal community or not.

“A virtual machine (VM) is an abstraction layer or the environment between hardware components and the end-user. Virtual machines have an ability to run any operating systems on them and in special cases it referred to as virtual hosts.

The interaction between the guest operating systems which are running in virtual machines and resources which are available for sharing between virtual machines, provided in two ways. One is by using the host operating system, or another, a piece of software which called as the hypervisor and acts like mini operating system, can run many virtual machines. Hypervisor also call as virtual machine monitors. They are able to share system hardware components such as CPUs, controllers, disk, memory, and I/O among virtual servers

V. VIRTUALIZATION CONCERNS

Virtualization enable single system to concurrently run multiple isolated virtual machines. If isolation of these Virtual Machines is not properly implemented, intruders may perform unauthorized communication with other

Virtual Machines in the system. Assaulter can use Trojans, malwares etc to tamper the functionality of guest OS, they can also use viruses and worms to exploit the guest OS in Virtual Machines. Attackers can even compromise the privileged host virtual machine Dom0 to tamper boot process of guest Virtual Machines or access all guest Virtual Machines including their memory disk space and network traffic, the attacker can create multiple virtual machines to consume all the system resources simply by controlling Dom 0. The saved state of guest VM appears as a disk file in plain text to Dom0

VI. VIRTUAL MACHINE ESCAPE ATTACK

Virtual machine escape is an exploit in which the attacker runs code on a VM that allows an operating system running within it to break out and interact directly with the hypervisor. Such an exploit could give the attacker access to the host operating system and all other virtual machines running on that host. Although there have been no incidents reported in the wild, VM escape is considered to be the most serious threat to virtual machine security.

Virtual machines are designed to run in self-contained, isolated environments in the host. Each VM should be, in effect, a separate system, isolated from the host operating system and any other VMs running on the same machine. The hypervisors an intermediary between the host operating system and virtual machines. It controls the host processor and allocates resources as required to each guest operating system.

If the attacker can compromise the virtual machines, they will likely have control of all of the guests, since the guests are merely subsets of the program itself. Also, most virtual machines run with very high privileges on the host because a virtual machine needs comprehensive access to the host's hardware so it can then map the real hardware into virtualized hardware for the guests. Thus, compromising the virtual machine means not only that the guests are goners, but the host is also likely lost.

To limit vulnerability to VM escape

Keep virtual computing device software patched. Install only the resource-sharing elements that you without a doubt need. Keep software installations to a minimum because every application brings its personal vulnerabilities

VII. CONCLUSION

Virtualization itself is no longer inherently unsecure, it is a technological know-how that has new vulnerabilities and requires restructuring of guide safety processes. One of the largest challenges is to preserve and invulnerable all of the VMs, on account that many situations and configurations can be swiftly created. The contents of every visitor OS is a digital disk, stored as a file. If this file is accessed, copied, or modified on the host by using an unauthorized party, then the privateness and integrity of the VM is compromised. Likewise, if an attacker accesses the host and immediately modifies the hypervisor, then he or she will be able to run arbitrary code, however the hypervisor has extra layer of abstraction from physical hardware and similarly restricts malicious tries to manage the laptop from the hardware. This abstraction encapsulates malicious attacks and permits external monitoring for malicious assaults on a VM. Since the hypervisor video display units each VM, it can record the states and allow the VM to return to a previous state, which has many backup and malware removal advantages. The hypervisor ought to strictly control conversation between VMs and restriction useful resource consumption of every VM to a finite certain to stop DoS attacks. All recognized vulnerabilities of VMs can be prevented, but it is definitely quintessential to impervious the host and each guest OS in order to create a invulnerable digital environment

VIII. REFERENCES

1. Hoffman, P., Scarfone, K., Souppaya, M.: Guide to security for full virtualization technologies. National Institute of Standards and Technology (NIST) (2011) 800– 125.

2. J.E. Smith and R. Nair, Virtual Machines. San Francisco, CA: Elsevier, 2005, pp. 369–443.

3. L.McLaughlin,“How to Find and Fix 10 Real Security Threats on Your Virtual Servers,” CIO, 2007.

4. J.Brodkin, “Virtual server sprawl highlights security concerns,” Network World, 2008.

5. T. Garfinkel, M. Rosenblum, “When Virtual is Harder than Real: Security Challenges in Virtual Machine Based Computing Environments,” USENIX Association, 2005

A BENCHMARK STUDY OF CASHLESS DIGITAL APPLICATIONS: PAYTM VS PHONEPE

Mitali Kumbhani[1] and Dr. Neelam Naik[2]
Student[1], MSc IT- Part I, Usha Pravin Gandhi College of Arts, Science and Commerce
Assistant Professor[2], Usha Pravin Gandhi College of Arts, Science and Commerce

ABSTRACT

Cashless transactions methods has paved a revolutionary change in 21st century. Online transaction processing typically facilitates and manage transaction-oriented applications. The physical currency is replaced by a number of methods that are powered by digital information technology and are capable to transfer money online through different applications. "Faceless, Paperless, Cashless" is one of professed role of Digital India. Digital India becomes a new program by the Government of India, which also promote cashless economy. The Main purpose for this research is to check out the most preferable application between Paytm and PhonePe, on the basis of their features and benefits. This paper focuses on comparison between Paytm and PhonePe to know the best out of two methods. Researchers found that people like Paytm in terms of its adapted features and its variety of reasons but there were few people who like PhonePe just because of its few features.

Keywords: Professed, cashless digital application, transaction.

INTRODUCTION

Digital transactions bring in better transparency, scalability and accountability. In conjunction with demonetization, the government has announced waivers on convenience charges, surcharge and service charge on digital payments by government departments and organizations to promote digital and payments. [1] Digital payments in India are projected to reach $1 trillion by 2023, up from about $200 billion currently, said a Credit Suisse Group cited by an Economic Times report. Its majority business is from peer-to-peer(P2P) transactions, and it can see an increase in business-to-business (B2B) over the years, which currently stands between 10-20 percent.

As of now the two cashless digital applications had already captured majority of shares in India. Not just that many other cashless digital services were launched including Amazon Pay, BHIM, Freecharge, Google Pay, Mobikwik, and many more in India. Although digital payment is a fairly recent trend in India, it has seen exponential growth due to a favorable regulatory environment, deeper smartphone penetration and growing internet access. So through this research project will come to know which cashless digital application is best among the two that is Paytm and PhonePe.

As, today Paytm and PhonePe are the applications used by most of the people. This two are the trending applications for cashless transactions that is Digital payments. While other digital payment companies maintain a minority share, PhonePe competes with the Paytm's first mover advantage, though it continues to diversify its array of services to remain the leader. So, the finding of the study that is research is that among these two applications which is most used by the people for cashless transactions. The research focus on which is the best among Paytm and PhonePe. Which is more trending among the people today

LITERATURE REVIEW

In research paper, [2] author have concluded, with limited cash in hand and an indefinite crunch in sight, most people are rushing to cashless transactions. Digital transactions bring in better transparency, scalability and accountability. The new move will compel more merchants to accept digital money. Cash may no longer be king. While you wait for the serpentine queues at ATMs to peter out and currency notes of Rs 100 denomination to become easily accessible again, the adoption of digital payment solutions is picking up at a furious pace. Everyone from the neighborhood vegetable vendor to the chai and bhelpuri-wala is embracing digital payment solutions to tide over the cash crunch. ET Wealth conducted an online survey to find out the level of adoption of digital payment solutions and user habits. The findings reveal that while people are getting comfortable with cashless payments, some mindset issues are holding back many from embracing the newer platforms. The findings also suggest that the usage habits of those who have taken to cashless modes could be exposing them to security threats.

[3] It was predicted that in forthcoming years there will be a sharp and unexpected rise in the number of digital payment transactions in India. The digital payments sector that is, cashless payments transactions will touch 1 trillion in succeeding 5 years. UPI and Aadhaar will help increase digital penetration, according to sources. UPI

fund transfer money can be sent using a virtual address without providing IFSC code or account number. The usage of UPI based transactions continue to grow exponentially touching Rs.542 billion value in Aug'18.

[4] Paytm has processed 1.2 billion merchant transactions in the first three months of the current financial year, the digital payments provider claimed, even as the battle for a larger share of the overall retail digital payments pie intensifies. On the other hand, Paytm's closest rival, the Bengaluru-based PhonePe, claimed it had recorded 90 million merchant transactions through the app in August alone. Paytm processes transactions across multiple modes like mobile wallets, UPI as well as its payment gateway. PhonePe is primarily into UPI payments. The Noida-based Paytm said these transactions have been processed at more than 14 million retail stores that accept Paytm payments across the country. PhonePe said it processed 90 million transactions at its 6.5 million merchant base in August alone, across 210 locations in the country. According to numbers published by the National Payments Corporation of India, the total UPI transactions recorded or August stood at 918 million. ET wrote in its September 5 edition that the number of payments recorded on UPI by Paytm was 157 million, while it was 340 million for PhonePe.

[5] Demonetization has triggered more usage of e-payment among public which increases the usage of cashless transaction. Transferring money through cashless modes would basically demand the usage of plastic money that is digital applications. This indicates a movement towards a cashless economy. The government initiative on Digital India to boost the adoption of digital payment system among the individuals.

RESEARCH METHODOLOGY

H0: There is no significant difference between consumers of Paytm vs. PhonePe.

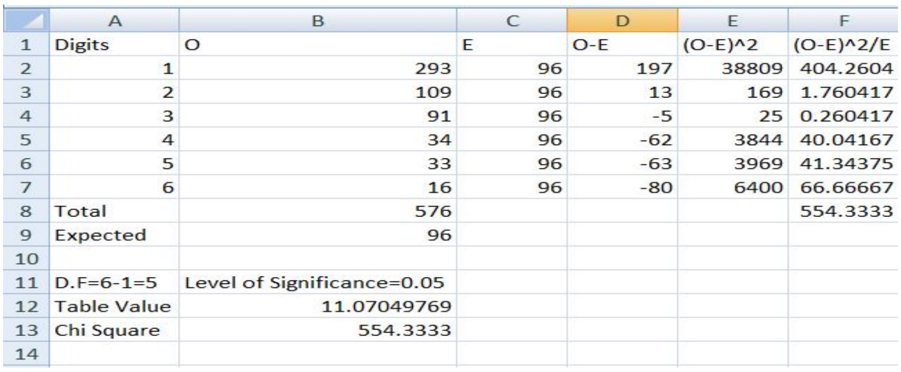
H1: There is a significant difference between Paytm vs. PhonePe that is Paytm is better than PhonePe.

In this section, the methods used to prove whether our NULL hypothesis or Alternate Hypothesis is accepted or rejected. For this we collected quantitative data and qualitative data through surveys that is Google form.

To prove Hypothesis, researcher kept a sampling frame of 60 people. For this, we use non-parametric test because the features in data are of nominal type.

The samples are collected from Maharashtra of different age-groups and gender.

On the basis of data collected Chi-Square test is used for testing NULL hypothesis. The reason behind using Chi-Square test was that for it we required more than 50 samples and also it's a non-parametric test in we used goodness of fit method to prove our hypothesis. The other test allows you to say either "We can reject the null hypothesis of equal means at the 0.05 level". A Chi-Square test allows you to say either "we can reject the null hypothesis of no relationship at the 0.05 level" or "we have insufficient evidence to reject the null at the 0.05 level".



	A	B	C	D	E	F
1	Digits	O	E	O-E	(O-E)^2	(O-E)^2/E
2	1	293	96	197	38809	404.2604
3	2	109	96	13	169	1.760417
4	3	91	96	-5	25	0.260417
5	4	34	96	-62	3844	40.04167
6	5	33	96	-63	3969	41.34375
7	6	16	96	-80	6400	66.66667
8	Total	576				554.3333
9	Expected	96				
10						
11	D.F=6-1=5	Level of Significance=0.05				
12	Table Value	11.07049769				
13	Chi Square	554.3333				
14						

Figure 1: Results of Chi square test on data collected

From above data, our Chi-Square value=554.3333.

Tabulated value of χ^2 **test** for 5 Degree of Freedom at 5% level of significance is 11.0704. So our calculated value of χ^2 **test** is 554.3333, it is highly significant and Null Hypothesis is rejected at 5% level of significant.

Hence we can conclude that **H1**(Alternate Hypothesis is accepted and its means there is significance difference Paytm vs. PhonePe that is Paytm is better than PhonePe.

The questionnaire selected for the survey covers the following aspects of cashless digital applications referred by the customers.

1. The frequent usage of online transaction applications is measured.

2. Most of the people use Paytm rather than PhonePe. Paytm was used by most of the respondent and has brand loyalty.

3. According to the users, Paytm provides better efficiency than PhonePe.

4. Other than that, Paytm provides many features to the customers for ease of use, gives more security, and it provides multipurpose usage.

5. At present majority of Paytm users are very satisfied and they believed in recommending it to other and spread positive word of mouth about Paytm.

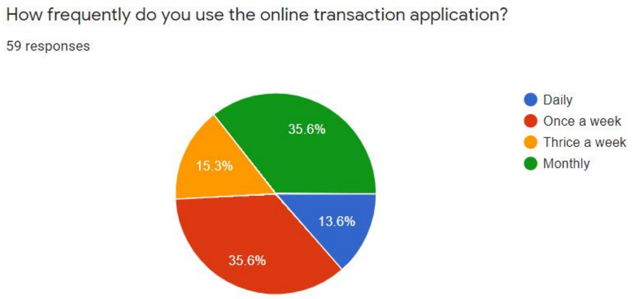


Figure 2:Choices of people using online transaction application.

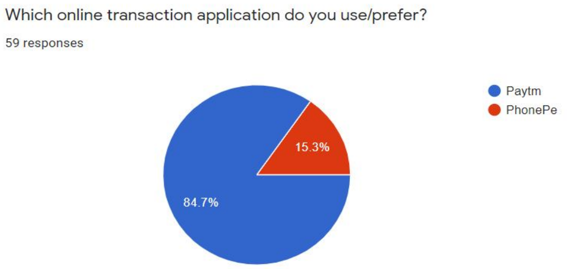


Figure 3:Results of most preferred application.

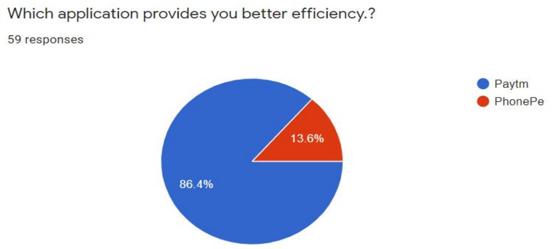


Figure 4:Application providing better efficiency.

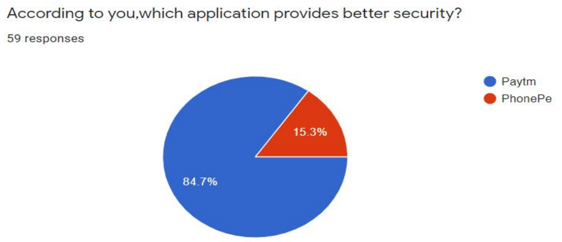


Figure 5:Application security.

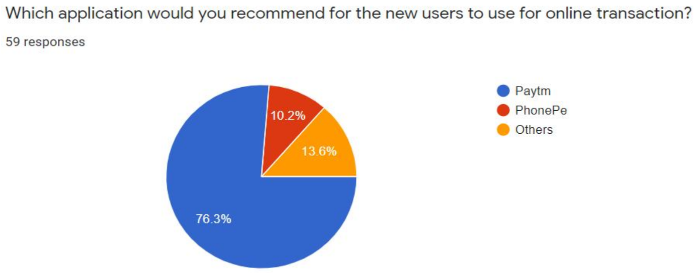


Figure 6:Application recommendation.

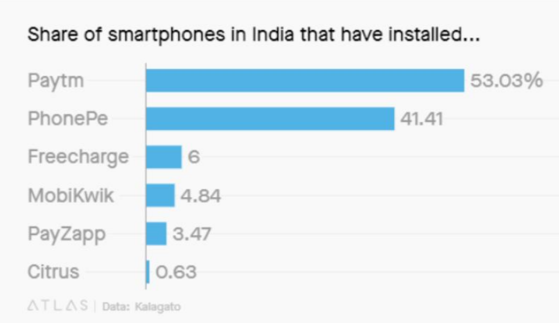


Figure 7:Comparison between Paytm and PhonePe.

CONCLUSION

In this paper researcher presented the significance difference of Paytm and PhonePe that both are giving the best features to tackle their competitors and attract more customers but still difference in availability in new updated available content, features makes Paytm stronger than PhonePe. But PhonePe's services that is UTP transactions, E-Wallet, Zero Banking Charges, etc. is constantly adding fuel in increasing its market share and value resulting in increasing number of customers giving tougher competition to Paytm in being cashless digital applications. The key feature of Paytm which over powers PhonePe is its KYC security in transaction. Paytm has shown a tremendous increase and growth in the E-payment sector, Market share of Paytm was maximum which is followed by PhonePe and others. As competition is growing Paytm need to be updated and provide more complementary services and features to keep his position strong in market. We also came to know that there are many people who still like other applications like BHIM, Google Pay, etc. So from the above research study, the most important conclusion is that Paytm is most widely used App with market share of 74%.

REFERENCES

[1] Shraddha Sharma, " As Paytm and PhonePe rule digital payments roost; Google Pay rapidly gaining market share",MoneyControl,2019.

[2] Ramya N.D SivasakthiM Nandhini, "Cashless transaction: Modes, advantages and disadvantages", IJAR 2017.

[3] C. Dhanalakshmi," A conceptual study on cashless economy: Digital India", IJCMR 2018.

[4] Pratik Bhakta," Retail e-payments: Battle between Paytm vs PhonePe intensifies", Ettech 2019.

[5] P. Sarika, S. Vasantha." Impact of Mobile Wallets on Cashless Transaction", (IJRTE)l 2019.

SENTIMENT ANALYSIS ON ABROGATION OF ARTICLE 370

Tanoop Aravindakshan[1] and Dr. Neelam Naik[2]
Student[1] and Assistant Professor[2], Usha Pravin Gandhi College

ABSTRACT

This document will help you to know the sentimental analysis of people regarding the decision made by the current government on the abrogation of article 370

Keywords: Sentimental analysis; Textblob; government; people

I. INTRODUCTION

“Will Abrogation of article 370 bring peace in Kashmir?” this was the question that made the researcher think on this topic. The researcher wanted to know what people are thinking of this decision taken by the current ruling government. So the researcher took this topic as my research topic that is a sentimental analysis on abrogation of article 370

What is Article 370?

In India, Article 370 of the Indian Constitution gave special status to the state of Jammu and Kashmir in the northern part of the Indian subcontinent administered by India as a state from 1954 to 31 October 2019 and part of the larger region of Kashmir, which has been the focus of the main conflict between India, China and Pakistan since 1947 giving it the right to have a separate state flag, constitution and control over the state's internal administration.

The article was drafted in Part 21: Temporary, Transitional and Special Provisions of the Constitution. Following its establishment, the Constituent Assembly of Jammu and Kashmir was empowered to recommend the articles of the Indian constitution that were to be applied to the state or to repeal Article 370 altogether. The Presidential Order of 1954 was issued after consultation with the State's Constituent Assembly, specifying the articles of the Indian constitution applicable to the State. Since the Constituent Assembly dissolved without recommending the repeal of Article 370, it considered that the article had become a permanent feature of the Indian Constitution.

This Article 370 and Article 35A established that residents of the Jammu and Kashmir state live under a different set of laws, including those relating to fundamental rights, property ownership, and citizenship in accordance with residents of other Indian states. As a consequence of this rule, in Jammu & Kashmir, Indian nationals who are from other states could not purchase land or property.

On August 2019, the Government of India issued a constitutional order that superseded the 1954 order and made all the provisions of the Indian constitution applicable to other states applicable to Jammu and Kashmir on the basis of a 2/3 majority resolution passed in both houses of the Parliament of India. After the resolutions passed in both houses of parliament on 6 August, he issued a further order declaring all the clauses of Article 370 to be inoperative, except for clause 1.

In addition, parliament passed the the Jammu and Kashmir Reorganization Act, enacting the division of the state of Jammu and Kashmir into two union territories to be called Union Territory of Ladakh and Union Territory of Jammu and Kashmir. On 31 October 2019 the reorganization took place.

After Abrogation of article 370

Advantages

J & K will be like any other Indian state or union territory, J&K residents will have single citizenship of India, Union Territory assembly tenure:5 years, Indian national flag prevails, Centre responsible for administrative, local regulations also, Kashmiris won’t need a permanent resident certificate, Any Indian can settle in Kashmir and buy property

Disadvantages

Instilled insecurity in the locals as they have to give away their dual citizenship, Kashmiri Muslims feel it threatens the state’s unity and integrity, Add on to the political vulnerability and instability in the Valley, Hampers the delicate relationship with Pakistan. It is like a nail in the coffin, Implants the seeds of insecurity in certain sections of citizens, The implementation of the abrogation of Article 370 is a threat to democracy. It is an attempt to polarize and appease the Hindu population in the valley, The safety of Kashmiri girls is questioned. Certain Hindu fascists have threatened to marry girls of the region. This is outrightly sexist.

What is sentimental Analysis in one sentence?
Machine learning and natural language processing is considered to be a category of sentmental analysis. It is used to recognize, extricate, or portray opinions from different content structures, including articles, audits and news and categorizes them as negative, neutral and positive.[1] conference

II. PROBLEM IDENTIFICATION

The main motive for making this research paper is to know what people think about the abrogation of article 370. There were a lot of controversies regarding this topic. Kashmir is a place where there is always a tense situation. There were many problems in Kashmir such as a conflicts between Pakistan and India for Kashmir, There were many terror groups forming in Kashmir, many people did not want to be a part of India, Kashmiris were having their own flag till 2011.

The actions took by the government to carry out this decision lead to a chaotic situation in Kashmir. So the researcher is using the sentimental analysis method to know how supportive are the people from India with this decision sentimental analysis will take the inputs from the people and find how supportive was this decision.

In this paper, the researcher has attempted to know the view of the people for abrogation for article 370 and performed a sentimental analysis on the text inputs taken from the people.

III. LITERATURE REVIEW

This part of the paper will be used to explain the related study of sentimental analysis on different domains such as tourism, Stock, Indian general election.

In this paper [2] authors García, A., Gaines, S., & Linaza, have studied the During the last years sentimental analysis has been extensively investigated for the English language. Two main group can be made with the currently existiong approaches: methods based on the Natural Language Processing (NLP) techniques and combination of lexical resources. This paper introduces Sentiment Review of user reviews in Spanish for the hospitality and food and drink markets using lexical repositories. Based on the positive and negative words a global score will be calculated which appear in the review and using the mentioned lexicon database. [3] Severyn, A., & Moschitti have studied on the results of their approach and the official test sets will be compared on the systems participating in the challenge and Say that their model could be placed in the first two places in both the phrase-level subtask A (between 11 teams) and the message-level subtask B (between 40 teams).

In this paper [4] Mittal, A., & Goel, A. have studied to find the correlation between " market sentiment" and "public sentiment" with the help of machine learning principles and sentiment analysis. Twitter data is used to predict public mood and to use the predicted mood, and the stock market movements will be predicted by the. DJIA days before. Using Self Organizing Fuzzy Neural Networks (SOFNN) on Twitter feeds and DJIA values from June 2009 to December 2009, their results were presented using a new method of financial data cross-validation and precision of 75.56 percent.In this paper [5] Singhal, K., Agrawal, B., & Mittal, N has studied to understand public opinion and trends with the help of Political analysis using social media is getting a lot of researchers ' attention, particularly during election time. During the Indian general election, we crawled the political tweets, and further tested our possible solution to the election results.

IV. PROPOSED RESEARCH WORK

In this section, the researcher has taken 73 samples from all over Mumbai with no age, gender or religious discrimination

In this research work, a google form to get data there are 8 questions that have options of Likert type scale. this scale contains strongly agree, agree, disagree, strongly disagree. The researcher has divided these questions into two which is one that supports the decision one which is not supporting the decision. The questions that support the decision are 1)" With this decision lead to cut down the terror groups forming in Kashmir?" this question is asked because there are many terror groups formed in Kashmir such as Harkat-ul-Jihad al-Islami, Lashkar-e-Taiba, Jaish-e-Mohammed, Hizbul Mujahideen after the decision India has more control over Kashmir will this reduce the terror impact over Kashmir. 2) "Will this decision lead to much worse relationship with our neighboring countries?" after this decision was taken Pakistan filed a case against India to UN but the case was won by India and presidents and ministers of other countries supported India but will Pakistan be quiet with this decision or will declare another war with India. 3) "Is it not mandatory to constitute a Delimitation Commission to de-limit the Assembly and Lok Sabha constituencies after every Census?" Delimitation Commission to Redraw Jammu and Kashmir Assembly Constituencies Will be Set Up by Central Government after the abrogation of Article 370 4)" Is it not a fact that Article 370 has created a republic within the Indian Republic and created an impression across the world that Jammu & Kashmir is a disputed issue that is still to be settled?"

Before the abrogation of article 370, India used to have less control over Kashmir which lead to many chaotic situations in Kashmir.

. The questions that do not support the decision are “1) Is this a move from the BJP government to make a Hindu Rashtra?” the BJP has always been accused of making India a “Hindu Rashtra”.2)” Wasn’t there a need to hear peoples voice before passing the decision? “ The people of Kashmir were not asked before taking the decision of abrogation of article 370)” Will there be fewer riots in Kashmir after the abrogation of article 370?” Some amount of Kashmiris did not want to be a part of India, will this situation create more riots and fights in Kashmir 4)” Did the government had other ways to deal with Article 370 instead of abrogation of article 370?” Instead of abrogating article 370 government could have reframed the Article was one of the views from a large number of people. These questions will be given the option of a Likert type. Which will find the sentimental analysis from these questions.

Sentiment analysis is contextual text mining that identifies and extracts subjective information in the source material and helps a business understand its brand, product or service social sentiment while monitoring online conversations. Analysis of social media streams, however, is usually limited to only basic analysis of feelings and count-based metrics.

Text messages will be used as data for sentimental analysis. The sentimental analysis was done with the help of jupyter in anaconda using the python programming language. In this code, the TextBlob class is used to get the sentiment function which shows the polarity and the subjectivity

The process of determining the writer's attitude or emotion is basically referred to as Sentiment Analysis, i.e. wh ether it is negative, neutral or positive.

Textblob's sentiment function returns two properties-polarity and subjectivity.

Polarity is a float value with a range of[-1,1] where -1 is negatve and 1 is positive. Subjective sentences generally refer to emotion, personal opinion, judgment while objective is factual information. Subjectivity is likewise a float value with a set of [0,1].

V. RESULT AND DISCUSSIONS

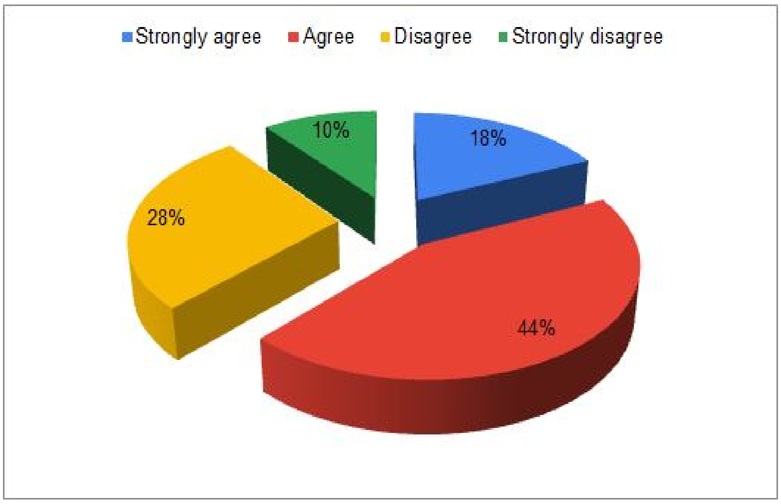


Fig no 1a: Reaction for questions asked against the decision

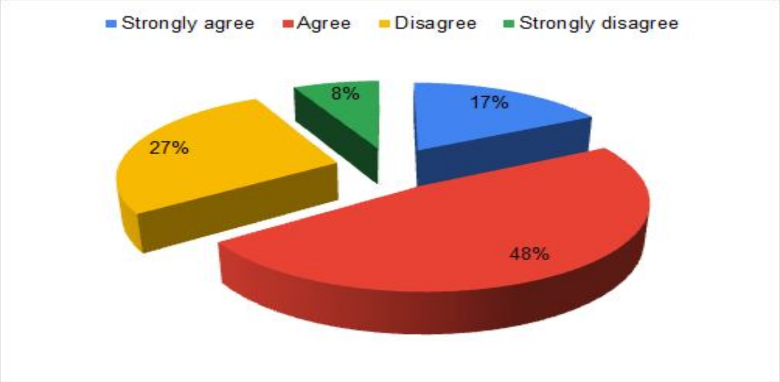


Fig no 1b: Reaction for questions which support the decision

These were questions asked and the results for those questions are been shown with the help of pie charts to conclude from these pie charts we have seen that maximum people are agreeing with the statements and minimum people are with the strongly disagree.

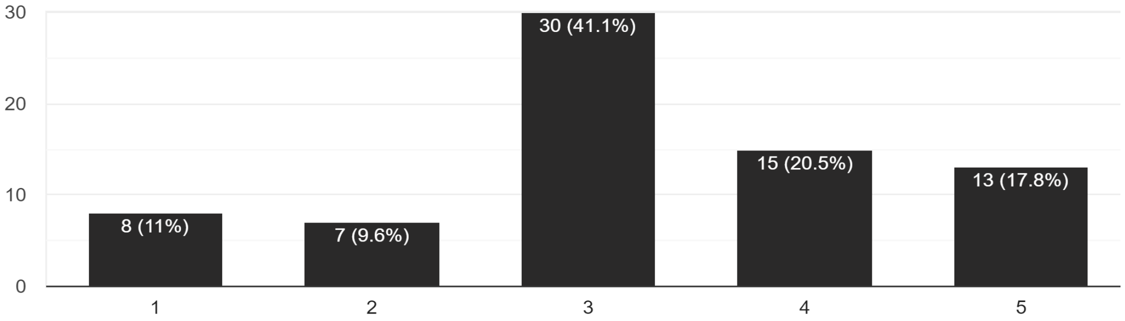


Fig no 2: Marks that are given by people from this bar chart, we can observe that the the highest mark given to this decision is 3 marks that are considered as neutral

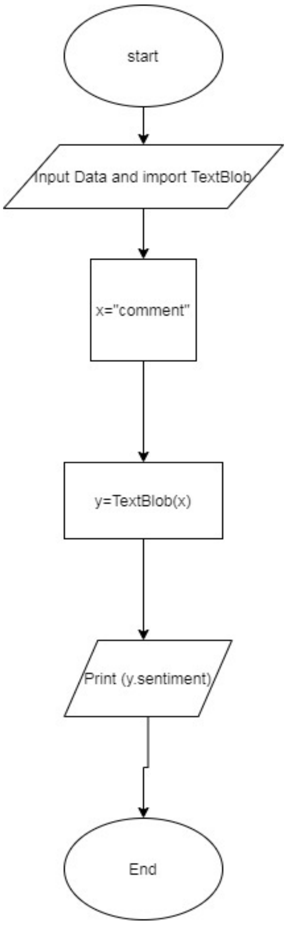


Fig no 3:Flow chart of the Code

```
ANS:Sentiment(polarity=0.182144490729396
4, subjectivity=0.5609840997576847)
```

The researcher has used python programming language to do the sentimental analysis with the help of anaconda

Step 1: import the textblob from the textblob

Step 2: assign a comment to a variable

Step 3: use the variable textblob function to get the output

The polarity value that we got from analyzing all the text with the help of textblob is 0.182 which is a positive response by calculating the sentiment from all the text taken from the people

VI. CONCLUSION

From the sentimental analysis done, the researcher concludes that the 73 samples got from the google form the decision taken by the government for abrogating article 370 has both pros and cons. The conclusion will vary as per the samples taken. The input was taken only from Mumbai and not from outside Mumbai. So this sample only talks about the people who are not facing the issue but has knowledge of the situation that has occurred.

As per the polarity and subjectivity analysis, the major opinion of the results is that they are supportive (i.e polarity of 0.182) of the decision taken by the government of abrogation of article 370 according to them every decision has their pros and cons, India had no control over the boundaries of Kashmir as well as inside Kashmir this decision of abrogation of article 370 will help Indian government to do so. The mentality of the people can be understood with the help of this analysis. These analyses will be useful for the government to know whether decisions are accepted by the people or not which will help them to know their winning chances in their next election.

To enhance this study we can use social media apps such as Twitter and Facebook to get much better data. The method that is used by the researcher to do sentimental analysis is rule-based data, but there are much better methods to do sentimental analysis which will provide more efficient results.

VII. REFERENCE

[1]. Prediction of Indian election using sentiment analysis on Hindi Twitter. Sharma, P., & Moh, T. S. (2016, December).

[2]. A lexicon based sentiment analysis retrieval system for tourism domain. García, A., Gaines, S., & Linaza, M. T. (2012).

[3]. Twitter sentiment analysis with deep convolutional neural networks. Severyn, A., & Moschitti, A. (2015, August).

[4]. Stock prediction using twitter sentiment . Mittal, A., & Goel, A. (2012).

[5]. Modeling Indian general elections: sentiment analysis of political Twitter data. Singhal, K., Agrawal, B., & Mittal, N. (2015).

ISSN 2394 - 7780

DATA SECURITY IN CLOUD COMPUTING USING ENCRYPTION

Sunita Vinod Gupta
Information Technology Department, Usha Pravin Gandhi College of Arts, Science and Commerce, Mumbai,

ABSTRACT

Cloud computing is new way of computing where computing resources and services are available online and are accessed by users via internet. This technology allows users to store huge amount of data in cloud environment and access those data from anywhere using internet. But this facility also leads to security issues such as confidentiality, privacy, data security. To provide security against these issues cryptography can be used. In this paper, various cryptographic algorithms used by authors are reviewed. Also, various cloud service providers and encryption facilities provided by them are studied.

Keywords: cloud computing; confidentiality; data security; cryptography; cloud service provider.

INTRODUCTION

Cloud computing is the new and fastest growing technology that delivers computing resources over internet. This technology allows consumers to access IT services and resources based on their needs regardless of where those services and resources are hosted by cloud providers and pay according to usage. As per official NIST definition, "cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [1]. This technology can offer not only services but also storage, networking, computing capabilities, IT infrastructure to run applications. These services are available to consumers through internet and offered on pay-per-use basis from cloud providers. Some of the important benefits of cloud computing are cost efficiency, increased storage capacity, backup and recovery, continuous resource availability and location independence [2]. One of the advantages mentioned here is increased storage capacity. This is offered to consumers on per-use-basis with no worry about maintenance. Hence consumers move their data on to the cloud. As users' data are now on cloud, the security of the sensitive data is a major concern for consumers [3]. Thus there is need to store information in secured way on these cloud servers. Main security issues in cloud environment are confidentiality and integrity of data [4]. To address these issues, solution is to use cryptography. Cryptography is an art of writing secrets. It consists of two parts: encryption and decryption. Encryption process converts readable data (plain text) into non readable data (cipher text). Decryption is the reverse process of encryption. In this paper, various cryptographic algorithms used by several authors are studied.

Related Work

Various cryptographic algorithms such as AES (Advanced Encryption Standard), RSA (Rivest–Shamir–Adleman), DNA, Homomorphic, blowfish are already studied by many authors. Encryption process will make data unreadable for attacker thereby providing security to data against various attacks. There are two places where encryption process can be carried out: server side and client side. Server side encryption will encrypt data at server end before storing it on storage servers and decryption will be done just before providing data to consumers. Client side encryption is the responsibility of client. Here consumer will encrypt its data and uploads encrypted data on the cloud making it unreadable for cloud provider also. In server side encryption, key management can be done completely at server side or client can also be involved but in client side encryption, key management is done at client side only. Cloud server has no role to play in key management at client side encryption. Table 1 lists various algorithms used by authors.

Table 1: Encryption Algorithms used for data security in cloud computing environment

Algorithm/s used	Encryption done by	Key Length	Key Management	Observations
Enhanced RSA with varying key size [3]	Client	Variable size keys e.g. 128, 256, 512, 1024	Client	Enhanced RSA algorithm is compared with High Speed and Secure RSA algorithm. Encryption and decryption time is much lesser in Enhanced RSA when compared with High Speed and Secure RSA. Computation is complex thereby increasing the security.
Blowfish [4]	Client	Variable length key	Client	File is compressed and then encrypted. Compression process decrease space requirement at cloud environment. Symmetric key is stored at client side

		(32 bits to 448 bits)		in key management module which can be accessed only by authentic user. If key management module is compromised, then key will be available to attacker who can then decrypt the files.
AES along with digital signature and Diffie Hellman Key Exchange [5]	Server	Not available	Server	Two servers are maintained in cloud environment one for key management and encryption and second for storing encrypted data. Key exchange and client authentication are done using Diffie Hellman key exchange and digital signatures respectively. For encryption AES algorithm is used.
RSA and AES [6]	Client (RSA) and server (AES) both	AES key-not available RSA 2048 bits	Server	Client and server both will do encryption of file. Sender will encrypt file using RSA algorithm and store it cloud storage. Server will also encrypt the already encrypted file using AES algorithm. Providing double encryption increases the security of file on cloud environment.
AES [7]	Client	128 bits	Client	AES-128 is used for encryption at client side. Trusting security provided by cloud providers may not be sufficient leading to loss of confidentiality and privacy. Importance of client side encryption is highlighted here. Keys are stored in servers available at client environment and security of this server is of paramount concern.
BDNA (Binary DNA) [8]	Client	7 bit key	Client	Client side encryption is used. For encryption process two encoding tables, encryption key and one random number are used. Cloud environment will maintain encrypted data along with list of valid users of this data thereby providing authorized access to data.
DNA [9]	Client	1024 bit DNA based secret key	Server generates asymmetric keys and data owner generates DNA based symmetric key	Symmetric and asymmetric encryption algorithms are used for securing data on cloud. 1024-bit DNA based secret key is used by data owner to encrypt data and asymmetric encryption algorithm is used for communication between data owner, user and cloud service provider. The proposed encryption scheme is secured against many attacks such as malware injection, side channel attack, phishing attack, insider attack, denial of service attack.
Homomorhic [10]	Client	Not available	Server	Client side encryption is done. Computations are performed on encrypted data and result is provided to client who can decrypt it and can get the original result. Not all operations/computations are supported. Only addition, subtraction and check balance operations are supported in this paper.

Encryption Facilities provided by Various Cloud providers

In this section, few cloud providers are studied to get information about which encryption algorithm/s they are using. Cloud providers provide two kinds of encryption facilities: first is for data which is going to be stored on cloud storage and second is for when transmission of data occurs from user to cloud provider to avoid data breach during transit. These cloud providers are listed in Table 2. Algorithms, hashing techniques and protocols used by these providers are Advanced Encryption Standard (AES), Rivest Cipher 4(RC4), Message-Digest algorithm 5(MD5), Secure Hash Algorithm (SHA), Transport Layer Security (TLS), Perfect Forward Secrecy (PFS), Internet Protocol Security (IPSec), Secure Sockets Layer(SSL), Application Layer Transport Security (ALTS), Hyper Text Transfer Protocol Secure (HTTPS), SSH File Transfer Protocol (SFTP). Few providers only provide server side encryption where as few offer both option. From the listed providers, only Google Cloud Platform offers server side encryption by default. In others, options of encryption need to be selected by consumers if they are interested in getting encryption facility otherwise data will be in plain text format at cloud environment.

Table-2: Various cloud providers providing encryption facilities

Provider	Encryption of data at rest	Encryption of data in transit	Place of Encryption
Amazon Web Services [11], [12]	Yes AES 256	Yes TLS and AES 256	Client side and/or Server side
Microsoft Azure [13]	Yes AES 256	Yes TLS and PFS and RSA 2,048	Client side and/or Server side
Google Cloud Platform [14], [15], [16]	Yes AES 256 or AES 128	Yes IPSec/ TLS/ Managed SSL/ ALTS and AES 256 or AES 128	Client side and Server side (server side encryption is done by default irrespective of client side encryption)
IBM cloud [17]	Yes IBM's SecureSlice which uses AES or RC4 along with hashing 1. RC4 128 with MD5 128 Hash 2. AES 128 with MD5 128 3. AES 256 with SHA 256	Yes TLS	Server Side
Salesforce [18]	Yes AES 128 or AES 256	Yes TLS	Server side
Dropbox [19]	Yes AES 256	Yes SSL/TLS and AES 128 or higher	Server side
iDrive [20]	Yes AES 256	HTTPS	Server Side with client key
Egnyte [21], [22], [23]	Yes AES 256	SFTP and TLS protocol and AES 256	Server Side Also integrates Client Side

It is overserved that cloud providers are using mainly Advanced Encryption Algorithm (AES) with 128-bit key or 256-bit key for encrypting data in rest. And for transferring data from user to cloud provider mainly Transport Layer Security (TLS) is used by most of the cloud providers.

Why AES?

There are many studies conducted by various authors to prove that AES algorithm is secure and faster algorithm as compared to other algorithms. This may one of the reasons for selecting AES algorithm for encryption by various cloud providers. AES, DES and RSA encryption algorithms are compared on various parameters like key size, encryption time, decryption time, security, rounds, simulation speed, etc. and found that AES is most secure algorithm. Encryption time for AES is lesser. So, it is concluded that AES is better in terms of security and encryption time as compared to DES and RSA algorithm [24]. In [25], AES and Blowfish algorithms are compared and authors concluded that AES can be used in situations where high security is needed. Also AES shows better encryption performance with images. Comparative analysis of RSA and AES algorithms concludes that AES is faster and safer algorithm. AES offers better security and has lesser implementation complexity. Hence, AES has emerged as one of the most efficient and strongest algorithm today [26]. Various metrics such as encryption-decryption time, throughput, and memory utilization are used to compare DES, 3DES, and AES algorithms. Encryption-decryption time is lesser in case of AES and thus it is faster algorithm and it also offers more throughput. But AES takes more memory when compared with DES [27]. Comparative analysis of DES, AES and RSA is tabulated in Table 3 based on key length, cipher type, block size, security, easiness in hardware and software implementation, encryption/decryption speed etc [28]. These features show that AES is more secure and faster.

Table 3: comparative study of AES, DES and RSA [28]

Features	DES	AES	RSA
Developed	1977	2000	1977
Key Length	56 bits	128,192,256 bits	More than 1024 bits
Cipher Type	Symmetric block cipher	Symmetric block cipher	Asymmetric block cipher
Block Size	64 bits	128 bits	Minimum 512bits
Security	Not secure enough	Excellent secured	Least secure
Hardware and Software Implementation	Better in hardware than software	Better in both	Not efficient
Encryption and Decryption	Moderate	Faster	Slower

CONCLUSION

Cloud computing is the technology that provide computing resources to the users in a manner similar to utilities like water, gas, electricity, etc. Huge storage facility provided by cloud providers is one of the biggest benefits provided to users. Using this facility, consumers are moving all of their data to cloud environment giving rise to security issues of sensitive data. To provide security against issues like privacy and confidentiality, various cryptographic algorithms are used. In this paper, various cryptographic algorithms like AES, RSA, DNA, BDNA etc. are studied and reviewed. Upon studying encryption facilities used by cloud providers, it can be concluded that AES is mostly used encryption algorithm as it fast and secure algorithm when compared with other algorithms.

REFERENCES

[1] P. Mell and T. Grance, "NIST definition of cloud computing," *NIST Special Publication 800-145*, 2011. .

[2] M. Arun Fera, C. Manikandaprabhu, I. Natarajan, K. Brinda, and R. Darathiprincy, "Enhancing security in Cloud using Trusted Monitoring Framework," *Procedia Comput. Sci.*, vol. 48, no. C, pp. 198–203, 2015, doi: 10.1016/j.procs.2015.04.170.

[3] I. G. Amalarethinam and H. M. Leena, "Enhanced RSA Algorithm with Varying Key Sizes for Data Security in Cloud," *Proc. - 2nd World Congr. Comput. Commun. Technol. WCCCT 2017*, pp. 172–175, 2017, doi: 10.1109/WCCCT.2016.50.

[4] A. Grover and B. Kaur, "A framework for cloud data security," *Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA 2016*, pp. 1199–1203, 2017, doi: 10.1109/CCAA.2016.7813924.

[5] M. P. Rewagad and M. Y. Pawar, "Use of digital signature with diffie hellman key exchange and aes encryption algorithm to enhance data security in cloud computing," *Proc. - 2013 Int. Conf. Commun. Syst. Netw. Technol. CSNT 2013*, pp. 437–439, 2013, doi: 10.1109/CSNT.2013.97.

[6] N. Khanezaei and Z. M. Hanapi, "A framework based on RSA and AES encryption algorithms for cloud computing services," *Proc. - 2014 IEEE Conf. Syst. Process Control. ICSPC 2014*, no. December, pp. 58–62, 2014, doi: 10.1109/SPC.2014.7086230.

[7] A. Sachdev and M. Bhansali, "Enhancing Cloud Computing Security using AES Algorithm," *Int. J. Comput. Appl.*, vol. 67, no. 9, pp. 19–23, 2013, doi: 10.5120/11422-6766.

[8] M. Sohal and S. Sharma, "BDNA-A DNA inspired symmetric key cryptographic technique to secure cloud computing," *J. King Saud Univ. - Comput. Inf. Sci.*, 2018, doi: 10.1016/j.jksuci.2018.09.024.

[9] S. Namasudra, D. Devi, S. Kadry, R. Sundarasekar, and A. Shanthini, "Towards DNA based data security in the cloud computing environment," *Comput. Commun.*, 2020, doi: 10.1016/j.comcom.2019.12.041.

[10] M. M. Potey, C. A. Dhote, and D. H. Sharma, "Homomorphic Encryption for Security of Cloud Data," *Procedia Comput. Sci.*, vol. 79, pp. 175–181, 2016, doi: 10.1016/j.procs.2016.03.023.

[11] "Encryption of Data at Rest - Encrypt Data at Rest with Amazon EFS Encrypted File Systems." [Online]. Available: https://docs.aws.amazon.com/whitepapers/latest/efs-encrypted-file-systems/encryption-of-data-at-rest.html. [Accessed: 13-Feb-2020].

[12] "Encryption of Data in Transit - Encrypt Data at Rest with Amazon EFS Encrypted File Systems." [Online]. Available: https://docs.aws.amazon.com/whitepapers/latest/efs-encrypted-file-systems/encryption-of-data-in-transit.html. [Accessed: 13-Feb-2020].

[13] "Azure encryption overview | Microsoft Docs." [Online]. Available: https://docs.microsoft.com/en-us/azure/security/fundamentals/encryption-overview. [Accessed: 13-Feb-2020].

[14] "Encryption at Rest in Google Cloud Platform | Documentation." [Online]. Available: https://cloud.google.com/security/encryption-at-rest/default-encryption/. [Accessed: 13-Feb-2020].

[15] "Encryption in Transit in Google Cloud | Documentation." [Online]. Available: https://cloud.google.com/security/encryption-in-transit/. [Accessed: 13-Feb-2020].

[16] "Data encryption options | Cloud Storage | Google Cloud." [Online]. Available: https://cloud.google.com/storage/docs/encryption/. [Accessed: 13-Feb-2020].

[17] "Security architecture: Data - IBM Cloud Architecture Center." [Online]. Available: https://www.ibm.com/cloud/architecture/architectures/securityArchitecture/security-for-data#dataencryptionandkeymanagement. [Accessed: 13-Feb-2020].

[18] "Shield Platform Encryption Architecture." [Online]. Available: https://www.salesforce.com/content/dam/web/en_us/www/documents/reports/wp-platform-encryption-architecture.pdf. [Accessed: 13-Feb-2020].

[19] "How Dropbox keeps your files secure | Dropbox Help." [Online]. Available: https://help.dropbox.com/accounts-billing/security/how-security-works. [Accessed: 13-Feb-2020].

[20] M. Road, "Online backup subscription - service and security overview Introduction Sign up Download IDrive State-of-the-art s security and e encryption Automated and i intelligent backups Versatile and intuitive."

[21] "Egnyte Information Security Procedures." [Online]. Available: https://www.egnyte.com/enterprise-tos/information-security-procedures.html. [Accessed: 13-Feb-2020].

[22] "Online backup subscription - service and security overview." [Online]. Available: https://egnyte-www-static.egnyte.com/assets/pdfs/white-papers/Whitepaper-Egnyte-SecurityArchitecture-sa.pdf. [Accessed: 13-Feb-2020].

[23] "Cloud File Server features: online storage, file sharing, FTP, file transfer, security and access control, user management - Egnyte Cloud File Server." [Online]. Available: https://www.egnyte.com/file-server/online-file-server-features.html. [Accessed: 13-Feb-2020].

[24] M. Prerna and S. Abhishek, "A study of encryption algorithms AES, DES and RSA for security," *Global Journal of Computer Science and Technology,* 2013.

[25] M. Anand Kumar and S. Karthikeyan, "Investigating the Efficiency of Blowfish and Rejindael (AES) Algorithms," *Int. J. Comput. Netw. Inf. Secur.*, vol. 4, no. 2, pp. 22–28, 2012, doi: 10.5815/ijcnis.2012.02.04.

[26] A. Al Hasib, A. Ahsan, and M. Haque, "A Comparative Study of the Performance and Security Issues of AES and RSA Cryptography," no. November 2001, pp. 505–510, 2008, doi: 10.1109/ICCIT.2008.179.

[27] K. Shaify, M. Meenakshi, "Performance evaluation of various symmetric encryption algorithms," *Proc. – 2014 IEEE Conf. Parallel, Distributed and Grid Computin,* no December, pp. 105-109, 2014.

[28] M. P. Babitha, "Secure Cloud Storage Using AES Encryption," pp. 859–864, 2016.

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

USAGE ANALYSIS USING PLAYSTORE AND APPSTORE APPS

Harshali Patil, Lavina Jadhav and Sarala Mary
MET, Institute of Computer Science

ABSTRACT

Mobile users' demands are increasing day by day. As per forecast by Statista, by end of 2020, worldwide smart phone users will be 6.95 billion. This tremendous growth of smart phone users increases demand for apps. The average numbers of apps used per day by each user are more than 10 and average time spend per day by android users is approximately 2.5 hours. This paper describes the trending apps for Android users. It also identifies the users' preferences and can recommend apps for users.

The trend analysis will help the developers to identify the choices of users and based on reviews enhancements can be done in the current versions. Most useful and least used apps can be identified and necessary actions can be taken to reduce the number of unnecessary apps on playstore. This paper also identifies most trending apps and categories. The trend analysis will not only useful to identify regular users' pattern but also it can identify category wise usage.

Keywords: trend analysis, playstore, appstore, data analysis, classification

I. INTRODUCTION

Mobile users are increasing extensively and hence the different services offered to users. Indian telecom subscribers as of December 2017 (in millions) are depicted in fig.2. In last 5 years Indian mobile users' growth rate is represented in chart shown in fig.2. The smartphone penetration rate in India from 2014 to 2019 as share of mobile phone users is represented in this statistic. Around 39 percent of all Indian mobile users expected to own a smartphone by 2019. The rate of increase in smartphone users is linear.

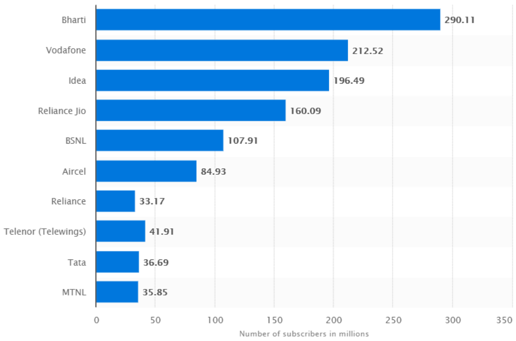


Fig. 1 Number of mobile telecom subscribers in India as of December 2017, by company (in millions) [1]

Smart phone users in India are depicted in figure 2. The graph provides an idea about increase in use of smart phone users. The graph shows 2014 to 2019 smart phone users in India.

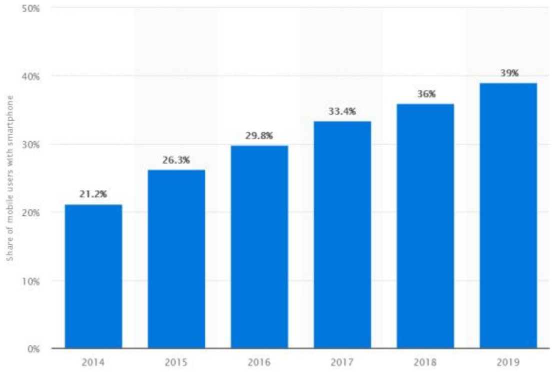


Fig.2 Share of mobile phone users that use a smartphone in India from 2014 to 2019* [2]

As per the comScore report, 70% of internet users in India access the web exclusively via mobile devices which demonstrates the importance of smartphones and feature phones to widespread internet access [3].

The report says that in Indonesia the percentage of mobile-only internet users is similarly high. The comScoe report provides the study of 13 markets and details about % of internet users [3] is shown in following table 1.

Table 1: % of internet users

Country	% of internet users
India	70
Indonesia	67
Mexico	37
Spain	32
Brazil	29
China	22
United States	12
Canada	8
UK	8
France	6
Germany	4

Percentage of number of mobile users in India in past one year is 51.07% where as desktop users are 44.84% and tablet users are 4.09%. The figures indicates that mobile users percentage is comparatively high than desktop and tablet users.

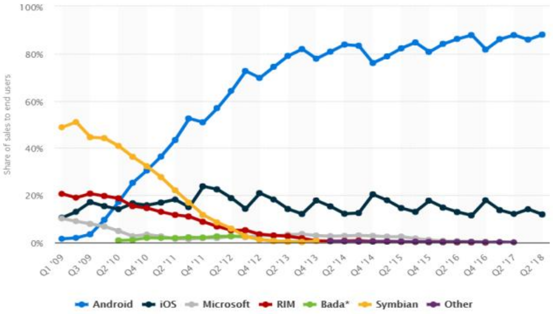


Fig. 3 Global mobile OS market share in sales to end users [3].

The above graph shows the growth of Android users. iOS users consistently approximately 20% in a decade. Hence the analysis of user trend can be done using Android apps and iOS apps. Number of Android users is more compared to iOS users hence the apps available on Android platform are increasing. When users download the app the privacy and security is major concern. The security of the mobile users is a crucial factor. The most common ways to secure your app are as follows

To overcome the server side vulnerabilities is to scan them. By scanning the app using automated scanner can solve many problems.

Binary protection is one of the important parts in security. Binary hardening techniques analyzes binary file and modify it to protect against the exploits.

The secure data storage across platform is useful to protect data. Additional level of encryption is provided by the operating system.

To secure transport layer by taking care of things like do not send password using SMS, MMS etc. if mobile app detects invalid certificate then provide alert to user. Use of SSL versions of third party analytics companies.

Unintended data leakage can be avoided by monitoring leakage points (like caching, browser cookies, logging etc)

II. PLAYSTORE DATA PROCESSING

To analyze trend, in this paper data used is from “kaggle” (www.kaggle.com) open dataset repository. The playstore data which is used in this research paper is having 10481 samples and dataset has 13 attributes. The open dataset consist of attributes as follows shown in following Table 2.

Table 2: Attributes of playstore data

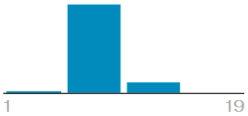
Attribute	Description
App	Name of application
Category	App category
Rating	User rating of the app when scraped
Reviews	Number of user reviews for the app when scraped

Size	Size of the app when scraped
Installs	Number of downloads/installs by users when scraped
Type	Free or Paid
Price	Price of app in dollars when scraped
Content rating	Age group of the app targeted at like Children / Mature 21+ / Adult
Genres	Apart from main category, an app can belong to multiple genres. e.g. a musical family game will belong to Music, Game , Family genres.
Last updated date	When the app was last updated on play store when scraped
Current version	Current version of the app available on play store when scraped
Android version	Minimum Android version when scraped

The preprocessing techniques are applied on the dataset. Data cleaning is done using manual method. The app names are having emoticons and special characters are manually processed and all special characters are replaced by null. Using WEKA tool unsupervised technique "replacemissingvalues" is applied on the Play Store data set. This method replaces all missing values of nominal and numeric attributes present in the dataset with the modes and means from the training data. "Numerictonominal" techniques are applied to convert the dataset into nominal attributes.

The data summary is listed in the following table 3

Table 3: Attribute summary of playstore data

Attribute	**Summary**
App	9669 unique values
Category	Family 18% Game 11% Tools 8% Medical 4% Other (30) 59%
Rating	
Size	Varies with device 16% 11M 2% Other (460) 83%
Type	Free 93% Paid 7%
Content Rating	Everyone 80% Teen 11% Mature 17+ 5% Everyone 10+ 4% Other (3) 0%
Android version	4.1 and up 23% 4.0.3 and up 14% Other (33) 64%
Installs	1,000,000+ 15% 10,000,000+ 12% 100,000+ 11% 10,000+ 10% Other (18) 53%
Genre	Tools 8% Entertainment 6% Education 5% Medical 4% Other (116) 77%

III. APPSTORE DATA PROCESSING

The App Store dataset has 7197 has samples and 16 attributes. The attributes are as follows shown in table 4

ISSN 2394 - 7780

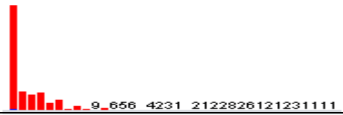
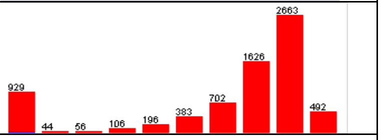
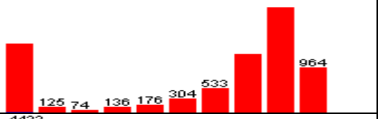
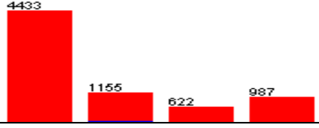
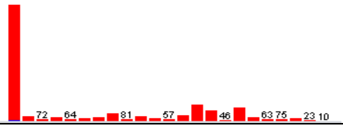
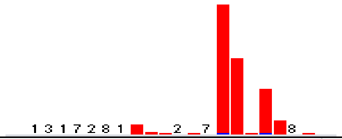
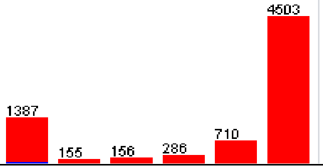
Table 4: Attributes of Appstore data

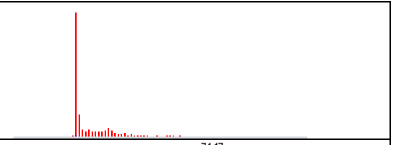
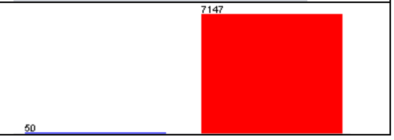
Attribute	Description
Id	App identification
Track name	Name of app
Size bytes	Size in bytes
Currency	Type of currency
Price	Price amount
Rating count tot	Rating counts of user for all versions
Rating count ver	Rating count of user for current version
User rating	Average user rating value for all version
User rating ver	Average user rating value for current version
User rating	Average user rating value for all version
Ver	Latest version code
Cont rating	Content rating
Prime genre	Primary genre
Sup_devices.num	Number of supporting devices
Ipadsc_urls.num	Number of screenshots showed for display
Lang.num	Number of supported languages
Vpp_lic	Vpp device based licensing enabled

The preprocessing technique applied on the dataset, numeric attributes are converted into nomial. The unsupevised preprocessing techniques "Numerictonominal" is applied on all numeric attributes.

The data summary is listed in the following table 5

Table 5: Attribute summary of appstore data

Attribute	Summary
Id	7197 unique values
Track_name	6356 (88%) unique values 814 (12%)
Price	
User_rating	
User_rating_ver	
Cont_rating	
Prime genre	
Sup_devices	
ipadSc_urls	

Lang	
Vpp lic	

Using apriori algorithm having support as 10% and confidence as 90%. The association rules for AppStore data is as follows

1. vpp_lic=1 7147 ==> currency=USD 7147
2. ipadSc_urls=5 4503 ==> currency=USD 4503
3. ipadSc_urls=5 vpp_lic=1 4488 ==> currency=USD 4488
4. cont_rating=4+ 4433 ==> currency=USD 4433
5. cont_rating=4+ vpp_lic=1 4417 ==> currency=USD 4417
6. ipadSc_urls=5 4503 ==> vpp_lic=1
7. currency=USD ipadSc_urls=5 4503 ==> vpp_lic=1 4488
8. ipadSc_urls=5 4503 ==> currency=USD vpp_lic=1 4488
9. cont_rating=4+ 4433 ==> vpp_lic=1
10. currency=USD cont_rating=4+ 4433 ==> vpp_lic=1 4417

Using Association rule mining technique and with apriori algorithm applied on PlayStore dataset having minimum support as 60% and confidence as 90% provides following set of association rules

1. Type=Free 10039 ==> Price=0 10039
2. Price=0 Content Rating=Everyone 8019 ==> Type=Free 8019
3. Type=Free Content Rating=Everyone 8019 ==> Price=0 8019
4. Price=0 Android Ver=4.1 and up 2320 ==> Type=Free 2320
5. Type=Free Android Ver=4.1 and up 2320 ==> Price=0 2320
6. Price=0 10040 ==> Type=Free 10039
7. Android Ver=4.1 and up 2451 ==> Type=Free 2320
8. Android Ver=4.1 and up 2451 ==> Price=0 2320
9. Android Ver=4.1 and up 2451 ==> Type=Free Price=0 2320
10. Content Rating=Everyone 8714 ==> Type=Free 8019

IV. DATA MODEL AND ANALYSIS

For PlayStore dataset using simple K-means algorithm with number of clusters as 2, with distance function as Euclidean distance and maximum iterations as 500. The dataset clusters are formed as follows

Cluster	Clustered instances
C0	9425 (87%)
C1	1416 (13%)

For AppStore dataset using simple K-means algorithm and applying similar parameters as like PlayStore dataset

Cluster	Clustered instances
C0	4276 (59%)
C1	2921 (41%)

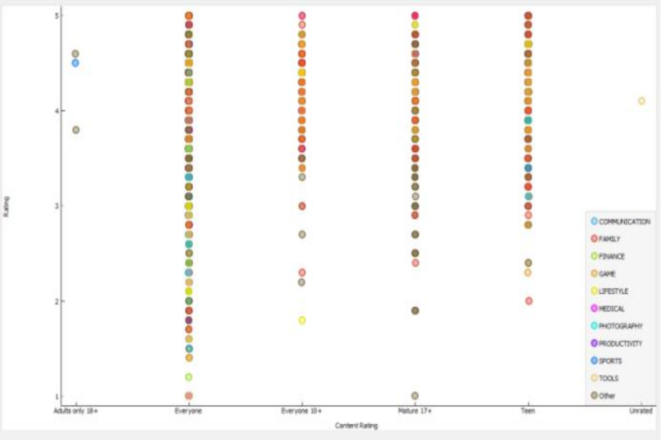


Fig 4. Content rating vs rating plot of Play Store dataset

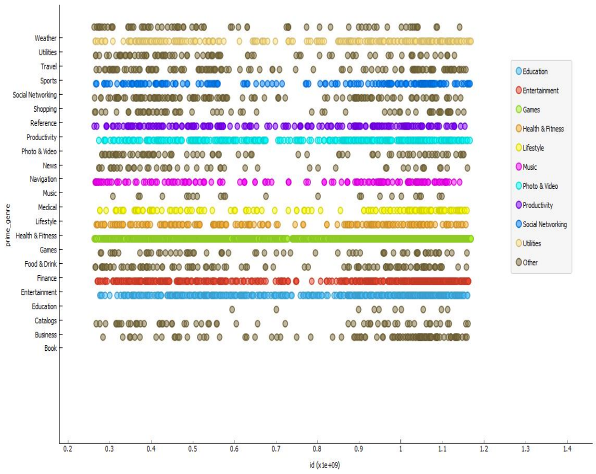


Fig. 5: id vs prime genre plot of App Store dataset

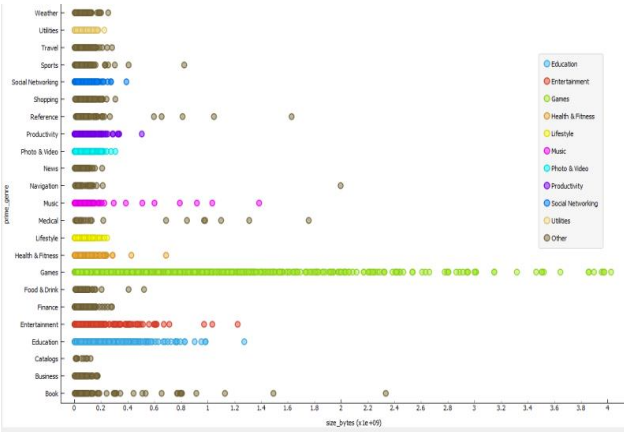


Fig. 6 Size bytes vs prime genre plot of App Store dataset

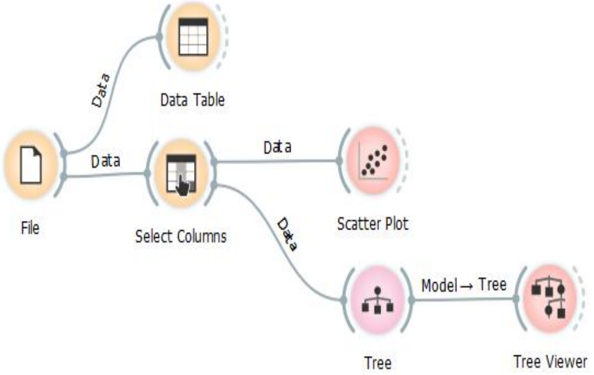
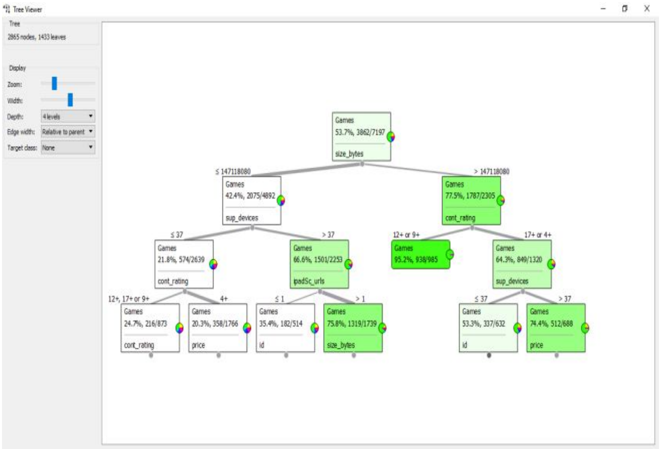


Fig.7 Tree model of App Store using ORANGE

ISSN 2394 - 7780



Tree Viewer

Tree
2865 nodes, 1433 leaves

Display
Zoom:
Width:
Depth: 4 levels
Edge width: Relative to parent
Target class: None

Fig.8 Depth 4 tree model of App Store dataset

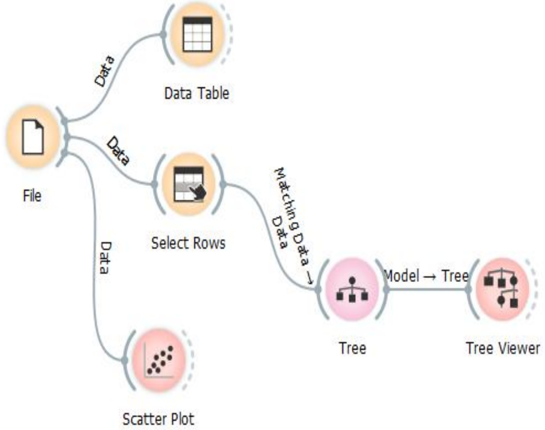


Fig.9 Tree model for Playstore dataset

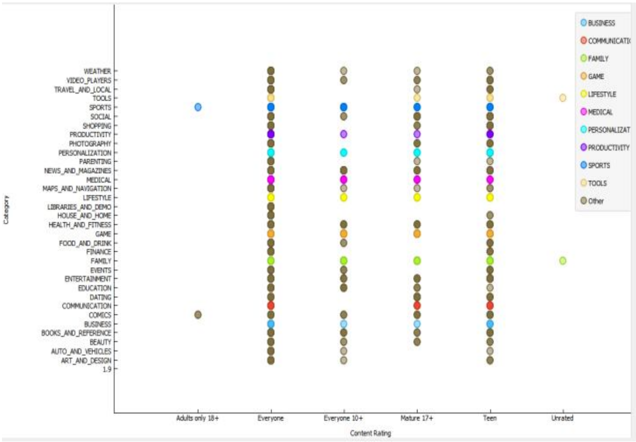


Fig.9 Content rating vs. category plot of PlayStore dataset

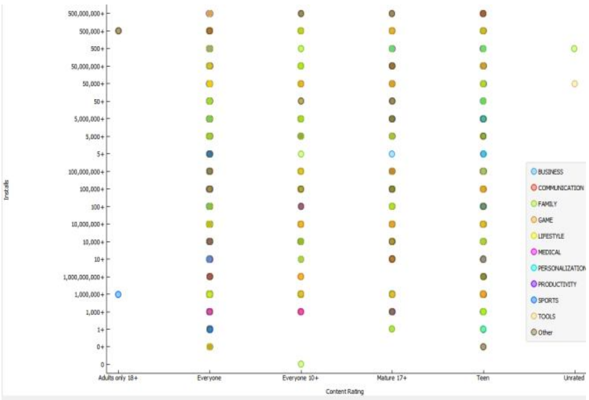


Fig.10 Content rating vs. installs plot of PlayStore dataset

V. CONCLUSION

The data modeling and analysis is done using WEKA and ORANGE tool. The Play Store apps data has huge potential to take app-making businesses to success. Data insights can be obtained for developers to work on current play store apps scenario and capture the Android market. Data visualization helps to get clarity of user liking. Appstore dataset analysis helps to identify that the dominant currency is USD, hence it helps to find that number of iOS users are more in those countries which trades in dollars. The numbers of android users as compared to iphone users are in 30 fold hence this analysis will definitely help the android market to analyze the user trend and will help developers, designers to come up with new ideas to sale products online.

REFERENCES

[1] Number of mobile telecom subscribers in India as of December 2017, by company (in millions), https://www.statista.com

[2] Share of mobile phone users that use a smartphone in India from 2014 to 2019* , https://www.statista.com

[3] Felix Richter, “Mobile Devices Put the Worldwide in WWW”, Article on internet usage worldwide , Nov 2017

[4] Web reference: www.kaggle.com

NETWORK FUNCTION VIRTUALIZATION OVERVIEW, ARCHITECTURE AND USE-CASES

Sooraj Shravan Prajapati
Information Technology, S.S & L.S Patkar Varde College, Unnat Nagar, Goregaon (W) Mumbai, Maharashtra

ABSTRACT

Network function virtualization is emerging as a network technology in telecom industry for providing agility and efficiency and flexibility in the deployment of network services to users. The network function virtualization can be easily created and migrated from one place to another without the installation of specialized hardware.it allows the faster deployment of services to user and gives great opportunities in the networking world.in this paper overview and architecture and use cases of the network function virtualization technology is presented.

Keywords: NFV architecture, use-cases, benefits, efficiency, reliability, security.

I. INTRODUCTION

NFV is an approach for telecom industry where the network node entities are traditionally used for hardware things.

The network Function Virtualization can be idea that utilizes the information technology fundamental to create or make virtualized network environment. This requires high volume(storage) serves, switches and network nodes that might be placed in information centres or we can say as centralized location. It involves the implementation of community options in such a way that

It involves implementing community options in a very code that it can run on enterprise trendy hardware and, which is generally moved to different locations within the community based on the necessity. It is able to provide flexibility and reliability to standardize the great deal of hardware. It virtualizes the major parts of networks and tries to minimize the less use and consumption of more hardware resources. There will be creation of network nodes and it is distributed all over the PC’s / servers. Using this network blocks will be created that will be forming an enterprise network and it will be providing the different-different services to tele-co industry. NFV uses high volume of servers, firewalls, switches and balancers to provide better quality of services.

Examples of the virtualized function includes: Load balancers, Firewalls, WAN-accelerators, routers.

Network functions virtualisation and software package outlined networking area unit terribly closely connected, however they're not constant. typically, the terms area unit incorrectly used synonymously

II. FRAMEWORK

A network, with any system mistreatment Network virtualizations techniques are often counteracted into variety of components.

A. Virtualized network fuctions (VNF’s):

It comprises the network functions to provide the variety of functions in the virtualized format, which can be easily deployable on to the hardware.

i.e. Network operate virtualization infra.

B. Network Function Virtualization infra (NFVI):

It consists of all hardware & software system parts that contain all the interval surroundings during the deployments of Network Function Virtualization infra.

One of the main benefits of this is it can be settled easily across the many physical location. These locations are the elements or entities of Network Function Virtualization infra.

C. Network virtulization management, orchestration subject field frame-work (NFV), (MANO):

It consists of many useful blocks that is use manipulate and exchange the data which is required to manage the control flow and supply the important elements.

It also manages and monitors the fail-overs and supplies effective security.

III. NFV & SDN

Network functions virtualisation and software package outlined networking area unit terribly closely connected, however they're not constant. typically, the terms area unit incorrectly used synonymously.

A. Software-Defined networks (SDN):

It deals with the replacement networking protocols with centralized management. It reduces the complexity in the distributed networking systems so that efficiency and flexibility of network can be increased.

B. Network-Function Virtualization (NFV):

It replaces the network parts of network environments with the algorithms which runs on different servers. In other words, we can say that it optimizes the network services and decouples the network-functions from the hardware. Including this algorithm on different- different machines will increase the complexity and increase the flexibility in the network services.

It can be divided into four layers

- A Virtualized-Network Functions is the basic block design in the block of NFV. It first virtualises the network and then provides the services

i. Virtualized-Network Functions (VNF)

It has two sub-parts i.e. VNF and EMS. A VNF that is virtual network is the basic block of design it virtualizes the network to perform routing process.

For example, once a router is virtualized router will set the base station and VNF similarly it can be a DHCP server or Firewall VNF.

A VNFs square measure deployed on Virtual Machines (VMs). A VNF could also be deployed on multiple VMs wherever every VM hosts one perform of VNF. However, the complete VNF can even be deployed get on one VM furthermore. Element Management System (EMS) is responsible for the useful management of VNF. The management functions embody Fault, Configuration, Accounting, Performance and Security Management. AN EMS could manage the VNFs through proprietary interfaces. there's also one EMS per VNF or one EMS which can manage multiple VNFs. EMS itself could also be deployed as Virtual Network perform (VNF).

ii. NFV-Infrastructure (NFVI)

NFV infrastructure is a combination of hardware and software system components that execute, manage, and kill VNFs. NFV's infrastructure consists of multiple locations, which provide network assets as part of NFV's infrastructure.

It includes both the things hardware resources and virtualization layers.

Hardware resource contains storage, computing, networks and processing speeds & also property to VNF through virtualization layers which is also called hypervisor layers. Hypervisor abstract all the hardware resources and separates the VNF code from hardware resources. Extracting and logically dividing physical resources into a hardware capture layer, so the code can be used on completely different physical resources.

iii. Operation-Support Subsystem (OSS)

It deals with network management, fault management, configuration management & repair management. BSS works with client management, product management and order management. In NFV design, the ANS operator's decoded BSS / OSS is also integrated with the common interface of NFV management and orchestration abuse.

iv. Management & Orchestration (MANO)

It manages the interactions of a VNF with computing, storage & network resources. For example, increasing the virtual machines & energy potency. It also takes care of allocating the virtual machines on hypervisors and also storage and network properties.

It also takes cares of VNF life cycle managements which involves installation, updating and query and termination. Its deployed for every new service and manages the multiple VNF.

Orchestrator manages the infrastructure of virtual network functions and also the system resources based on the network realization services. There is always one free block identified as free-lance block, this includes the data sets that contains information for VNF templates, forwarding graphs & service infrastructure.

IV. USE CASES

The Network Function Virtualization provides variety of branches in use-cases. It provides the platforms for purchasing and deploying the network packets. It is associated with end to end Orchestration and management in enterprise infrastructure.

It provides better agility and efficiency in automation of networks and also provides the flexibilities in deploying the services it improves the business potency in capital and operation by providing business needs and fulfilling their requirements.

It also cuts the complexness from services and operations and provides the many possible solutions for the requirements.

Network Virtualization

The main use with that NFV technologies are getting used by in the main medium corporations round the world is, of course, for network virtualization. As already mentioned before, NFV separates the hardware from the computer code. the method creates a virtual network on prime of the physical network. This decoupling of hardware and computer code permits service suppliers to expand and accelerate the event and innovation of services.

It conjointly helps to enhance essential network needs like provisioning. In order to optimize their network services, customers look to network virtualization to separate their network functions like DNS, caching, IDS, and firewalling from the proprietary hardware that was, till recently, the dominant solutions. This resolution conjointly permits them to run on computer code instead.

Mobile edge Computing

When it invoves mobile, edge computing comes into picture like raido towers, mini data and native knowledge centers.

NFV takes a number of these mobile network service functions and interprets them from hardware to software system.

Network perform virtualization, aboard alternative technological and network advances and developments like software-defined networking and AI, can seemingly become the prime solutions for the network challenges of tomorrow thanks to their early integration and combination with one another.

Security

Just like the tools we have a tendency to use to farm our crops or manufacture our cars, the tools we have a tendency to use to shield our physical and virtual tools have evolved due to the varied leaps in technological progress that have occurred over the last decade.

Many security vendors area unit already providing virtual firewalls to shield VMs. The F5 Gi Firewall VNF Service, for instance, is one amongst the foremost widespread NFV solutions that encompasses firewall capabilities. however really, firewalls area unit only 1 of nearly each security device or element that may eventually be virtualized mistreatment network functions virtualization similarly as software-defined networking.

One of the most attractions to mistreatment virtualized Security is that the plan of centralized management mechanisms and equality of distributed social control.

These two edges alone have seen corporations trying to bolster their security flock to research these varieties of security solutions.

Video analytics

Another technology that has seen an enormous increase in its potential, since the origination of the web of Things, is video analytics systems and computer code. Now, corporations square measure able to capture huge amounts of information exploitation IoT video and sensible devices put in in their factories, stores, offices, and even farms.

But most of the time, superior AI video analysis is performed solely cloud-native applications or powerful servers placed on the cloud. therefore, having to transfer these giant amounts of information for analysis from on-premises to the cloud becomes a true challenge.

The trendy network sometimes faces AN end-to-end network latency, that poses a true challenge for the apps and network services that square measure very sensitive to network delays, like video analytics.

ACKNOWLEDGMENT

The present research work is an effort to show some light on “Network Function Virtualization Overview, Architecture and cases”. This work would not have been done without the supervision and help to me by number of people.

I would like to thanks to all my friends and all the staff of my college who has helped me a lot to complete this research.

REFERENCES

[1]. M. Wegmuller, J. P. von der Weid, P. Oberson, and N. Gisin, “High resolution fiber distributed measurements with coherent OFDR,” in Proc. ECOC’00, 2000, paper 11.3.4, p. 109.

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

[3]. Published E2E Arch, REQ, Use Case, Terminology documents in ETSI NFV Open Area.

[4]. NFV and SDN—key technology enablers for 5G networks.

SHOULD PEOPLE GET CHIP IMPLANTED UNDER THEIR SKIN?

Tvisa Manish Kadakia
Master of Science in Information and Technology, SVKM's Usha Pravin Gandhi of Arts, Commerce and Science, Affiliated with Mumbai University, Mumbai

ABSTRACT

Microchip implants for the humans is not new. Today, the potential for the microchip to be embedded inside a body of Humans has been realised. The concept of implantation has been widely accepted by the people of Sweden who favours the convenience rather than focusing on the potential health risk. RFID (Radio Frequency identification)will be storing the 16-digit identification number that will be unique identifier for the lifetime .The purpose of this study is the way to interpret and explore the perceptions of Respondents towards RFID. (J. Technol. Manag. Innov. 2008)

Keywords: Microchip ,Implantation, Embedded, Unique Identifier, RFID

I. INTRODUCTION

RFID (radio frequency identification) Tag had its humble beginnings from very start.I t had shown its usage from the very start of WorldWar II where it was being used to identify the Airplannes, track the nuclear weapons,materials and animals .From there onwards it has spread it roots to the various aspects such as Material Managemnt, Logistics ,EmployeeTracking, Access Control and many more to continue.

When it comes to Human being the concept of RFID becomes highly appreciated since it provides digital interface to the real world.UIL(Unique Identifier for Lifetime) is a 16-digit identification number assign to each RFID that will be acting as a primary key to mark each individual unique from others. The burdensome of carrying your personal documents or storing it within your personal devices will be reduce to a greater extent by letting your chip be storing all of your information such as credit card details, medical reports ,library card and many more other sources of information.Along with pros there are always the cons that comes hand in hand so in this case they are in terms of privacy issues basically it includes being able to track the person previous and current location, privacy and legacy concerns.

II. PROBLEM

There are always benefits and problems associated to the human microchipping.The important problem associated towards this microchipping is the person privacy.Since most of the people around the world have given the convenience the first priority over the potential health risk ,so for them the privacy infringement doesn't matters but for some of the people what matters is the fact that if they are storing their personal information within a tiny chip than the some security policy should be assign .As the purpose of this chip is to make the life of the people easy .The person embedding the chip inside his or her body they are the one who are liable towards the fact that who has the access towards their data.It so vary depends upon the type of the information stored within the chip.There is no deny to the fact that various hackers are their lying all over the world that are trying to hack your data by accessing your previous and current location,monitoring your bank transaction and many more .They are to questionable enough whether they hold the right to sell your data to the third party without your concern.A second potential problem is associated with the fact that who will be storing the information and to whom the access should be provided .There also lies the fact which can not be neglected is the health problems that comes along with the microchipping.

III. HISTORY

The British Scientist Kevin Warwick carried out the first experiment in the year 1998.He is the professor of cybernetics at Reading university.Underwent an operation by placing a chip under the skin at his forearm in order to conduct a study on how to control of intelligent buildings run by the computers.The microchip emitted an identifying signal that the computer would be recognising and would operate the various types of an electronic devices such as room lights,opening of a door and many more.The size of the microchip was 23mm long and 3mm wide.

Professor Warwick said that "The potential for this technology is enormous.It is quite possible for an implant to replace an Access or Visa Card.There is very little danger in losing an implant or having it stolen." (professor has worlds first silicon chip implant, 26 August 1998).The implant was removed after nine days and since then it has its place within the Science Museum in London.

Amal Graafstra a hobbyist in early March,2005 implanted a 125khz EM$102 bioglass-encased RFID Transponder into hid left hand.The purpose of the implantation was access control system in order to gain entry to his office soon after he implanted an advance microchip called as HITAG S 2948 which was generally a low frequency transponder which he used to access his home,

open car doors and to log on to his computer. (Amal Graafstra-Technologist,Author & Double RFID Implantee, 2017-05-26). In 2007 he authored the book RFID Toys, (Dangerous Things, 2017-05-26).

On 16 March 2009 the British scientist Mark Gasson was the one who surgically implanted RFID device into his left hand.In April 2010,Gasson's team demonstrated how a computer virus can be wirelessly infect the chip and also be transmitted to other systems. (Gasson, 2010).

Gasson stated that with the implanted technology is the separation between man and the machine that can certainly become theoretical .He has no plan to remove his implants and was the first person to be credited with the computer virus. ("Could you become infected with a Computer virus?", 2019-07-24).

IV. RADIO FREQUENCY IDENTIFICATION TECHNOLOGY

RFID is generally developed for the purpose of identification. The size of the RFID is small and it can be attached to any of the physical objects as well as human beings. The working of the RFID Tag generally takes place by emitting an identification signal in the form of the radio waves which is queried by an RFID reader. The information will be captured and utilized by the reader and will be used for futher processing. It is generally considered as an replacement to that of an Barcode which provides various advantages. Provides us with multiple scanning and allows us to store various other information rather than that of identification number. RFID has been widely applicable in the chains of Logistics and Enterprise but it has also shown its way towards Medical, Agriculture and many more domains.

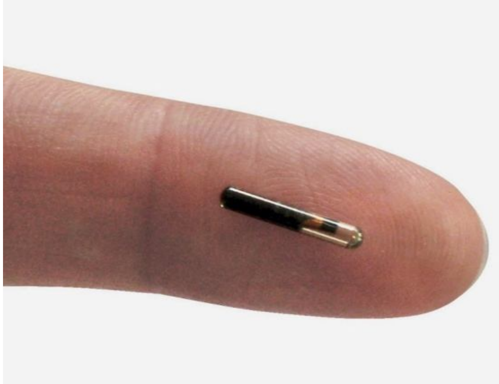


Figure-1: A RFID Tag (by Andrew Brown(freeCodeCamp))

A. RFID Implants

The concept of human microchipping is not something new that people are not familiar with. We generally have came across the concept where the RFID implanted under their skin is been connected to some other sort of devices for an easy access.To overcome this technique we will rather focus on the concept where the RFID would not be connected to some other devices for an easily access rather would be access individually.

The first implant called VeriChip came into the picture on October,2004. Recieved an approval from US Food and Drug Administration. It will storing an identity number(identification number) and can be read from the distance of 10-15 cm.The ID will be the unique identifier as an unique identify for everyone residing across the world.No other form of the data will be stored on the chip rather than that of ID .Centralized Database will be storing the data that is related to the owner only.

VeriMed is the first ever designed commercial application designed to identify the patients in the database .The concern is about who will be accessing those centralized database.An authorized person should be the one who can be provided an access based upon the concept of role-access based system.In the case of centralized database that contains the information about the patient medical history can be accessed by an authorized doctor through the password-protected website or some other forms of security.

They are the passive tags .They are not in requirement of the batteries to be operational .Therefore,they can be functional for much longer period of time.They are small in size and lacking in terms of internal power thus by limiting the performance of the device in terms of memory,processing power and communication. (3226936_RFID_implants_Opportunities_and_challenges_for_identifying_people).

B. Concerns

- ***Security Issues***

There exist the permanent bridge between an RFID Tag and person privacy infrigment.The people across the world are quite impress with the wonders that are generally being shown by this technology. But when it comes to using any of the technology there exist the various cons it eventually depends upon the every individual whether to take into consideration those cons or just leave it aside and take the advantage .

At times it becomes mandatory to take into consideration the fact that the above stated technique breaches the privacy of an individual. Since the privacy infringement is done by tracking or accessing the current or previous location of an individual ,capturing of the stored information within the chip and many more. But what matters the most is the question that “Is the information accessed by an Authorized Person?”,”Is third party allowed to make the use of your information without your concerns?”.

- ***Privacy and Ethical Considerations***

EGE(The European Group on Ethics in Science and New Technologies)published an opnion on the use of implants .They stated that there should pose the question regarding the privacy concern whether there are any legal or legislation on data protection and if not needs to be considered. (3226936_RFID_implants_Opportunities_and_challenges_for_identifying_people)

- ***Health Risk***

Inserting something into your body is tremendously danger and life threatening.So before you give a thought of implanting yourself with the chip think about the potential health risk that are associated with them.

For example ,Nonlonizing Radiation from microwave radio frequency and magnetic fields (Covacio, 2003),electrical hazards,tissues infections,infections related to the medical equipment and allergic to the metals that has been used within RFID.

Implanting something within the human bodies is problematic according to the medical condition since the scanner would not be able to detect the actual placement of the chip and would create the major problems.

Since the microchip implantation is like any other computing machine which can easily catch viruses since in computing machine it is possible to detect the viruses with the help of various net protectors but when it comes to the humans this viruses that attack the chips needs to be detected and solved for that a huge cost needs to be taken into consideration for finding an alternative solutions or replacement in some of the cases.

- ***Lack of Universal Standards.***

There are different identification standards and protocols followed across the globe which would require different and various types of microchip to be taken into consideration fitting into the predefined standards and protocols.

V. REVIEW OF LITERATURE

A. Social Acceptance

Social concerns are plague to this technology. Identification Technology generally have a low acceptance rate. In Sweden, more than 4000 of people have adopted this technology. Since than this technology has gain its importance in Sweden Country. “Swedes are less concerned about data privacy than people in other countries. High level of trust for Swedish companies, banks, large organization and government institutions. ”,said by Osterlund’s. (thousands-of-swedes-are-inserting-microchips-under-their-skin, 22/10/2012)

In United States ,legislation has been crafted in order to maintain the benefits associated with the RFID technology along with the disadvantage of privacy and security risk. California ,Georgia, Wisconsin and many more are among the states of U.S. which have passed legislation to prohibit forced implantation of RFID in humans.

Microchip Consent Act,2010 which became effective from July 1,2010 in state of Georgia where no person shall be required to be implanted with microchip ,but also voluntary implantation to be done by the physician under the authority of Georgia Composite Medical Board. (Michael).

Researchers continue to investigate social acceptance of the implantation of this technology in human bodies. A survey was conducted in which it is reported that the 15 % respondent “Very and Somewhat willing” to have a chip inserted under their skin in order to receive the benefits associated with it and most of the respondent belong to the age group of 45-55 .However ,60% of respondent “Not at all” prepared to have chip inserted under their skin in any circumstances and most of the respondent belongs to age group of 18-35.

B. Methodology

An analysis of the qualitative data obtained through the aforementioned survey is been included. Asking participants ,"Would you accept Microchip Implantation Technology within your body?".The collection and analysis of data was gleaned from closed ended questions as well as open ended questions. The survey explored the perspective of Indians as well as Non-Indians. Participants included both Indian and Non-Indian businessmen or businesswomen, employee and students who have currently enrolled in the college or still in the college.

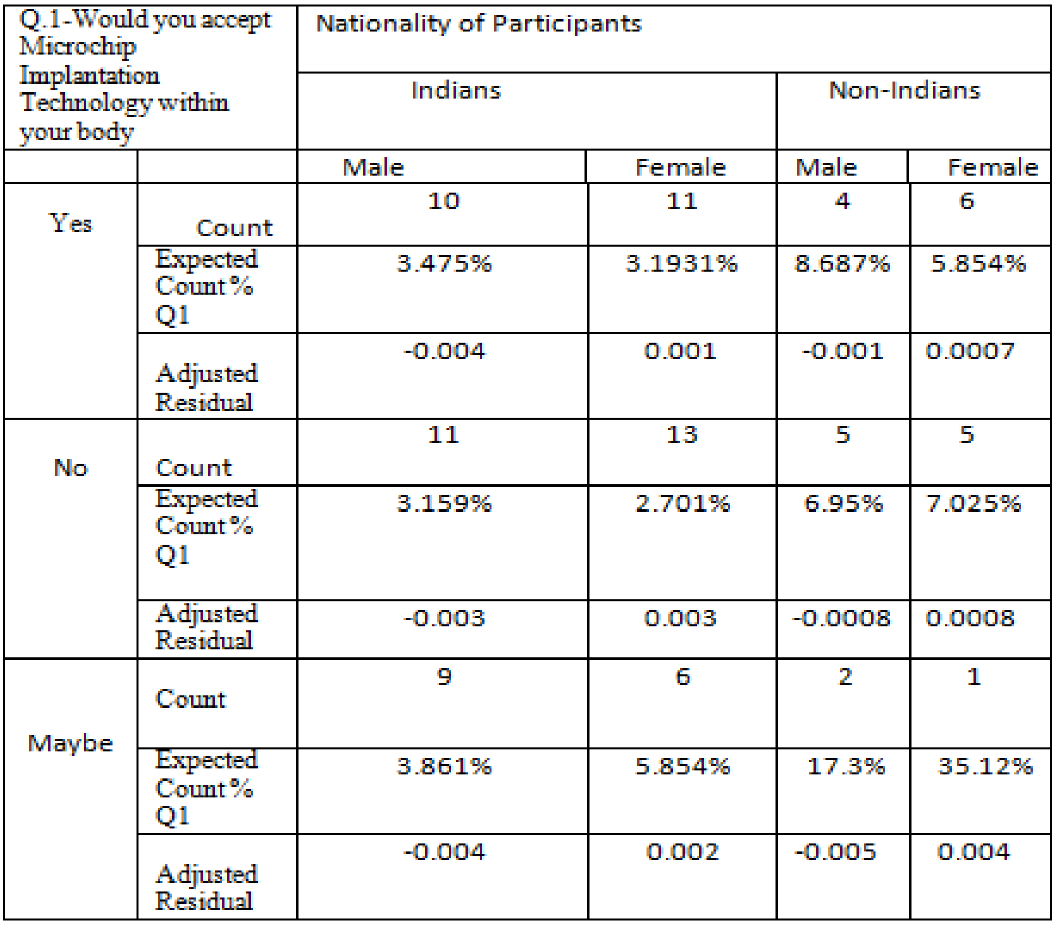
C. Findings

In first phase of study ,the frequency of data compared the ratio of male to female who are willing to accept the human microchipping. The participants was explored relative to the perception of surgically getting transponders under their skin (yes-no-maybe).The significant chi-square indicated that there lies a relationship between the participants opinions and their nationality. Using the rule of the adjusted residuals greater than 2.0.Examination of adjusted residual indicates that the relationship is mostly created when the female participants from both India and Non-India responded "Yes " than expected(Indians Female:11 vs. 3.1931,adjusted residual =0.001,Non-Indians Female: 6 vs.5.854,adjusted residual=0.0007).In addition the participants were more tending towards the "No"(Indian Females:13 vs.2.701,adjusted residual=0.003,Non-Indian Female:5 vs.7.025,adjusted residual=0.0008) and fewer participants responded "Maybe"(Indian Female:6 vs.5.854,adjusted residual=0.002,Non-Indian Female:1 vs.35.125%,adjusted residual=0.004).

Thus, concluding that in relation with male to female ,more number of female are not ready with the approach of getting microchip implanted in their skin. Most of the Female participants expressed negative comments and unwillingness in mild manner. The comments included ," It's not advisable to implant any chip in one's body" and "I won't agree to it". When considering the positive perceptions relating to this techniques some of them stated," This will be of great benefit to our future and future generations" and "It's something new".

When considering the theme of RFID implant in terms of the positive perceptions and that to related with health ,one participants commented," ,"Its a good way to keep a check on health.".Some of the participants stated," Not much aware of".

Table 1



Q.1-Would you accept Microchip Implantation Technology within your body		Nationality of Participants			
		Indians		Non-Indians	
		Male	Female	Male	Female
Yes	Count	10	11	4	6
	Expected Count% Q1	3.475%	3.1931%	8.687%	5.854%
	Adjusted Residual	-0.004	0.001	-0.001	0.0007
No	Count	11	13	5	5
	Expected Count% Q1	3.159%	2.701%	6.95%	7.025%
	Adjusted Residual	-0.003	0.003	-0.0008	0.0008
Maybe	Count	9	6	2	1
	Expected Count% Q1	3.861%	5.854%	17.3%	35.12%
	Adjusted Residual	-0.004	0.002	-0.005	0.004

D. Discussion

More than expected, the ratio of the male to female ,females were leading ahead with great number of margin and that to towards negative perception. In comparison of both nationality ,Non-Indians were more towards the denial of microchip implantation than Indians. When using this data received from the closed ended question lead to the qualitative findings, that participants frequently expressed and/or attached negative or not so aware about the meaning when describing their feelings towards this emerging technology .It came to surprise that the people residing outside India are in no mood to appreciate this technology or take the advantage against the health risk.

Thus, after discussion it is easy to conclude that people outside India are giving first priority to the Health over to convenience especially the Female participants with comparison to Male participants.

E.Conclusion

In conclusion ,it can be concluded that the demographics of participants as well as the psychographics in terms of the generation factors appear to affect towards the perceptions of human microchipping . One limitation to this study is psychographics of the participants.A second limitation is the amount of the time required to collect the data and understand the demographics of the participants.Third limitation was the religious beliefs that participants were forbid towards microchipping.

ACKNOWLEDGEMENT

This study is done in order to fulfil Dr.Wolk’s wishes to continue the investigate the social implication for such an emerging technology .

REFERENCES

[1] Pawel Rotter.”RFID Implants:Opportunities and challenges for identifying people”,IEEE Technology and Society Magazine,2008.

[2] Call for Papers,”www.katinamichael.com”.[Online]. Available:”https://www.katinamichael.com/call-for-papers/”

[3] Christine Perakslis,Katina Michael.”Indian Millennials:Are microchip implants a more secure technology for identification and access control?”,2012 IEEE Conference on Technology and Society in Asia(T&SA),2012.

[4] Microchip implant(human),Wikipedia.[Online].Available:”en.wikipedia.org”

[5] Journal of technology management & innovation v.3.n.3 Santiago,2008.Available:”scielo.conicyt.cl”.

PERSONALIZED LEARNING

Siddhi Thakkar[1], Rahul Rathod[2] and Smruti Nanavaty[3]
Student[1] and Co-ordinator[3], Usha Pravin Gandhi College of Arts, Science and Commerce
Student[2], M.Sc.Big Data Analytics, ST. Xaviers College

ABSTRACT

21st century is becoming the era of massive transformations. The field of education and knowledge has given users more ease of education, they can study any topic of their choice at any time and from any place of their choice over e-learning platforms. Despite of having a lot of flexibility of choice of topics, learners are at times not really satisfied with the e-learning platforms. The drop off rates of students from an online learning platform is becoming a major concern of educators round the globe. The digital era provides a lot of knowledge in various forms such as hardcopy, audio visual, etc. .This paper focuses on this problem to find out the reasons for drop outs in learning from a particular topic and also their preferences of learning mode.

INTRODUCTION

Online learning provides a lot of scope of unlimited learning. Students have a wide range and variety of topics to choose from as per their interests. There are several e-learning platforms that provide contents on similar topics in their own different way. Not only students, but also instructors who wish to share their knowledge and understanding of concepts to a wide range of students, they can choose to turn towards an e-learning platform where they can record and upload their lectures as instructors. E-learning platforms also provide a way to these instructors to become learners and learn new skills at their ease. This has accelerated the learning and knowledge reach for learners all around the globe.

However these online learning platforms are criticised in comparison to traditional classroom learning. Traditional classroom learning has more interactions as the learner instructor relationship is dynamically built as per each student’s level of understanding and interest in the topic being delivered. The instructor knows whether the learner is able to grasp the concept or not and accordingly adjusts its methods of delivering the lecture that automatically reduces the drop off rate of learners and increases the learning index of the learners. Researchers have an interest in knowing how to control these parameters with respect to an online learning platform. [1]

In this paper authors have worked around some parameters that gives an insight towards what the learner is more comfortable with while grasping concepts. Which platform is preferred more? Also which factors affect the selection and enrolment of a learner in an online course, and which factors affect the drop off rate of these learners. Analysis on the factors that affect the drop off of the learners can be used to provide a better understanding. The data collection is primary data collection. This data is collected from a Google form which was circulated via a link.

LITERATURE REVIEW

A. Learners Platform

Digital era brings up endless access to various learning resources of various learning platforms. Despite of this, many learners prefer going old school. They would prefer taking a traditional classroom as compared to an online classroom for various factors such as ease of communication with the instructor.[1] Learners can instantly ask questions and solve their problems.[2] On the other hand many people prefer going for online platform because it is more cost effective (some courses do not apply charges) also they provide certificates that hold a lot of value. A learner can go through the course at his or her own pace and choose the learning platform of its choice. There are many other factors due to which a learner who is used to traditional classrooms may opt for an online learning platform like the availability of classroom learning for the topics the learner wishes to enrol for. If there are no classrooms available and then it encourages the learner to select an online learning platform. It is also seen by the researchers that when learners are personally recommended about some courses they would usually give it a try and enrol in it. This is also another reason that a learner would choose an online learning platform.

B. Knowledge seeking material

Multiple learning resources are available, such as journals, books, videos, audios, soft copies (PDFs), online blogs etc. Hard copies have always been a preference, it is observed that because of its medical benefits for the reader’s eyes. Authors observed that experience is the best teacher, so when learners visualise the concepts they can understand better. [3] like audio visual medium. Soft copies are also a highly preferred choice as they can be read via any digital medium such as tablets, machines, etc. It becomes a more preferred choice when hard copies are unavailable and its alternative resource becomes a soft copy which is readily available over the internet. Digital era has opened more doors of communication for the world, anyone can connect with other people for doubt solving or simple finding answers or better explanation on different topics. The most preferred medium for these kind of communication are blogs. Knowledge seekers can now find individual blogs for a variety of topics that are discussed in depth. Professionals and experts also take part in these discussions so that the understanding of concepts become much clearer. To an extent, online learning platforms use this technique as a module for their e-learning courses popularly known as virtual chat rooms. Here everyone who has taken part in the course can talk to each other the online platform itself in isolation. This helps in solving doubts with the instructor and encourage in-house discussions as well.

C. Discontinuing a course.

When taking up a course learners tend to look after many things like the cost of the course, is the course available online or offline, if the course has received good reviews, what are the post course benefits? etc. When such parameters become a deciding factors along with time available with the learner, they may tend to choose an online learning platform. Observations helped to understand that many learners tend to leave or discontinue an online course.[4] Analysis showed that there might be some reasons like the learner finds the content of course irrelevant after taking up the course. Hence the user may quit or exit the course, with respect to such situations many online learning platforms have a refund scheme where if the courses is paid then the course fees is returned after a trial period in which they can go through the course. Also, if the learner is not interested in the topic after taking up the course, he or she may not go through the course quiet often and as a result leave the course. Many more factors like the instructor of the course and monotonicity affect the course continuation rate of learners. It is important that the learners are well known of the pre-requisites if any required for a course.

METHODOLOGY

A. Data Collection

Data is collected using a Google form which asks for the learner’s preferred educational platform between a traditional learning platform and online learning platform.[5] If they prefer an online learning platform then at what frequency. The learners were also asked about their preferred mode for grasping concepts. This helps us to understand more about the learner and we can provide them more with those preferred mode than other. Learners we also asked that if they choose to opt out of any course or choose to discontinue any online learning course then what could be the reasons for it. Sampling was performed over the data because taking the entire population into consideration was very unlikely. Random sampling was chosen, the data points were assigned random values and 50 random data values were picked up for analysis.

B. Age wise chart

The data collected gave us a variety of age groups ranging from under 18, 18 to 25, 26 to 30, and 31 & above. Authors saw which age groups choose an online learning platform the most. The data was sorted as per age groups and clusters were formed on the age groups. This gave us the most likely times an age group would select an online platform.

C. Hypothesis testing

Digital era has bought a lot of advances in education. Using the predictions received from the above method of analysing age groups preferred learning platform, hypothesis was formed. Null hypothesis (H0) stated that “People who prefer E-Learning material more usually prefer e-learning platform.” An alternate hypothesis (H1) was formed with respect to the null hypothesis that, “People who do not prefer e-learning material more would not prefer an e-learning platform.” Researchers used z test to find out whether to accept null hypothesis and reject alternative hypothesis or whether to reject null hypothesis and accept alternative hypothesis. A z-test is a statistical test used to determine whether two population means are different when the variances are known and the sample size is large. [6]

We used two parameters, one the preference of choosing an online learning platform and the other being the methods used as a learning tool. Taking a mean of these two parameters z test was performed at 95% confidence level. At 95% confidence level our alpha becomes 0.05 that is if the probabilistic value i.e. p value that is received performing a z test is less that 0.05 then we need to reject the null hypothesis and accept the alternative hypothesis. On the contrary if the p value is equal to or greater than alpha value i.e. 0.05 then accept the null hypothesis and reject the alternative hypothesis.

C. Drop off reasons from an e-learning course.

It has been observed that there are many times where a learner wishes like quitting or opting out of an online course after enrolling in it. A learner might be continuing a course only for the sake of completion. We tried predicting the most likely reasons due to which learners discontinue an online course.

D. Learning tools preferred.

When learners are taught from their preferred mode of learning then they are able to understand the concepts better and are engaged in the topics being taught. It is observed that these preferences are based on the learner's comfort of grasping knowledge. It was observed that inappropriate resources is one factor that affects the drop-out rate of learners. This can be solved by providing learners whatever learning tools or mode they are comfortable with.

E. Data Security Aspects.

In this era of digital platform connections one's personality can be known just by the name of the person. Digital era has given almost everyone an identity. But, this platform can expose threats at the same time. If a person's learning pattern is recognised or leaked then it can be misused as by competition of that person to target its vulnerability. Also the motive of the research is to focus more on knowledge and beyond the comparative world give a learner a platform to be able to recognise its interests and educate itself accordingly. For such a motive, authors of the research plan to proceed with steganography. A method via which the data cannot be easily decrypted without having complete knowledge about the encryption done.

Analysis

After the sample data was selected from our population pool, the data was cleaned and sampling points were set to 50 samples for making the analysis more precise.

The first thing that was observed while analysis and prediction was performed was that, young teens, that is the age group of 18 to 25 prefer using online e-learning platforms more than the elderly and below 18. It is certain that traditional classrooms have provided the extreme age groups more flexibility than the online platforms. This prediction can be seen in the graph below.

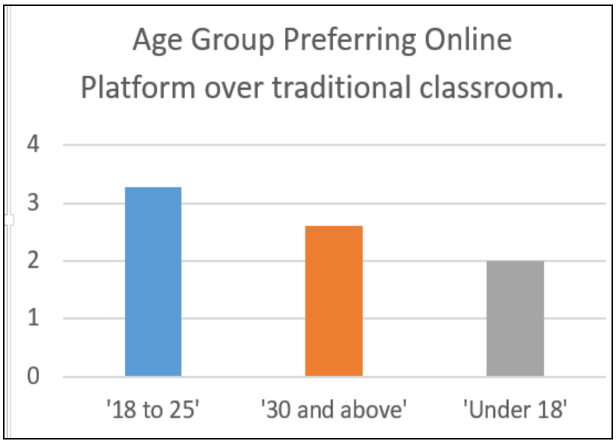
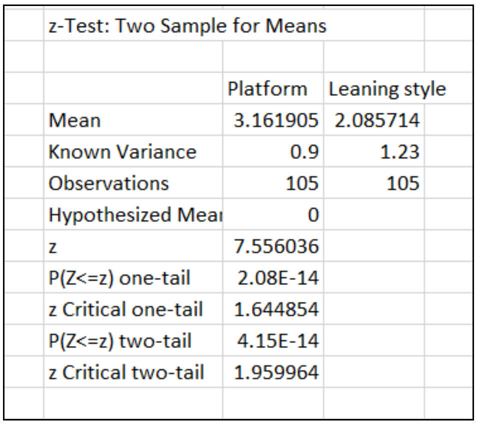


Figure 1: Age group preferring online platform over traditional classroom

The scale chosen is 0-4 indicating highest selection of online platform at 4 and least at 0. The next step taken was to prove our hypothesis. According to the data we collected, observations suggested that the Null hypothesis H0 would be, "People who prefer E-Learning material more usually prefer e-learning platform." The Alternate hypothesis for the same would be like, "People who do not prefer e-learning material more would not prefer an e-learning platform." As a result this hypothesis was made using Z-test over 2 ranges of input variables, the first being the choice of selecting an e-learning platform and the second being their preferred learning mode like, hard copies, soft copies, videos, blogs, etc.

After performing a z-test on the sample it was observed that the p value is less than the alpha value (0.05) talking about 95% confidence level. As the result it can be said that our null hypothesis, i.e. "People who prefer E-Learning material more usually prefer e-learning platform." is rejected and our alternate hypothesis, "People who do not prefer e-learning material more would not prefer an e-learning platform." is accepted. Z test is useful where the data size is over 30 data value points. It take two sample mean values and makes a comparison between them. A Z-test is any statistical test for which the distribution of the test statistic under the null hypothesis can be approximated by a normal distribution. Z-test tests the mean of a distribution in which we already know the population variance σ^2. [7]



z-Test: Two Sample for Means		
	Platform	Leaning style
Mean	3.161905	2.085714
Known Variance	0.9	1.23
Observations	105	105
Hypothesized Mea	0	
z	7.556036	
P(Z<=z) one-tail	2.08E-14	
z Critical one-tail	1.644854	
P(Z<=z) two-tail	4.15E-14	
z Critical two-tail	1.959964	

Figure 2: Hypothesis result after performing z test

In order to analyse the cause of people turning more towards traditional classrooms than e-learning portals despite of being a favourable choice we analysed the cause and reasons associated with these drop off rates. The following chart gives a graphical visualization.

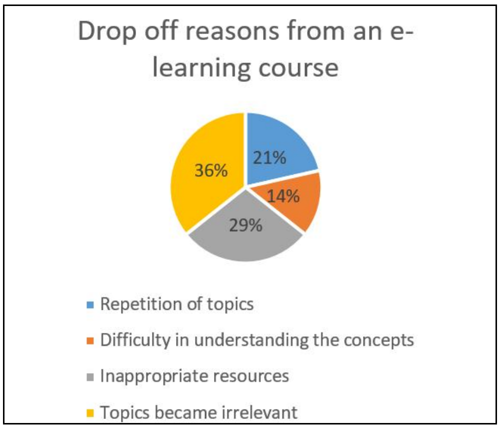


Figure 3: Drop off reasons from an e-learning course.

E-Learning portals may tend to go in depth but those topics may feel like a repetition to people who are quick learners. On the contrary people who tend to face difficulties in learning new topics may need continuous revision and in-depth exploration of topics. As a result it may cause a drop off from the e-learning platforms. Moreover it is also observed that the cause of inappropriate resources occurs when people who wish to find answers on the site itself may not get enough help and wander for solving queries as opposed to traditional learning. Also, when the interest of the user starts reducing from the topic, he/she may tend to find the topic boring and irrelevant and hence causing a drop from the e-learning scheme.

Forming clusters to be more specific about our target and analyse precisely about which gender prefers which learning platform the most, we can create separate clusters of Male and Female data and find out the ratio of their preferences. A visual representation is given below.

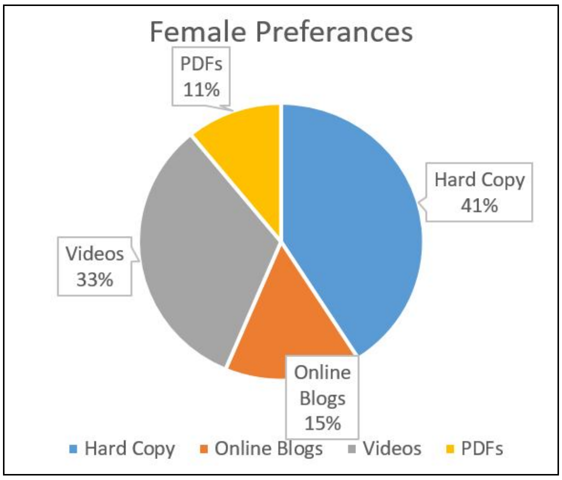


Figure 4: Learning resources female preferences

As observed, females prefer hardcopy learning material more over other resources.

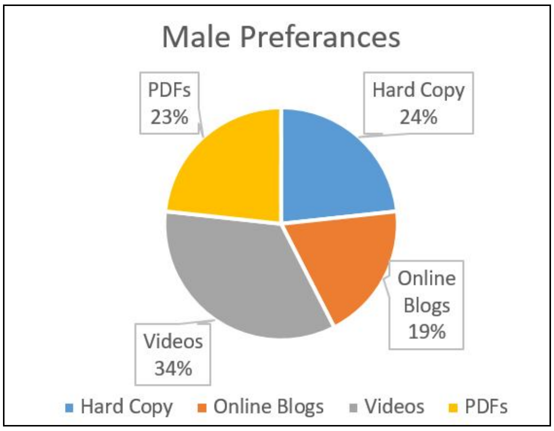


Figure 5: Learning resources male preferences

As observed, male prefer visual video learning resources more over any other resources.

These analysis have helped authors figure out various aspects that can affect a learner's learning.

CONCLUSION

A proposed solution for this can be to customise the e-learning platforms as per the preference, interest and learning styles of the users so that it can provide the users with much more opportunities of gaining the right knowledge.

If the learner is given a customised path where it can learn the concepts more precisely with respect to high interests towards a particular topic. This can make the e-learners learning be according to their wish and remove the irrelevancy factor that affects the discontinuing of the course.

Also, being able to provide learners the learning resources via which they understand better and are more suitable for grasping concepts. The more the learning resources are better the more the learners will be encourage towards continuing the course. Such affirmative responses can accelerate the use of online learning platforms that provide a vast pool of knowledge on various topics of interest suitable for every category of learner.

BIBLIOGRAPHY

[1] K, Tyler-Smith, "Early attrition among first time e-learners: A review of factors that contribute to drop-out, withdrawal and non- completion rates of adult learners undertaking eLearning programs."

[2] Weiyu Chen, Mung Chiang, "Early detection prediction of learning outcomes in online short- courses via learning behavior".

[3] Kew Si Na, Zaidatun Tasir, "Identifying at-risk students in online learning by analysis learning behaviour: A systematic review."

[4] Jui-Long Hong, Brett Shelton, "Improving Predictive Modeling for at-risk student identification: A multistage approach."

[5] https://docs.google.com/forms/d/ 1nqaa7kLe2YA3Tw2c0upKSyoz8sYxOMklnScGTAerylc/edit?usp=drive_web

[6] https://www.inssssvestopedia.com/terms/z/z-test.asp

[7] https://en.wikipedia.org/wiki/Z-test

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

A HOLISTIC STUDY ON EFFECTS OF VIDEO GAMING ON THE MENTAL HEALTH

Dhara Vara[1] and Prashant Chaudhary[2]
Student[1] and Guide[2], College Name: Usha Pravin Gandhi College of Arts, Science and Commerce

ABSTRACT

Video Gaming is an extremely used activity for relaxation with more than one billion users all over the world. In this years there has been a firm increment in the number of people and their increasing interest in Video Games. In this present research, we aimed to put some light on the link between online video gaming and gamer's mental health. It also focused on security issues in online gaming. Questionnaire Survey on psychological health as well as video gaming habits and other subjective tests were conducted with the individuals to know their perspectives. The responses were found more negative and less positive with respect to effects on mental health caused due to more time invested in games. Moreover, gamer's control over their playing time and whether they are aware of the adverse effects of excessive playing was observed through the survey. Our results are useful to develop a guideline as well as intervention strategy for the gamers to deal with video games and their playing time.

Keywords: cyber-attacks, behavioural addiction, video gaming

INTRODUCTION

Gaming disorder has become a significant issue in psychological health. While gaming is an important form of entertainment, excessive gaming is causing serious consequences for the players is broadly found overall through media and many games has increased slowly from past few decades due to emerging expansion in the gaming world. In this paper it was observed that frequently the people who has control over gaming spend time on playing games for Less than 2 hours and the other highest for 2-3 hours.. It includes issues such as loss of control, irritability, insomnia due to addiction, anxiety and other negative consequences of excessive gaming. The developed gaming world this days accompanies many security issues like cyber-attacks, leak of player's credentials and various other types of security breaching. Many software related fields are prone to cyber-attacks. But, the online video gaming world is always the target for the unauthorised practisers. From the beginning there were only single player games but after the online games came into expansion game applications put attention on transference of information and person-to-person interaction. When the networks are concerned for transmission of some data it brings a big amount of possible information security risks. It is still in process to find whether excessive video gaming should be considered a behavioural addiction or not. It was observed that people are more addicted to internet games to a higher extent as compared to basic offline games. Addiction of video gaming has found to be harmful not only to the mental health but also to the personality such as low self-esteem, aggression, sadness, fear of losing, low self-efficiency, etc. Spending half of the day in video gaming has its own potential consequences which are marked as the probability of player becoming socially awkward, losing interest in other relations like with family and friends, choosing to stay lonely. In this study it was also studied that players neglect other activities (studies/work) because of gaming.

LITERATURE REVIEW

Video Gaming is the most emerging source of entertainment to the people especially to the young generation. In recent studies it was found that the maximum video gaming players fall in the age criteria between 18-25 years of age. People are more addicted to games with video gaming console or computer. It can be played at home, your friend's place, and cyber cafe or even while walking down the street with a portable gaming device. It was learned by the experts that there will be no house where teens reside and there is no video game hardware or any other type of video gaming console. Online Video Gaming comes with lot of entertainment but it also carries some security issues with it which is taken into consideration by many game users. The issues mostly found were information-leak and cyber-attacks. This problems in any of the ways direct or indirect will harm the players. Common security threats found in online video gaming are:

Misusing Users Credential

Cyber offenders strive to gain profits by exploiting the gaming environment in several ways. Unauthorised practices like focusing gamers with illegitimate software and attempting phishing attacks to obtain game records, retrieval of credit card data, and other financial reserves. Plug-in softwares are used which makes the virtual characters perform in pre-designed way by designing them in such a way where they do not need the operation from the players which helps to gain virtual currency by beating the superiors and also retrieves the gamers credentials and forwards it to the cyber attackers.

Financial Loss and Interruption through Point Of Sale Attack
When the user wishes to purchase an online game, cyber attackers try to violate the essential source of credit card data during or after the payment is done by the user. There exists many gaming industries who run online gaming businesses which holds international clients and large numbers of international transactions which lead to fraud transactions. Attackers attempt to use Point of Sale also called as POS systems through great number of industries with special toolkits to obtain credit card numbers from game vendors. Gaming industries are advised to make sure that credit card details are to be kept tightly secured.

Other security issues can be Hacking issues through Trojans.
Artificial Hacking
Well-designed Phishing websites are created which looks alike the official websites of the games. The artificially created website will take player's account information along with the password. Later the hacker logs into the player's account with the acquired information to transfer virtual tools, virtual game currency, etc. and sell them offline for actual money.

This video gaming slowly becomes addictive in nature and begins to show its effects on player gradually. With Reference to [2] first, in order to understand video game addiction, we have to know how it all started. What do researchers say can cause a teen or an adult to become addicted to video games in the initial phase? It has been observed that video games can deprive a person from the actual world. Gamers spending more time on games isolate themselves from the other social contact, and pay attention almost entirely on gaming world achievements. According to the master's study files, various effects were also found such as sleep problems, depression, anxiety, fatigue, etc. Similarly in this paper study, we observed people suffering from irritability, insomnia, sadness, anxiety as mental effects on the players. It was also seen that video gaming habit normally affects males. If the person adjust with not touching the games or control to play there will be no other serious mental treatment required.

RESEARCH METHODOLOGY

In this section, we present the methods used to prove whether our Null Hypothesis or Alternate Hypothesis is accepted or rejected. For this we carried out a Questionnaire Survey using Google Form, to collect quantitative data through responses.

H0: Gaming does not affect mental health.

H1: Gaming affects mental health.

To prove our Hypothesis we kept a sampling frame of 100 people and the exact samples were collected from 94 people

Questionnaire used for survey:

1. How many hours per day you spend on playing video games?
2. Have you ever had to conceal or lie about the extent of your playing time?
3. What effect does playing video games have on your level of stress?
4. Have you found that playing video games increase (irritability, insomnia, anxiety, etc.?)
5. Have you neglected other activities (studies/work) because of gaming?
6. Do you often think of video games when you are away from your games console/PC/mobile phone?
7. Have you attempted to cut down your gaming time or even stop playing because you feel it is getting out of hand?
8. Overall, have you found that gaming highly affects mental health?

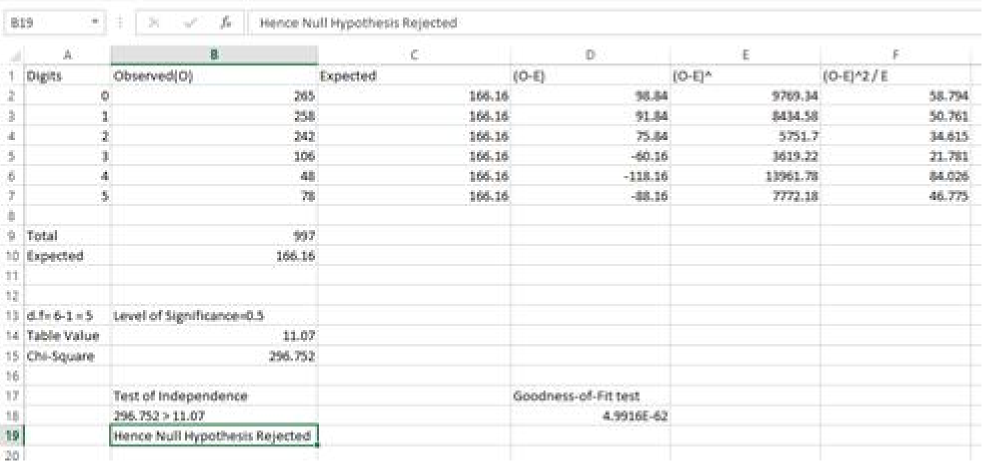
Result: On the basis of above collected data and to prove our hypothesis we used Chi-Square Test for Hypothesis Testing. The reason behind using Chi-Square test was that for it we required more than 50 samples and also it is a non-parametric test. We used Chi-Square Goodness of Fit method and Test of independence method to prove our hypothesis. The other tests allows you to say either "We can reject the null hypothesis of equal means at the 0.05 level". A Chi-Square test allows us to say either "We can reject the null hypothesis at 0.05 level" or "We have insufficient evidence to reject the null hypothesis at 0.05 level".

[4]What is Goodness-Of-Fit?

The goodness of fit test is a statistical hypothesis test to see how well sample data fit a distribution from a population with a normal distribution. Other way, this test shows if your sample data represents the data you would expect to find in the actual population or if it is somehow skewed. Goodness-of-Fit builds the variation between the observed values and those that would be expected of the model in a normal distribution case.

[5]What is Chi-Square Test of Independence?

The Chi-Square test of independence is used to establish if there is a significant relationship between two nominal (categorical) variables. The frequency of each category for one nominal is compared across the categories of the second nominal.



	A	B	C	D	E	F
1	Digits	Observed(O)	Expected	(O-E)	(O-E)^	(O-E)^2 / E
2	0	265	166.16	98.84	9769.34	58.794
3	1	258	166.16	91.84	8434.58	50.761
4	2	242	166.16	75.84	5751.7	34.615
5	3	106	166.16	-60.16	3619.22	21.781
6	4	48	166.16	-118.16	13961.78	84.026
7	5	78	166.16	-88.16	7772.18	46.775
8						
9	Total	997				
10	Expected	166.16				
11						
12						
13	d.f= 6-1 = 5	Level of Significance=0.5				
14	Table Value	11.07				
15	Chi-Square	296.752				
16						
17		Test of Independence		Goodness-of-Fit test		
18		296.752 > 11.07		4.9916E-62		
19		Hence Null Hypothesis Rejected				
20						

From the above data, Chi-Square value is seen to be calculated as = 296.752.

Chi-Square Tabulated value for 5 Degree of Freedom at 5% Level of Significance was 11.07. Since our Chi-Square Calculated value= 296.752, is much greater than the Tabulated value, it is highly significant and Null hypothesis is Rejected at 5% level of significance.

Hence, we conclude that H1 (Alternate Hypothesis is accepted) and it means Gaming Affects Mental Health.

CONCLUSION

In this paper we presented effects of gaming on mental health. It was studied that video games addictiveness can have various effects. Study show’s video games usage increase during 18 years or age but slowly decrease as the range of age rises. The current study adds to the knowledge on gaming by disclosing specific connections between video gaming and its gradual effects on person’s mental health. It also focused on security issues in online gaming world such as intruding virtual data, exploitation of users account, hacking, etc. which are found to be having no solution. In concern with gamers professionals suggest to provide them with more data encrypted system with secured identification technology which are less prone to threats. Potentially video gaming was found to be sometimes with positive effects as some people responded as gaming decreases their stress level but moreover it comes with psychological symptoms. Negative effects like low self-esteem, neglection of other activities, turning socially-awkward, keeping isolated, irritability, anxiety and other harmful effects. Gamers had to sometimes conceal about their excessive playing which shows they themselves were aware about the negative consequences later but could not stop themselves from playing. This shows the association between the video gaming and the psychological health of the player. Future studies may help to answer the question whether the link between video gaming and its effects on mental health and its functioning is moderated by gender, the reasons for playing, or the preferred game genre.

REFERENCES

[1] The Association between Video Gaming and Psychological Functioning https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6676913/

[2] http://albertgamboa.blogspot.com/p/literature-review.html

[3] https://www.performanta.com/resources/prevalent-threats-in-online-gaming/

[4] Goodness-Of-Fit https://www.investopedia.com/terms/g/goodness-of-fit.asp

[5] Test-of-Independence https://www.statisticssolutions.com/non-parametric-analysis-chi-square/

ACNE ASSESSMENT AND IT's INFLUENCE ON SELF-ESTEEM

Siddhi S. Kamath[1] and Smruti Nanavaty[2]
Student[1] and Co-ordinator[3], Usha Pravin Gandhi College of Arts, Science and Commerce

ABSTRACT

Adolescence is an important period for emotional and psychosocial development, and the self-consciousness is subject to substantial changes. Low self-esteem in adolescence has been associated with delinquency, aggression, antisocial behavior, impaired physical and emotional wellbeing and poorer adult economic success compared to high self-esteem individuals. This research focuses on how in this patient population acne, a disease process that affects more than 85 per cent of adolescents, influences self-esteem.

Keywords: Assess, Self-Esteem, Acne, Adolescence

I. INTRODUCTION

Symptoms and severity of the skin disorders vary greatly. They may be temporary or permanent, and painless or painful. Some of them have causes of the disease and others may be inherited. Many skin conditions are mild, and others may pose life-threatening conditions. The various types of skin disorders are acne, coldsores, blister, hives, carbules, ezema, psoriasis, cellulitis, contact dermitis, vitiligo. Many chronic skin conditions are present from birth, while others suddenly appear later in life. It is not always clear what causes these conditions. Many permanent skin disorders have effective treatments which allow for extended remission periods.

Acne is a chronic condition of inflammatory skin that causes spots and pimples, especially on the face, shoulders, back, neck, chest and upper arms. Acne is a long-term skin disease that occurs when follicles are clogged with dead skin cells and skin oil. It has blackheads or whiteheads, pimples, oily skin and potential scarring. The resulting presence may lead to anxiety, reduced self-esteem and, in extreme cases, depression or suicidal thoughts.

Adolescence is a complex life-cycle stage characterized by many striking biological, physical, psychologic, and interpersonal changes. This era is usually conceptualized as consisting of 3 phases. (1) Early adolescence, between the ages of 10 and 13, coincides with the onset of puberty and with many shifts in school life. (2) In middle adolescence, between the ages of 14 and 16, there may be increased evidence of changes in parent-adolescent relationships and often increased disputes in those family relationships, along with the commitment by the adolescent to his or her growing individual autonomy. It is also in mid-adolescence that peer relationships and romantic relationships are becoming more prevalent. (3) In late adolescence, between the ages of 17 and 21, teenagers often make important life-oriented decisions — committed lifelong relationships, career / work and continuing education.

The degree of happiness is an important indicator of results in life. Low self-esteem is also a potential risk factor for teenage depression, and is associated with increased suicide. In comparison, high self-esteem is often linked to increased resilience through difficult times, closer interpersonal relationships, and increased feelings of security, all of which help a person achieve career, academic, and social success.

The goal of this research is to present the existing self-esteem and acne data to help us understand the underlying psychological issues impacting patients better. Reviewing strategies that can assist in delivering more comprehensive care to patients with low self-esteem.

II. LITERATURE SURVEY

A. Two health-related quality of life questionnaires were used in a cross-sectional study of 200 adolescents between the ages of 15 and 18 years: Children's Dermatology Life Quality Index ® (CDLQI) and Cardiff Acne Disability Index © (CADI)[6]. Acne was registered in 83 percent of children, 54 percent of men and 46 percent of women. The total CDLQI score ranged from 0 to 19 of a maximum of 30, with an average of 1.7 representing a 6 percent disability. Nine subjects scored between 5 and 9, indicating a moderate disability (17 to 30%), while three scored above 10, implying a serious impairment (> 33%). Similarly, the total CADI score ranged from 0 to 15 of 15, with an average of 1.9 (CI 0 to 1), indicating a 13 percent disability. Twelve subjects scored 5 to 9 (33 to 60% impairment), one scored 10 + (> 67% impairment) and one scored 15 (100% impairment) to the limit. The authors concluded that in some teenagers acne significantly affects QOL.

B. Pawin et al used a CADI adaptation to interview 1,566 French teenagers who phoned a youth help line from November 2004 to January 2005[7]. Callers were split into those who had acne at present, had it before, or never had it. The respondents included 79.4% females and 20.6% males between 10 and 37 years of age, with an average age of 16.1 years. 48% of patients with acne reported being affected in their daily lives, which varied with the average prevalence of acne (39% for mild acne, 52% for moderate acne, and 67% for severe acne; p<0.0001). The number of acne patients who felt lonely (58.2%) or anxious (56.5%) was greater than or equal to that of patients with more serious diseases such as diabetes, obesity, autism, psychiatric disorders and cystic fibrosis, further demonstrating the significant psychological impact of acne.

III. RESEARCH METHODOLOGY

A. Participants and study design

A environment is a location where the analysis is performed. The three typical research settings-natural, partially regulated and highly controlled. This study is conducted in natural setting seleted community of Mumbai.

A pilot experiment was conducted to assess feasibility, cost-adverse events, and a (statistical variability) is an attempt to predict an appropriate sample size and improve pre-results design analysis. In the urban community, pilot study was conducted with the sample size of 10 adolescents and the pilot study was successful in defining the problems encountered.

Sampling criteria refers to a list of chararterisctics essential for exclusion or inclusion in target population. In this study, the sample size were 100 adolescent girls and boys from a selected urban community.

I) Inclusion Criteria

- People who are willing to participate.
- People who are from urban community.
- People who come in the age group from 12-24 years.
- People who are suffering from acne.

II) Exclusion Criteria

- People not suffering from acne.

B. Questionnaire and Variables

A questionnaire survey was conducted including the following variables:

- Demograhic data such as Age, Gender, Monthly Household Income
- Self-Grading tool
- Self-Esteem Scale

The full questionnaire contained details about mental health, lifestyle and behaviour in healthcare and can be referred from the following https://forms.gle/tJWqUhoyiGjGrYfu7.[1]

C. Statistical Analysis

The statistical analysis was based on Z-Test. Data were analyzed with frequencies to check for the maximum likelihood estimate. The significance level was set at P less than 0.05, and confidence intervals of 95 per cent were determined, as shown in Fig. 1.

I) Z-Test Hypothesis

- A z-test is a statistical test to determine whether two population means are different when the sample size is large and the variances are defined.
- Use it to test hypotheses where the z-test fits a normal distribution.
- A z-statistic number, or z-score, is a number that represents the z-test result.
- Z-testing is closely linked to t-testing, but t-testing is best performed when the sample size of an experiment is high. T-tests always presume that the standard deviation is unknown while z-tests assume that variance is known.

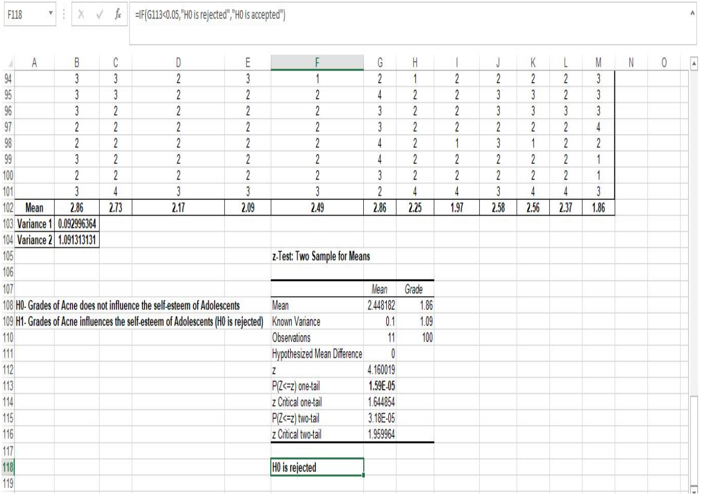
D. Results

Here, H0 is the null hypothesis stating Grades of Acne do not influence the self-esteem of adolescents, H1 is the hypothesis of this study stating Grades of Acne influences the self-esteem of adolescents.

For a hypothesis to be accepted the value of P(Z<=z) from the table should be <0.05 else the hypothesis is rejected.

ISSN 2394 - 7780

The test returns the value of P(Z<=z) one-tail as 1.59E-05 which is smaller than the set value of level of significance P=0.05 and thus H0 is rejected.



F118 =IF(G113<0.05,"H0 is rejected","H0 is accepted")

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
94		3	3	2	3	1	2	1	2	2	2	2	3		
95		3	3	2	2	2	4	2	2	3	3	2	3		
96		3	2	2	2	2	3	2	2	3	3	3	3		
97		2	2	2	2	2	3	2	2	2	2	2	4		
98		2	2	2	2	2	4	2	1	3	1	2	2		
99		3	2	2	2	2	4	2	2	2	2	2	1		
100		2	2	2	2	2	3	2	2	2	2	2	1		
101		3	4	3	3	3	2	4	4	3	4	4	3		
102	Mean	2.86	2.73	2.17	2.09	2.49	2.86	2.25	1.97	2.58	2.56	2.37	1.86		
103	Variance 1	0.092996364													
104	Variance 2	1.091313131													
105						z-Test: Two Sample for Means									
106															
107							*Mean*	*Grade*							
108	H0- Grades of Acne does not influence the self-esteem of Adolescents					Mean	2.448182	1.86							
109	H1- Grades of Acne influences the self-esteem of Adolescents (H0 is rejected)					Known Variance	0.1	1.09							
110						Observations	11	100							
111						Hypothesized Mean Difference	0								
112						z	4.160019								
113						P(Z<=z) one-tail	1.59E-05								
114						z Critical one-tail	1.644854								
115						P(Z<=z) two-tail	3.18E-05								
116						z Critical two-tail	1.959964								
117															
118						H0 is rejected									
119															

Fig. 1 Z-Test Hypothesis

As shown in the Fig. 1, the Z-test was performed on two sample means and variances of the populations Grades and Self-esteem scale.

IV. DATA ANALYSIS

This section provides a detailed insight to the results that are obtained from the experiment.

TABLE-I: DISTRIBUTION OF ADOLESCENTS ACCORDING TO THEIR AGE IN YEARS

Age	No.of Occurances	%
12-15 years	4	4%
16-19 years	30	30%
20-24 years	66	66%

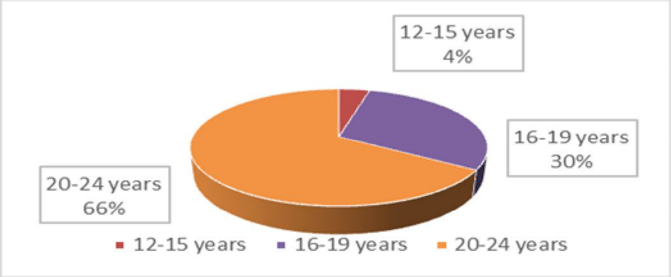


Fig. 2 Age wise distribution of individuals affected with acne

In the Fig. 2, 4% of adolesents from age group 12-15years, 30% belongs to age group 16-19years and 66% belong to 20-24 years are affected with acne.

TABLE-II: DISTRIBUTION OF GENDER

Gender	No.of Occurances	%
Male	44	44%
Female	56	56%

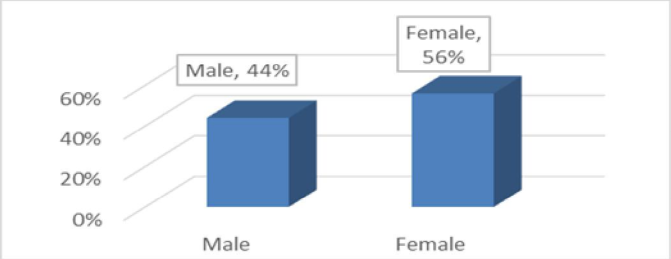


Fig-3:Gender distribution of individuals affected with acne

Fig 3. depicts that 44% males affected with acne and 56% females affected with acne.

ISSN 2394 - 7780

TABLE-IV: DISTRIBUTION OF TIMESPAN OF ACNE

Timespan of acne	No.of Occurances	%
Less than 1 year	51	51%
1-3 years	27	27%
More than 3 years	22	22%

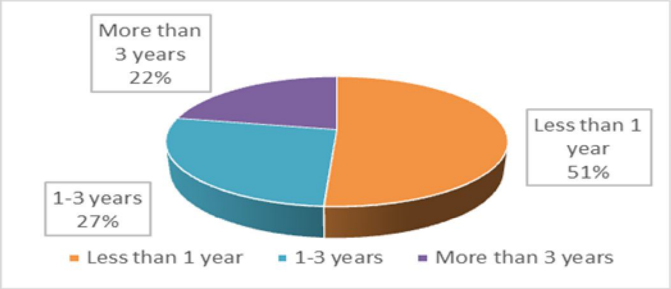


Fig. 5 Distribution of timespan of acne

Adolescents suffering from acne less than one year is 51%, 1-3 years are 27% and more than 3 years is 22%, as shown in Fig 5.

TABLE-V: DISTRIBUTION OF USE OF FACEWASH

Use of Facewash	No.of Occurances	%
Yes	83	83%
No	17	17%

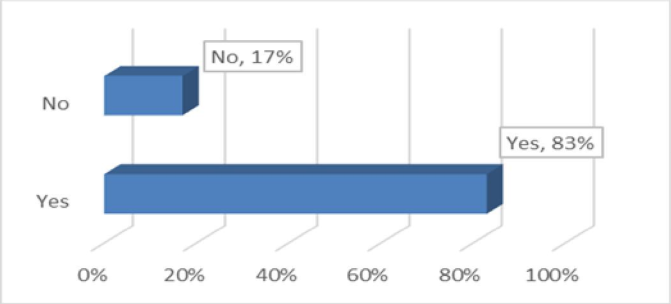


Fig. 6 (a) Distribution of use of Facewash

Fig 6 (a). shows that 83% adolecents use face wash and 17% do not use face wash.

TABLE-VI: DISTRIBUTION OF FREQUENCY OF FACEWASH USAGE

Frequency of use of Facewash	No.of Occurances	%
Daily	32	32%
Twice a day	32	32%
2-3 days a week	20	20%
Never	15	15%

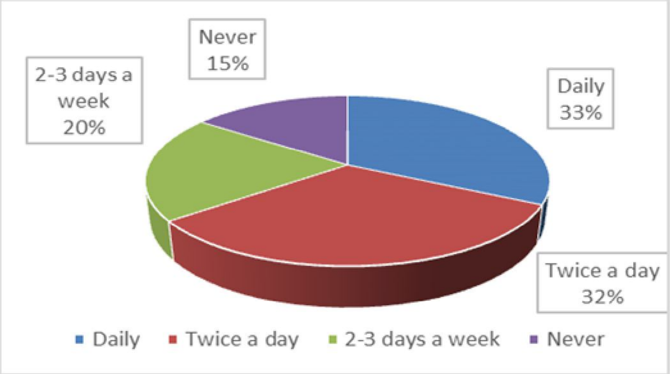


Fig. 6 (b) Distribution of frequency of facewash usage

Fig 6 (b). show 33% adolescents use facewash daily, 32% use twice a day, 20% use 2-3 days a week and 15% never use facewash.

TABLE-VII: DISTRIBUTION OF ADOLESCENTS SUFFERING FOLLOWING MEDICAL CONDITIONS

Medial conditions	No.of Occurances	%
Thyroid	12	12%
PCOD	9	9%
All of the above	1	1%
None of the above	77	77%

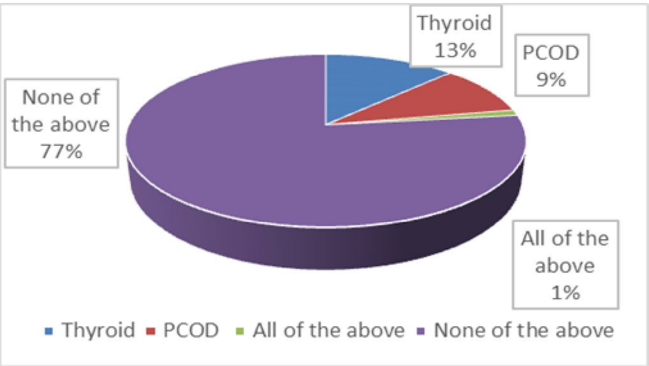


Fig-7: Distribution of Adolescents suffering medical conditions

Fig. 7 shows that 13% of adolescents suffer from Thyroid, 9% from PCOD (Poly Cystic Ovarian Disorder), 1% from All of the above and 77% from None of the above.

TABLE-VIII: DISTRIBUTION OF ADOLESCENTS USING MEDICINE

Medicine	No.of Occurances	%
Yes	18	18%
No	81	81%

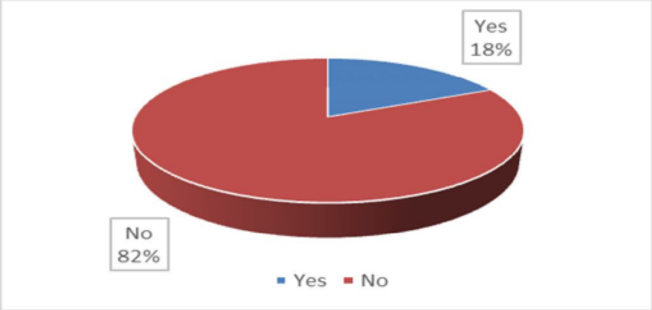


Fig-8: Distribution of Adolescents using Medicines for treating acne.

Fig. 8 shows that 18% take medication for acne and the remaning 82% do not cosume any medication for acne.

TABLE-IX: DISTRIBUTION OF ADOLESCENTS USING HOME REMEDIES

Home Remedies	No.of Occurances	%
Yes	58	58%
No	41	41%

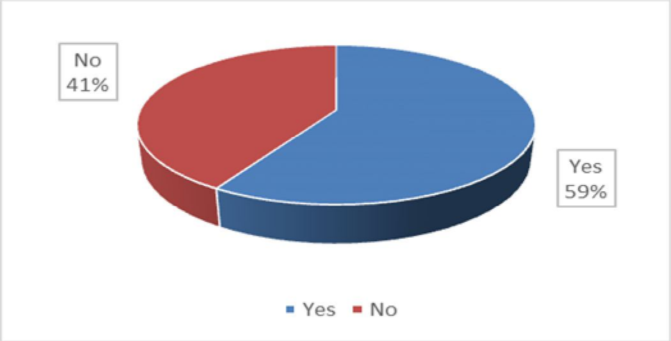


Fig-9: Distribution of Adolescents using Home Remedies for treating acne.

58% of adolescents use home remedies whereas 41% do not.

TABLE-X: DISTRIBUTION OF ACNE AFFECTED FAMILY MEMBERS

Family History	No.of Occurances	%
Yes	36	36%
No	64	64%

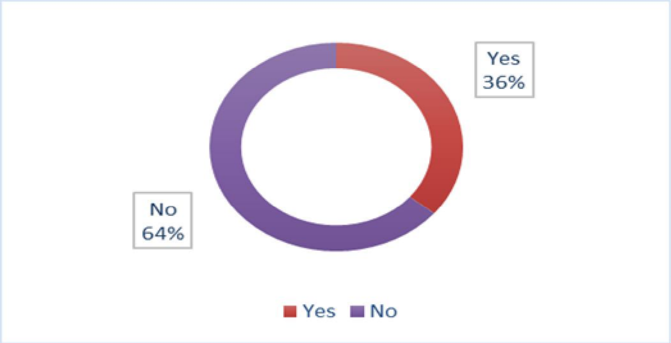


Fig-10: Family History- shows 36% of adolescents have a family history of acne whereas 64% of adolescents do not.

TABLE-XI: DISTRIBUTION OF ADOLESCENTS WITH DIFFERENT SKIN TYPES

Skin Type	No.of Occurances	%
Normal	22	22%
Oily	29	29%
Dry	13	13%
Combination skin	35	35%

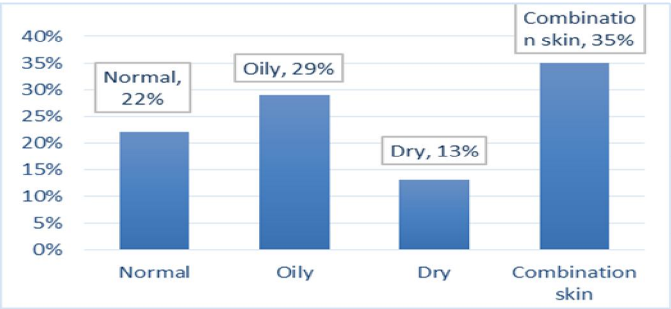


Fig. 11 Distribution of adolescents having different skin types

Fig. 11 shows the degrees of variations in the skin type of an individual. 35% of the adolescents have combination skin type, 29% oily, 22% normal and 13% dry skin type respectively.

TABLE-XII: DISTRIBUTION OF ADOLESCENTS ACCORDING TO THEIR GRADE OF ACNE

Grade	No.of Occurances	%
Grade 1	52	52%
Grade 2	20	20%
Grade 3	18	18%
Grade 4	10	10%

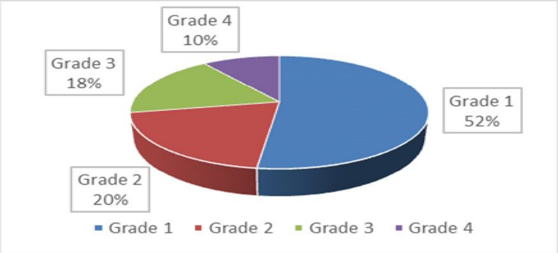


Fig-12: Distribution of Adolescents according to their grades of acne.

Fig. 12 shows us Grade 1 is the highest form of acne constituting 52% of the population. It includes black heads and white heads with small red boils and rashes.

TABLE-XIII: DISTRIBUTION OF ADOLESCENTS ACCORDING TO THEIR SELF-ESTEEM

Questions	Strongly Agree (%)	Agree (%)	Disagree (%)	Strongly Disagree (%)
I feel very anxious about my acne over these last few months/years.	22%	46%	28%	4%
I feel irritated when people remind me of acne.	18%	43%	33%	6%
I stay away from my friends and avoid social setting due to acne.	5%	18%	66%	11%
I think others avoid close contact with me such as hugging	2%	19%	65%	14%
I don't feel good about myself because of acne.	10%	40%	39%	11%
I'm satisfied with my appearance.	22%	44%	32%	2%
I feel my acne can never get better.	10%	21%	53%	16%
I use scarf to cover my face because of acne.	4%	11%	63%	22%
I use a lot of home remedies to treat my acne.	16%	36%	38%	10%
I use photo editing apps to cover my acne.	20%	30%	36%	14%
My facial features are neglected.	11%	27%	50%	12%

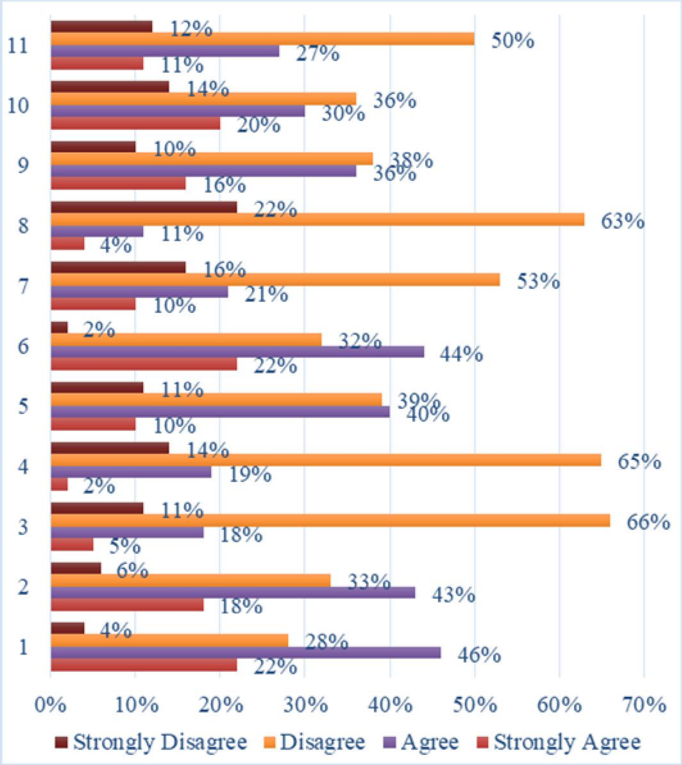


Fig-13: Distribution of Adolescents on their self-esteem

TABLE-XIV: AGGREGATED % OF ADOLESCENTS AFFECTED SELF-ESTEEM

Affected adolescents	No.of occurances	%
Low	1.4	13%
Moderately Low	3.35	30%
Fairely high	5.03	46%
High	1.22	11%

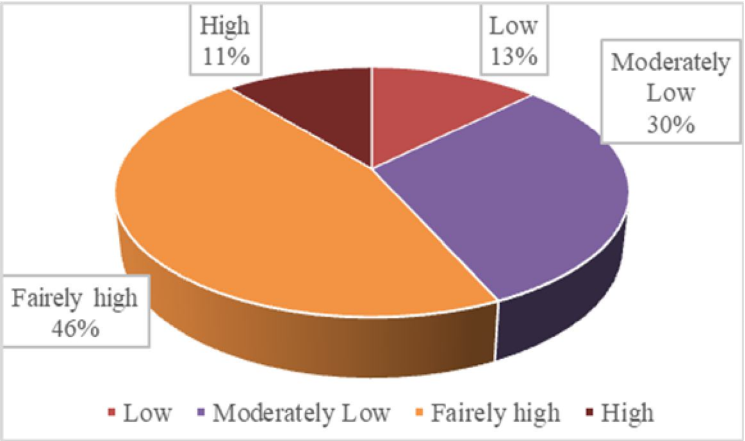


Fig-14: Dsitribution of affected self-esteem of adolescents.

Fig. 14 clearly shows us that 13% adolescents have low self-esteem, 30% have moderately low self-esteem, 46% have fairely high and 11% have high self-esteem.

V. CONCLUSION

In this community-based suvey the findings showed us that self-esteem of the adolescents were affected more by Grade 1 acne (including black heads and white heads with small red boils and rashes) by 52%. The study also concluded that adolescents within the age group 20-24 years were significantly affected by 66%.

To sum up, the research has identified the correlations of acne with the self-esteem of adolescents at a community level. Acne is complexly linked to the self-evaluation of adolescents and young adults and these interactions should be kept in mind when treating young adults with this disorder.

REFERENCES

[1] J Invest Dermatol. 2011 Feb;131(2):290-2. doi: 10.1038/jid.2010.375. Consequences of psychological distress in adolescents with acne.

[2] Misery L[1].University of Brest, Laboratory of Skin Neurobiology, Brest, France. laurent.misery@chu-brest.fr

[3] Curr Clin Pharmacol. 2018 Aug 21. doi:10.2174/1574884713666180821143545. [Epub ahead of print]

[4] Evaluation the effects of oral and topical simvastatin as adjunct therapy in the treatment of acne vulgaris.

[5] Ahmadvand , Yazdanfar , Yasrebifar , Mohammadi , Mahjoub , Mehrpooya MSchool of Pharmacy, Hamadan University of Medical Sciences, Department of clinical pharmacy. Iran.

[6] PLoS One. 2018 Sep 28;13 (9):e0205009. doi:10.1371/journal.pone.0205009. eCollection 2018. Stigma predicts health-related quality of life impairment, psychological distress, and somatic symptoms in acne sufferers.

[7] Davern J, O'Donnell AT Department of Psychology and Centre for Social Issues Research, Faculty of Education and Health Sciences, University of Limerick, Limerick, Republic of Ireland.ww.pubmed

[8] Puberty information [Internet].physical changes for girls and boys. Available from http//www.pamf.org

[9] https://www.aad.org/public/diseases/acne-and-rosacea/acne American academy of acne

[10] https://medlineplus.gov/magazine/issues/fall08/articles/fall08pg22-25.html

[11] https://www.dermnetnz.org/topics/psychological-effects-of-acne/

[12] https://www.mayoclinic.org/diseases-conditions/acne/symptoms-causes/syc-20368047

[13] www.medscape.com/index/list_6894_1 acne psychological effect

[14] https://www.jaad.org/article/S0190-9622(15)02614-6/fulltext

[15] https://journals.lww.com/jdnaonline/Abstract/2014/11000/Acne_Vulgaris__A_Review_of_Causes_and_Treatment.7.aspx

[16] https://www.webmd.com/skin-problems-and-treatments/acne/features/emotional-impact-acne#1

ISSN 2394 - 7780

A COMPARATIVE STUDY OF CURRENT TRENDING ONLINE STREAMING GIANTS (NETFLIX VS. AMAZON PRIME)

Mayur N Jadav[1] and Prashant Choudhary[2]
Student[1] and Guide[2], UshaPravin Gandhi College Arts, Science and commerce

ABSTRACT

Netflix and Amazon prime is a streaming service that offers a wide variety of award-winning TV shows, movies, anime, documentaries and more – on thousands of internet-connected devices.There's always something new to discover, and new TV shows and movies are added every week.We can watch anywhere, anytime, on an unlimited number of devices. Sign in with your Netflix or Amazon prime account.The Main purpose for this research is to check out the most preferable stream between Netflix and Amazon prime, on the basis of their features and benefit provided by them.Comparison between Netflix and amazon prime to know the best out of two. We found that people love Netflix in terms of adapted content, itsvariety and video quality, but there were few people who love Amazon Prime just because of low cost, less number of vulgarity and complimentary services.

Keywords: Complimentary services, Documentaries, Streamingservices.

INTRODUCTION

A decade ago on April 14th,1998 an American services provider named NETFLIX was launched with its primary business of subscription-Based streaming service which offers online streaming of a variety of Films and Televisions programs and series etc.Netflix expanded its business in 2010 with the introduction of streaming media while retaining the DVD and Blu-ray rental business. The company expanded internationally in 2010 with streaming available in Canada, followed by Latin America and Caribbean. And in 2016 Netflix plan to launch its service in India, the second-most populous country in the world.

Followed by launch of Netflixcame another online giant Amazon Prime who launched its services in 2005 include paid subscription services that give users access to services that would otherwise be unavailable or cost extra, to the typical Amazon customer. It includes free 2 day delivery, streaming music and video etc. And on 26thjuly, 2016 launched its globally popular Amazon Prime membership programme in India.

As of now the two online streaming giants already captured majority of shares in India. Not just that many other streaming services were launched including Hotstar, Zee5, SonyLiv and MXPlayer in India. Indian market started flooding with lots of streaming company trying to gain large pie from this growing market through launching variety of Series, shows and even various strategies to outfit their competitors. So through this research project will come to know which online streaming medium is the best among the two i.e Netflix and Amazon Prime.

LITERATURE REVIEW

With Reference to [1] conclusion; In 2016, two giants Netflix and Amazon have evolved in online streaming business. The similarities between the two are converging in terms of market access and online and offline access of the content. Amazon is chasing to acquire broadcasting rights of sports events as live sports telecast would be a game changer. However, both players will continue to strengthen their respective content offerings and competition between them may force viewers to subscribe to both.

[2]Gave an overview about Hulu vs. Netflix vs. Amazon Prime Binge-watching TV thenshows has develop a traditional phenomenon appreciationstoward services identical Netflix , Hulu, and Amazon Prime Movie . As for eachoriginal consumption behavioursremain to emerge, consequentlyconsume the old-style means of employing them. Sporting occasions, livingshows, then the news endure the support of TV networks, which numerous video facilities such by way of Netflix takenormally not replicated.Further services for example Hulu and YouTube have startedtowards offer live television streaming facilities in conjunction by their video facilities. The convenience of watching your favourite program on your individual leisure proceedingsomewhat device makes these “video streaming”facilities a mainstay meant forcurrentconsumers. Videofacilities such as per“Netflix, Hulu, and Amazon Prime Video” offer a big library of TV programs and cinemas that have previously aired or else premiered in cinemas, lettingclients to view programs they requiredbefore missed. Contractual agreements by major networks let these facilitiesallocateprevious and present seasons of standardTV programs. Likewise, main networks have designed their individual platforms meant for reaching customersover digital technology. (AT&T's HBO, for sample, offers HBO GO and HBO Now complete a

variety of distributionchannels)Presently,"Netflix, Hulu, also Amazon Prime Video worknow the similar industry by extensive libraries that frequently overlap. Though, the foundation, financials."

RESEARCH METHODOLOGY

In this section, we will discuss the methods we use to prove whether our NULL hypothesis or Alternate Hypothesis is accepted or rejected. For this we collected quantitative data through surveys i.e.Google form.

H0: There is no significant difference between consumers of Netflix vs. Amazon prime.

H1: There is a significant difference between Netflix vs. Amazon Prime i.e. Netflix is better than Amazon prime.

To prove our Hypothesis we kept a sampling frame of 80 people and the exact sampling we collected was from 70 people. For this, we use non-probability sampling method given below:-

1. Convenient sampling
2. Judgmental sampling
3. Snowball sampling.

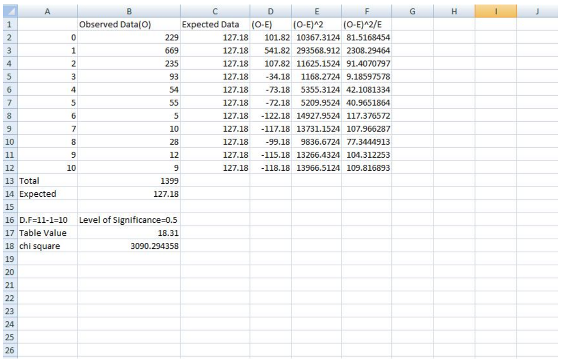
List of questionnaire used in my survey

1. How often do you watch videos on any online streaming medium?
2. Which one of the streaming medium provides you with what you are looking for in your watching wish list? (Ex-your favouriteshows, moviesetc).
3. Indicate which online streaming services you currently use.?
4. So now rate your experience with your current favourite streaming medium?
5. Which one provides more entertaining contents and adapting to local Indian culture well?
6. Which features you preferred the most in your favourite streaming medium?
7. Will you choose other streaming medium if they provide the same feature as your favourite streaming medium?
8. Which one is more costly than other?
9. Will you switch to other streaming medium, if other is less costly or more features available than your favourite one?
10. If yes, than what features are more entertaining that you won't switch to other?
11. Do your favourite online streaming mediums regularly provide new Content, Shows, Series & movies?
12. Do your current streaming mediums allow sharing your subscription with others for multiple screening in other devices?
13. If Yes, do you share with others?
14. Do your paid subscription comes with other complimentary services?
15. If No, will you prefer Amazon prime over other online streaming medium?
16. How much satisfied you are with current favourite streaming medium?
17. Do you prefer recommending your online streaming medium to others?

Result-On the basis of data collected above and to prove our hypothesis we used Chi-Square test for hypothesis testing. The reason behind using Chi-Square test was that for it we required more than 50 samples and also it's a non parametric test in we used goodness of fit method to prove our hypothesis. The other test allows you to say either "We can reject the null hypothesis of equal means at the 0.05nlevel". A Chi-Square test allows you to say either "we can reject the null hypothesis of no relationship at the 0.05 level" or "we have insufficient evidence to reject the null at the 0.05 level".

[3]What is Goodness-Of-Fit?

The goodness of fit test is a statistical hypothesis test to see how well sample data fit a distribution from a population with a normal distribution. Put differently, this test shows if your sample data represents the data you would expect to find in the actual population or if it is somehow skewed. Goodness-of-fit establishes the discrepancy between the observed values and those that would be expected of the model in a normal distribution case"



	A	B	C	D	E	F
1		Observed Data(O)	Expected Data	(O-E)	(O-E)^2	(O-E)^2/E
2	0	229	127.18	101.82	10367.3124	81.5168454
3	1	669	127.18	541.82	293568.912	2308.29464
4	2	235	127.18	107.82	11625.1524	91.4070797
5	3	93	127.18	-34.18	1168.2724	9.18597578
6	4	54	127.18	-73.18	5355.3124	42.1081334
7	5	55	127.18	-72.18	5209.9524	40.9651864
8	6	5	127.18	-122.18	14927.9524	117.376572
9	7	10	127.18	-117.18	13731.1524	107.966287
10	8	28	127.18	-99.18	9836.6724	77.3444913
11	9	12	127.18	-115.18	13266.4324	104.312253
12	10	9	127.18	-118.18	13966.5124	109.816893
13	Total	1399				
14	Expected	127.18				
15						
16	D.F=11-1=10	Level of Significance=0.5				
17	Table Value	18.31				
18	chi square	3090.294358				

From above data, our Chi-Square value=3090.294.

Tabulated value of χ^2 **test** for 10 Degree of Freedom at 5% level of significance is 18.31: So our calculated value of χ^2 **test** is 3090.295, it is highly significant and Null Hypothesis is rejected at 5% level of significant.

Hence we can conclude that H1(Alternate Hypothesis is accepted and its means there is significance difference Netflix vs. Amazon Prime i.e. Netflix is better than Amazon prime).

And also from survey we got that

1. Netflix is costlier than Amazon prime but still people chose Netflix over Amazon prime because it more entertaining and updated contents, series and movies and also we came to know that Netflix is better in adapting to local Indian culture.

2. And in India people preferred more variety of content rather more on availability of quality and price became the secondary reason that's why Netflix Loved by majority of people.

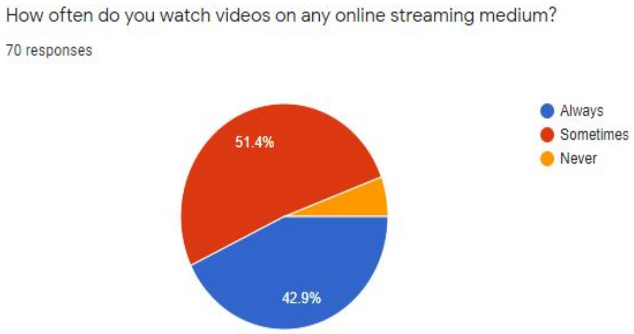
3. Other relevant thing to know was that people could easily switch to other medium if they provide better feature even if they are loyal to the own Preferred Streaming medium.

4. At present majority of Netflix customers are very satisfied and they believed in recommending it to other and spread positive word of mouth about Netflix.

DATA COLLECTION

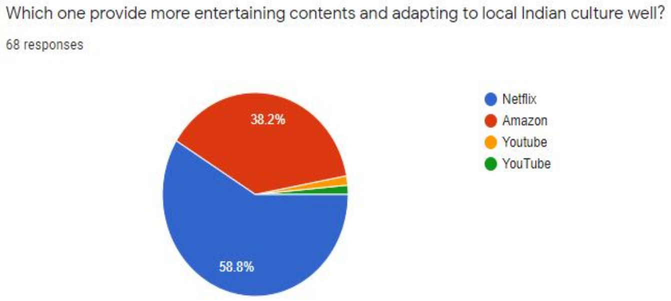
This data was collected through Google form and responses were collected from 70 people. There were total 17 questionnaire from which we selected this 10 most important question needed to prove our objective.

Detailed Description of questionnaire and responses collected are given below:



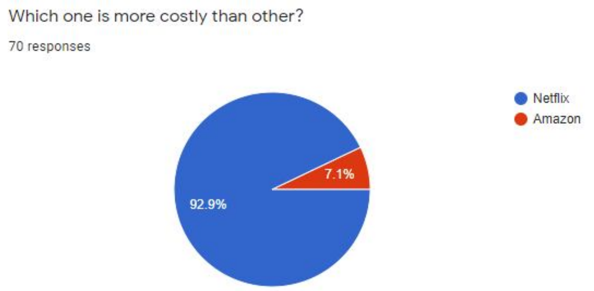
How often do you watch videos on any online streaming medium?

70 responses



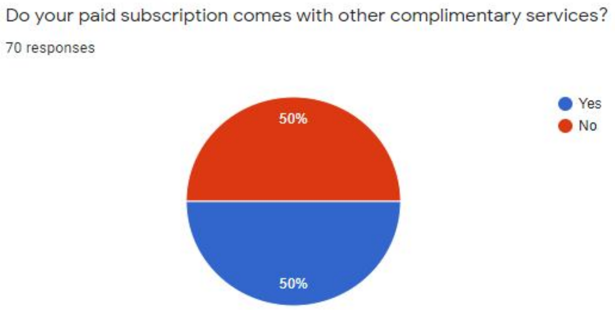
Which one provide more entertaining contents and adapting to local Indian culture well?

68 responses



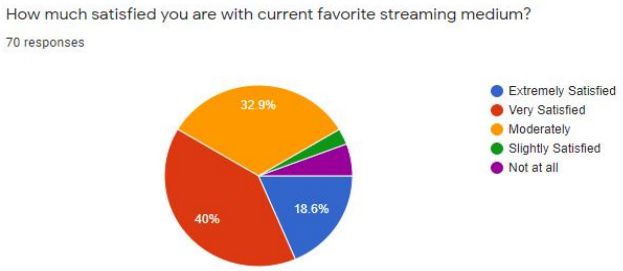
Which one is more costly than other?

70 responses



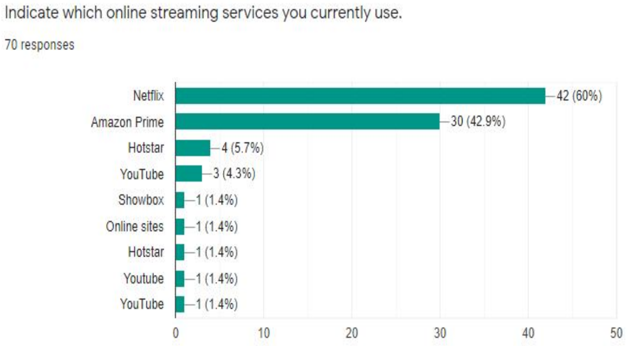
Do your paid subscription comes with other complimentary services?

70 responses



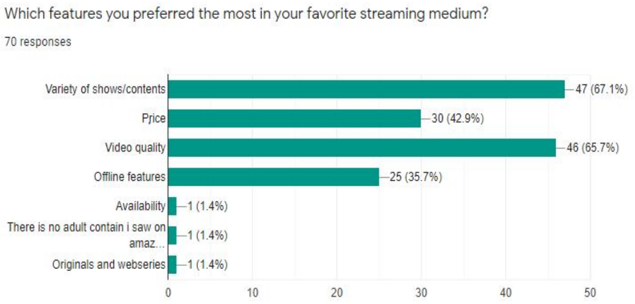
How much satisfied you are with current favorite streaming medium?

70 responses



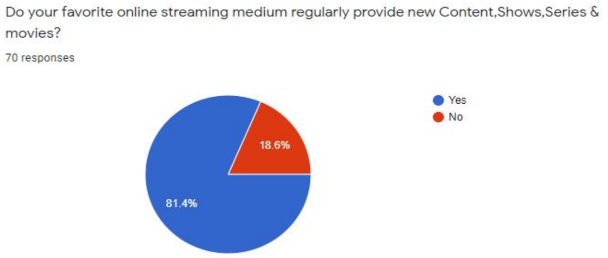
Indicate which online streaming services you currently use.

70 responses



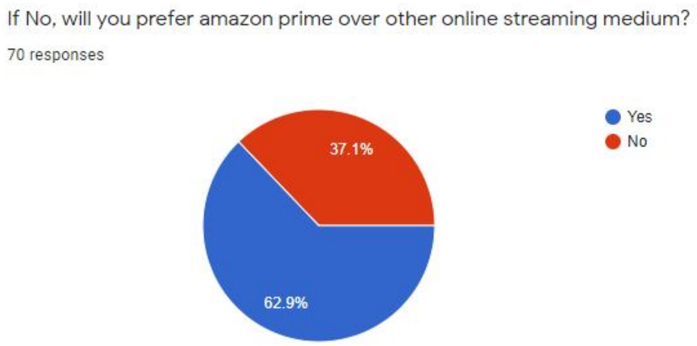
Which features you preferred the most in your favorite streaming medium?

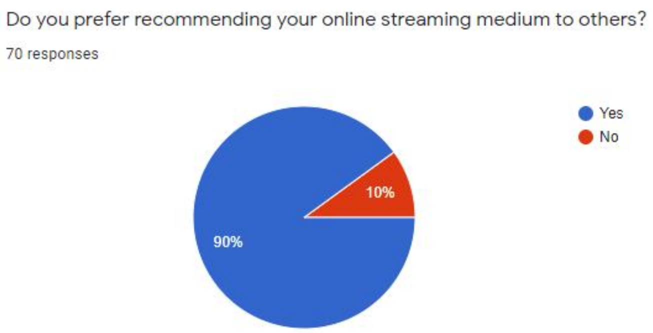
70 responses



Do your favorite online streaming medium regularly provide new Content,Shows,Series & movies?

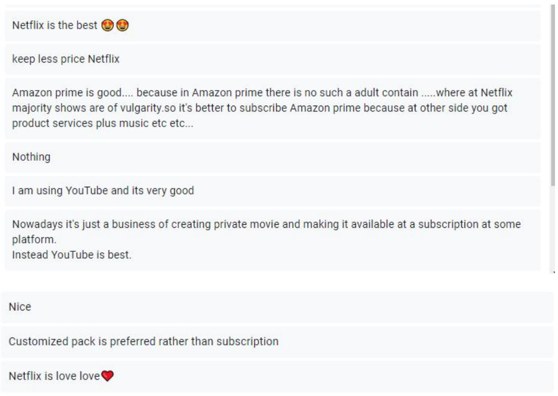
70 responses





CONCLUSION

In this paper we presented the significance difference Netflix and Amazon Prime that both are giving the best features to tackle their competitors and attract more customers but still difference in availability in new updated available content makes Netflix stronger than Amazon prime. But Amazon Prime‘s complementary services i.e. Prime video, Prime music, Faster delivery availability etc is constantly adding fuel in increasing its market share and value resulting in increasing number of customers giving tougher competition to Netflix in being online streaming giants. As competition is growing Netflix need to be updated and provide more complementary service to keep his position strong in market. We also came to know that there are many people who still love other mediums like Hotstar, YouTube etc. And finally want to conclude my research with some Suggestions/Feedback given by respondents in image given below:



Netflix is the best 😍😍

keep less price Netflix

Amazon prime is good.... because in Amazon prime there is no such a adult contain .....where at Netflix majority shows are of vulgarity.so it's better to subscribe Amazon prime because at other side you got product services plus music etc etc...

Nothing

I am using YouTube and its very good

Nowadays it's just a business of creating private movie and making it available at a subscription at some platform.
Instead YouTube is best.

Nice

Customized pack is preferred rather than subscription

Netflix is love love❤️

REFERENCE

[1] International Journal of Multidisciplinary Research and Development Online ISSN: 2349-4182, Print ISSN: 2349-5979 Impact Factor: RJIF 5.72 www.allsubjectjournal.com

[2] https://www.investopedia.com/articles/personal-finance/121714/hulu-netflix-and-amazon-instant-video-comparison.asp

[3] https://www.investopedia.com/terms/g/goodness-of-fit.asp

ONLINE SYSTEM IMPLEMENTATION IN INDIA: ASSESSMENT OF SECURITY ISSUES (A TRANCE ON BALLOT MORE COMPELLING THAN BULLET)

Bhakti Desai[1] and Dr. Manisha Divate[2]
Student[1] and Assistant Professor[2], Usha Pravin Gandhi College

ABSTRACT

This paper provides information about Online Voting system. There are major techniques of conducting elections. Each technique had its own some short comings. The existing system for casting a candidate's vote is LCD panel also called as first-past-the-post electoral system. Online Voting technique can be thought as an alternative to overcome the issues encountered by the candidates eligible for casting a vote. The aim of the proposed research is do a survey of willingness of population towards online voting and understanding the problems in implementing those in India.

I. INTRODUCTION

India is constitution with parliamentary system of government, and to decide the government holding free and fair elections are required. Elections plays a critical role for functioning of democracy. Elections helps the country to solve problem of accomplishing leadership and also to put the contribution of democracy[1]. For selecting a proper constitution every citizen of nation who fulfills the eligibility criteria is allowed to cast their vote for a party. The current voting system is LCD panel voting system where a candidate needs to visit to a polling station to cast there vote on a monitor located there.

There are 5 major Electoral Techniques which are:

1. Indelible ink
2. Electronic voting
3. NOTA
4. Absentee voting
5. Postal voting

And there is a practice going on to implement Online voting/Electoral system in India.[2]

II. LITERATURE SURVEY

The current election system which is first-past-the-post are conducted by choosing government schools and colleges as polling stations. The government employees are appointed to the polling stations to guide the voters. Electronic Voting Machines (EVMs) are used rather than the ballot boxes in order to save the frauds via booth capturing[3]. The voting day is the day when all the voters are asked to go to nearby polling station and cast their vote. The machines have names of candidates and party symbols standing for their respective party and voters are asked to select any one of their choice. After the time limit is ended the EVMs are scaled and within 2-3 days the votes are counted. And final results are announced on TV channels and other official social media sites.

Online Voting system is the technique used to cast a vote by just sitting at home by using software platforms used to safely conduct elections[4]. Online voting system provides a digital platform as it excludes the need to cast votes using paper or going out to polling station. This software application helps to save time, hold on to best practices, and experience internal claims and external managements, such as third-party vote administration needs. Online voting system is essential in India so that complexity, finances, human-resources and time consumption is eliminated. It will have several advantages as distinguished with the existing system such as: Accessibility, Cost Efficiency, Security, Confidentiality, Transparency and Accuracy[4].

In India in September 2010, Gujarat was the first state to use Online voting system technique maintained by Spain based firm Scytl. After that it was also used in municipal elections of April 2011. The advantages of this was fact that votes were non-traceable and authenticity of voters being well established.

III. RESEARCH METHODOLOGY

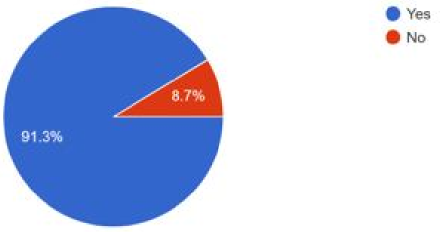
The null hypothesis to this research can be stated as **H0: No significant change in responses are observed before and after implementing secure integrated E-voting system.** People will opt for Online voting if its integrity is more secured. And **H1: There observed a significantly increasing trends for E-voting than LCD system.** People will not opt for Online voting even if its integrity is secured.

For this we conducted a Survey by Questionnaire method. We had designed a google Form which consist of ---- questions. Responses to those questions are our primary level quantitative data.

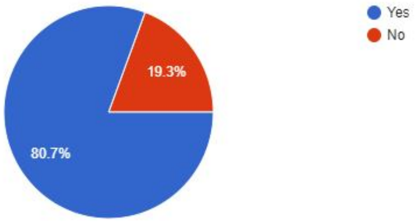
Based on random sampling the authors have carried out pivot study. About 150 samples are taken into consideration from a metropolitan city Mumbai.

Respondents were asked following questions:

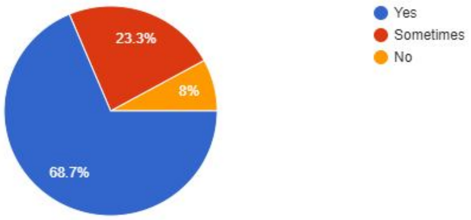
1. Do you have a valid voter-id issued by Govt. of India?



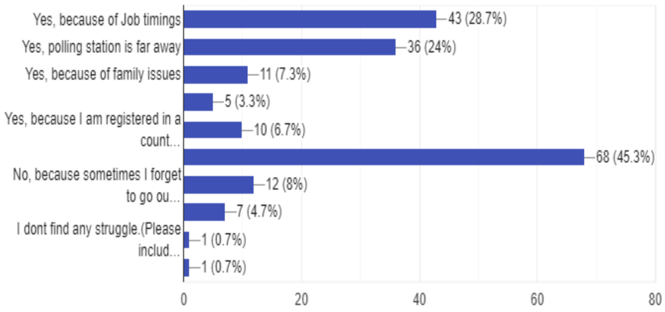
2. Are you registered to vote at the current address you reside at?



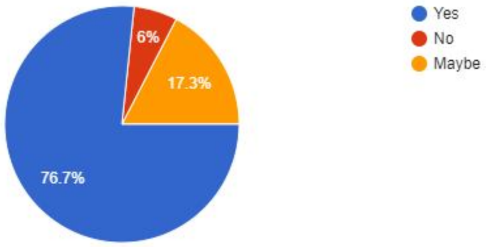
3. Do you vote?



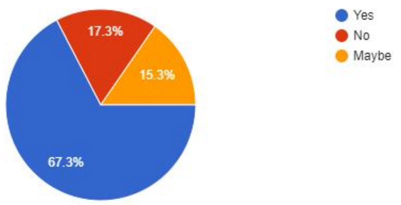
4. Do you struggle to find time to vote on polling day?



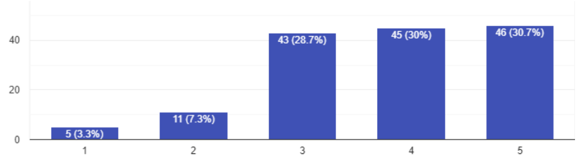
5. If Online Voting becomes a option, would you choose to use it?



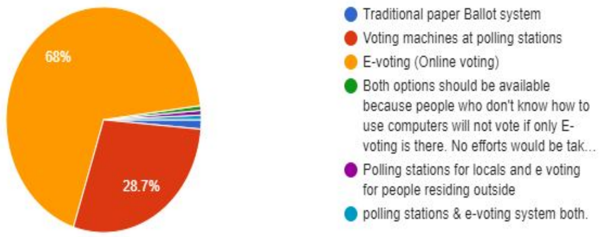
6. Do you think India should start Online Voting system?



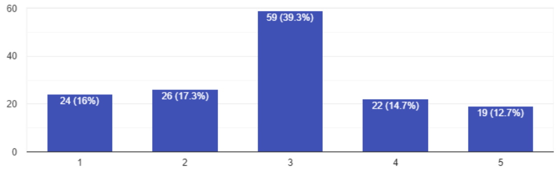
7. Do you think Online Voting would significantly increase election turnout?



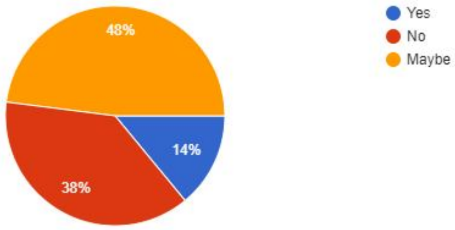
8. According to you which voting system is more efficient & feasible?



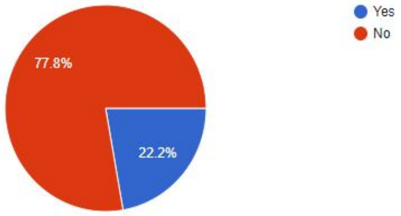
9. Do you think Online Voting system will be feasible to all age groups?



10. Do you think Online Voting system can be easily implemented in India?



11. Do you think Online Voting system will be secured in India?



As we can see that respondents are willing to use online voting system but are scared of their votes getting manipulated. So it gives rise to two different kinds of vulnerabilities which can be determined as:

1. Primary Vulnerabilities

2. Secondary Vulnerabilities [5]

Some respondents feels that implementation of Online voting system will be easy, flexible and feasible to use. Also it will be accessible from anywhere. One just needs to register themselves with their proper identity to the software application which makes the resources handy. And from the government point of view it becomes easy to calculate the votes as software itself calculates the responses.

While the other ones think that Online voting system will not be feasible to all age groups, and it is not suitable to use in rural areas as population there is illiterate or less privileged to use the application software.

Talking about the existing system, it has some shortfalls such as:

1. If the respondent is travelling for a business meet or is stating in hostel for future studies may not be at their hometown at the time of voting so their vote can be wasted.
2. If the respondent have moved from one city to another and is still registered to their previous residence address then their polling station will be at previous location which leads to not casting their votes.
3. Also, if the person is ill, hospitalized or due to some physical illness is not able to get down to the polling station, their votes can be wasted as well.
4. Respondent's less willingness to go to the local polling station, wait in a queue till their chance to cast their vote. Because it takes time to wait in queue. And are bit lazy to step down.
5. There are some respondents who don't even know where their polling station is located. This may also give a intention of not casting a vote.

IV. SECURITY ASPECTS IN ONLINE VOTING

Online voting systems are Internet based voting systems and are vulnerable to attack at three dominant fleck to be considered as the **Primary Vulnerability** might be faced in India:

1. Server side vulnerability
2. Client side vulnerability
3. Communication path interface

Diffusing attacks target the client or servers directly (here target client is voters and target server is the government officials) whereas denial-of-service (DOS) attacks target and discontinues the communication channel between both.

1. Client and Server voting (platform end):

Diffusing or penetration attacks associates the adoption of a transmission mechanism to include a nasty haul to the target host in the form of a Trojan horse or remote control program. Once carried out, it can spy on ballots, block voters from casting ballots and even worse, tweak the ballot bestow to its direction. What causes the terminal threat distinctly duplicitous is that it can be adept beyond revelation, and specified security mechanisms as encryption and authentication are ineffective counter to this kind of attack in that its target is beneath the level of contemplation at which those security protocols operate (browser). Virus and infraction revelation software is likewise reasonable to be incapable contrary to this threat because revelation mechanisms commonly inspects for familiar signatures of malicious programs or other clues of unauthorized action. These malicious attacks are commonly evoked from anonymous or altered programs, and amends system files to productively "authorize" the modifications built (posterior which they might harm added virus stability). The attacks could commence from everywhere in the world.

2. Communication Path

It speaks about the path between the client (device on which candidate votes) and server(site where votes are stored and calculated). For remote voting systems this procedure is required to be reliable and secured completely during the period from vote is casted till it gets counted. This depends on both an substantial connection link between client and server, along with the encryption of the data being transmitted to safeguard acquaintances. Preserving an authenticated connection linkage, despite can't be assured. Reasonably the remarkably important threat in this concern is a denial of service (DOS) attack, and that includes the usage of one or other added computers to disturb connection between a client and a server by overflowing the target with other extra requests that it can hold. This action adequately forbids the target machine from connecting up till the corresponding attack stops. In this aspect, an attacker can approach the bandwidth of other extra computers to overflow and destroy the planned target.

Alternate internet voting vulnerabilities i.e **Secondary Vulnerabilities** are essentially from :

1. Social engineering

In provision of voting system it is the most accepted term to regulate attacks that involves fraud voters into adjusting their security. Some people are not familiar with the use of software. They might not be alert of what they are doing while casting a vote. User interfaces (GUI) are usually indigent and also builds turbulence, comparatively reducing processes. A remote voting scheme will have some interface. For the system to be secure , there need to be some way for voters to understand that they are interacting with the election server. The better deliberate attack is probably by aiming the Internet's Domain Name Service (DNS). The DNS is practiced to preserve a mapping from IP addresses, which computers use to indicate each other to domain names, which people use to identify computers. The DNS is known to be sensitive to attacks, such as cache poisoning, which alters the information applicable to hosts about the IP addresses of computers[6].

2. Digital break down

Online voting brings the hidden aspect of a "digital break down", which may result in two approaches. Digital break down amongst those who have computers at home with Internet connection and those who do not have both. Secondly, amongst those who have quick access and those who have week connection and thus lower quality access. High Income group population can afford the software installation whereas Lower Income groups living in rural areas are not able to afford it. These dormant breaks perhaps be unsettled for participation and representation. Also the senior most population will restrict themselves from using the online services.

V. CONCLUSION

There should be Online voting system implementation in India. But along with the existing system. So that all the people can participate in voting. Online voting systems to be built with more security, integrity and transparency.

So that people gets an option to cast their vote while sitting at home or at work place and others if not able to cast their vote online can go to polling station and cast their vote. Although, there are no completely defensive techniques to confront these attacks, but as discussed there are certain precautions which can be used to avoid denial of service attacks.

So we conclude that both the system which is **Online voting system and LCD panel voting which is first-past-the-post electoral system should be used parallelly.**

VI. REFERENCES

[1] About Elections

- https://eci.gov.in/about/about-eci/the-functions-electoral-system-of-india-r2/www.slideshare.net

[2] Types of Electoral Techniques in India.

- https://en.wikipedia.org/wiki/Elections_in_India

[3] Existing System

- https://en.wikipedia.org/wiki/Electronic_votin_in_India

[4] About Online Voting System

- https://www.eballot.com/votes-and-elections/what-is-an-online-voting-system

[5] Security Aspects in Online Voting

- Issues in implementing of Online Voting System in India-Kanupriya Aggarwal

[6] About DNS services

- https://en.wikipedia.org/wiki/Domain_Name_System

DATA LAKE: A NEW PHILOSOPHY IN BIG DATA ERA

Shrikant Ghanshyam Pandey
Department of Information Technology, S. S. & L. S. Patkar College of Arts & Science, and V. P. Varde College of Commerce & Economics

ABSTRACT

As we know that the Data Lake is one of the arguable standards appeared in the era of big data. Data Lake original thinking is originated from enterprise discipline rather of tutorial field. As Data Lake is a newly conceived thinking with revolutionized concepts, it brings many challenges for its adoption. However, the doable to exchange the records landscape makes the lookup of Data Lake worthwhile.

Keywords: Big Statistics, Data Lake, Data Warehouse, Challenges, Miscellaneous

1. INTRODUCTION

In the technology of massive data, a new time period called "Data Lake" got here into view of the digital universe. The easiest intention of facts lake is to munge each and every records produced with the aid of an two organization two to supply more precious perception in finer granularity. Big Data applied sciences are from time to time regarded as damaging applied sciences as they revolutionized the common approaches of doing things in this information intensive era. Concepts from disbursed and parallel gadget are reapplied as two the two basis two of big two statistics such as two Map Reduce paradigms for managing the massive Vs characteristics – volume, velocity, variety, value and value. The incumbent SQL databases with ACID characteristics are challenged (and now and again even replaced) by NoSQL databases with BASE characteristics. Now, Data Lake thinking is attempting to task the reliable, traditional statistics warehouses for storing heterogeneous complicated data. The thinking of Data Lake was once two first initiated by Pentaho CEO James Dixon. If a facts warehouse or statistics mart is seem as a bottle of water cleaned and equipped for consumption, then "Data Lake" is whole lake of data which is cleaned for geared up use delivered extra in-depth definition for Data Lake as "a Data Lake stores disparate statistics whilst ignoring nearly everything". Some believe that new information architecture is required in the age of massive records as this computational intensive technology requests for new ideas and techniques for storing and manner changing and evolving data. All data generated by an company regardless of types, structures, or codecs will be saved in Hadoop clusters or other similar framework in their authentic forms. When parts of the organizations want to use the data, that stored information two will be loaded and modified as required through that business enterprise parts. Due to these, ideas in "Data Lake" appear to be challenging to the standard approaches of storing information i.e. statistics warehouse and information marts.

2. DATA LAKE NOTIONS

As Data Lake is a incredibly new concept, there are solely a few educational literatures which are centered to Data Lake. Data Lake as "a methodology enabled by a massive data repository based on low cost technologies that improves the capture, refinement, archival, and exploration of raw data within an enterprise." Data Lake may additionally contain raw, unstructured or multi-structured records where most part of these statistics might also have unrecognized fee for the organization. The basic concept of Data Lake is simple, all records emitted via the business enterprise will be saved in a single records structure referred to as Data Lake. two Data two will two be two saved two in the two lake in two their unique format. Complex pre-processing two and two transformation two of loading data into records warehouses two will two be two eliminated. two The two upfront two costs of statistics ingestion can additionally be reduced. Once information are placed in the lake, it's accessible for analysis via absolutely everyone in the agency suggested more specs for Data Lake mainly from the perspective of enterprise domain rather of lookup community.

All records are loaded from source systems.

1. No Data is grew to become away.

2. Data are saved at the leaf level in an untransformed or nearly untransformed state.

(implementation) and utilization are some distance extra popular in web articles and practitioner blogs than tutorial papers. Different facts lake concepts from the opinion about information lake is reviewed in from "Yesterday's unified storage is today's enterprise data lake" to "a massively scalable storage data repository that holds vast amount of raw data in its native format which is ingested by processing system (engine) without compromising the data structure."

DATA LANDSCAPE IN THIS ERA

When comes to core, there are only two operations in statistics processing – Transactional and Analytical. Daily operations such as Online Transactional Processing (OLTP) are usually work with CRUD-Create, Replicate, Update, Delete operations of the facts for every day routines. Data will be structured and saved in SQL databases. In big data era, not only structured information but additionally semi- structured and unstructured statistics will be stored in NoSQL databases. Nonetheless, facts in these databases will additionally be selected, cleaned, integrated, summarized, and modified according the shape of the statistics warehouse schema for the analytical purpose. Currently, statistics warehouses are the dominant method for presenting analytical data. Only transformed information will be stored in the data warehouse. However, as Data warehouses are very large and take time to create, "Data Marts" can be created. "Data Marts" are smaller two than data two warehouses, two and two supposed to store the information of a section of The business enterprise (i.e. a branch in the enterprise). Data warehouse will store data of the entire enterprise. These facts marts can be constructed separately. Or a phase of the information warehouse supposed for precise functionality or branch can be extracted to create a data mart. There are two famous definitions of Data Warehouse which come from their physical implementation. In frequent defined records warehouse is a subject-oriented (represent real-time object),In common approach, a top-down method for implementation, makes the creation two of records warehouse two as a first two step. two In his two approach, "data two mart", two focused for departmental use, in a two way a smaller element of records warehouse, is two created after two the institution of two the two foremost records warehouse. Kimball strategy is distinct as it encourages the creation of statistics mart. first. Then these departmental records marts are combined to create an organizational information warehouse. Therefore, Kimball strategy is known as bottom-up approach. Data warehouses work better with bit-map indexes, applied materialized view for better optimization.

DATA LAKE (Architectural Implementation)

"A data lake uses a flat architecture to store data in their raw format. Each data entity in the lake is associated with a unique identifier and a set of extended metadata, and consumers can use purpose-built schemas to query relevant data, which will result in a smaller set of data that can be analyzed to help answer a consumer's question". In quintessence, Data Lake is a information respiratory where all data in an agency i.e. structured, semi-structured, unstructured data + binary data are saved altogether regardless of types, format, or structure. The grasp of the records nature is delegated to the facts consumer at the time of information retrieval two (i.e. question time). When records are retrieved, person will transform that information in accordance to the parts of the agency to accumulate commercial enterprise insight.

SUGGESTIONS FOR DATA LAKE IMPROVEMENT

1. Align innovation initiative with company strategy. The precedence of the corporation approach are Business Acceleration, Operational Efficiency, Security and Risk. Data Lake implementation ought to focus on the key precedence of the corporate strategy.

2. Apply stable records integration strategy. The technological know-how for data integration might also be altering overtime in massive data. The first Data Lake solutions are based on Hadoop. To deal with real-time and streaming data, many DL options are now the use of Streaming framework such as Spark, Flank. Data Lakes want to preserve tune of evolving best-practices for metadata management. Data analytic pipeline need to automate the method of statistics extracting, loading, cleaning, transforming, and performing analytics.

3. Establish a modern-day onboarding strategy. Data Lake can get two crammed in batch load or in trickle feed. Simplify the procedure for loading data (regardless of types, sources or complexity) into the lake by means of enabling and setting up repeatable processes. Meanwhile preserve the suitable degree of records governance. Pay attention to the technique of on the fly metadata injection.

4. Embrace new facts administration techniques by way of adopting of early ingestion and adaptive execution processing such as MapReduce, Spark, or Flank that allow for flexibility.

5. Apply Machine studying algorithms to force real business Value.

III. DIFFERENCE BETWEEN DATA LAKE AND DATA WAREHOUSE

Both Data warehouse and Data information lake are statistics repositories. However, they are distinctive in many aspects from concepts, structures, and implementation. Data warehouses have well-defined regulatory features and storage capacity. In theory, Data Lakes have no restriction for storage capacity. two Any kind of records with any amount can be loaded into the records lake storage repository. "Data Lakes enable enterprises to look past the type and structure of data, giving them the chance to collect as much data as they desire" gives the records lake distinctions in distinction to statistics warehouse's processing of enormously structured data, pre-built layout earlier than question time, slowly altering statistics as follows. Rapid arrival of unstructured records volumes Use of dynamic analytical purposes (for query), Data turn out to be handy as quickly as it is created (as facts are changed based on question operation and software domain). The detail of differentiation between records lake and records warehouse is defined in subsequent sub sections.

COMPARISON FOR DATA LAKE AND DATA WAREHOUSE

Assessment between Data warehouse and Data Lake

	Data Warehouse	**Data Lake**
Data	Well-thought-out, administered data	Structured/semi- designed, formless data, fresh data, unrefined data
Processing	Schema-on-write	Schema-on-read
Storage	Lavish, reliable	Less Costing
Agility	Less nimble, fixed configuration	High agility, flexible configuration
Security	Developed	Growing
Users	Commercial professional	Data Scientists

IV. DESTRESS AND DEFIES IN DATA LAKE

- **Data Lake Concerns**

Marketing Hype: Adversary argument is that Data Lake two in reality is just a Hadoop's advertising and marketing hype of the Business Intelligence solution developers.

Data Swamp: Even for the supporter of statistics lake observed the pitfalls of Data Lake. One of the largest one is becoming a facts swamp. No one is aware of what will be put into the lake. Moreover, there is no strategies from stopping them such as entering wrong data, repeated data, or fallacious records. Data feed into the information lake do not guarantee the veracity considering the fact that their extraction.

V. ARGUMENT AND ASSUMPTIONS

Data Lakes try to remedy two issues – facts silos (old problem) and challenges imposed with the aid of large facts initiatives (new problem). Instead of having independently information collections, all information to be stored are accumulated in Data Lake to manage the historical silos problem. The new hassle is handling the challenges of large facts generation i.e. statistics lakes strive to resolve the challenges imposed with the aid of massive records V's characteristics – volume, velocity, verity, variety and value.

If information generated or produced from special departments within an company are stocked solely in their statistics stores, the probabilities of turning into facts silos are very likely. Data Lake tries to integrate information from these special stores in a single location trying to end the possibility of data silos. Traditional information warehouse with structured format can't take care of variety of information with exclusive latency need. In huge facts context, Data Lake might also be capable to handle volume and variety if aforementioned challenges have been handled. Data lakes are additionally growing in reputation due to the fact of IoTs (Internet of Things) increase However, presently Data Lakes are no longer threatening to replace facts warehouses as they have no longer treated the cited issues and challenges yet. additionally, offers very fascinating opinions of Data Lake. If the techniques had been located that will make ensure data lake options are well worth changing in location of warehouse, then the reply would be the evolution of the massive records warehouses. It also coincide the opinion for tableau large facts forecast. The forecast predicts that Data Warehouse and Data Lake concepts may additionally be mixed in close to future i.e. totally, Data Warehouse and Data Lake can grow to be solely one notion as soon as once more by means of improving and adding each other's capabilities. If Data Lake can efficiently deal with the challenges prompted by huge information and give up the issues of data silos, the complete landscape of facts storage architecture may additionally change once more in coming future.

VI. ACKNOWLEDGEMENTS

I would like to thank all the researcher, by referring their research paper and comments on an I am able to made this Review paper.

VII. REFERENCES

[1]. www.pwc.com/us/en/technology- forecast/2014/cloud-computing/assets/pdf/pwc-technology-forecast-data-lakes.pdf

[2]. https://www.blue- granite.com/blog/bid/402596/top-five-differences-between-data-lakes-and- data-warehouses.

[3]. http://www.pentaho.com/blog/5-keys-creating-killer-data-lake.

[4]. https://www.forbes.com/sites/ciocentral/2011/07/21/ big-data-requires-a- big-new architecture/#66609cb61157.

A REVIEW ON VARIOUS SOFT COMPUTING TECHNIQUES IN THE DOMAIN OF HANDWRITING RECOGNITION

Muzammil Aleem Khan
Department of Information Technology, Patkar Varde College, Mumbai

ABSTRACT

Handwriting identification is been one in all of the many fascinating and difficult analysis sectors within the field of image process and pattern identification within the last few years. Currently, the analysis work on handwriting recognition is focusing towards evolving new way of carrying out particular task that would scale back the process clock where as maintaining higher recognition accuracy. A review of various soft Computing techniques concerned within the method of written Words Recognition is represented during the paper. The work done by different researchers within the domain of handwriting recognition by victimization mathematical logic, Genetic Algorithms and Artificial neural Networks is additionally reviewed and the outcome is mentioned in terms of accuracy and speed

Keywords: OCR; Soft Computing Techniques; Artificial Neural Networks; Genetic Algorithms; Fuzzy Logic

INTRODUCTION

This advancement of AN automation method will improve the interface between man and machine in tons of applications. Finding out the address on envelope, signature verification, vehicle unique plate recognition, handwritten documents like historical documents square measure some of the applications of handwriting recognition. This is very complicated drawback involving varied problems like variability of someone’s writing overtime, totally different shapes and writing designs, isolated text or absolutely connected text and the similarity among characters etc. It’s tough to efficiently acknowledge the handwriting system having diversity within the input. The square measure state of affairs once we have imprecise or scant input and optimization is the that the measure needed for deciding the potency of handwriting recognition system. The structure of this paper is as follows: Introduction to Handwriting, soft computing techniques square measure delineate, short elaboration of ways in which to unravel handwriting reorganization problem victimization soft computing techniques and also the work already done to date during this field.

HANDWRITING RECOGNITION

The two major classification of handwriting recognition is online and offline within the offline recognition, the writing is typically captured optically by a scanner and the completed writing is obtainable as a picture. But, in the online system the 2D coordinates of sequential points square measure delineate as a perform of your time and therefore the order of stroke created by the author are on the market. The online methods are shown to the superior to their offline counter parts in recognizing written characters thanks to the temporal info on the market. The online methods are shown to be superior to their offline counterparts in recognizing written characters thanks to the temporal information the market with the previous.

Soft computing approach provides higher answer for handwriting recognition drawback than the ancient applied math approaches. The neural network conception are often use for the classification stage, which, is that the higher cognitive process apart of any recognition system and can use the options extracted in the previous stage. The neural network classifier are often used for handwriting recognition drawback. Fuzzy logic involves the popularity accuracy for the cases where ambiguous or general knowledge is employed. The concept of genetic formula additionally helps in providing best handwriting recognition system.

SOFT COMPUTING TECHNIQUES INVOLVED IN HANDWRITING RECOGNITION

Therefore considering all aspects of handwriting recognition, soft computing techniques emerged to be best way for resolution these kinds of issues. Soft computing could be a branch, in which, it is tried to make intelligent and wiser machine. Intelligence provide the facility to derive the answer and not merely arrive to the answer. The application of soft computing have verified 2 important advantages. Firstly, create resolution non-linear issues, in which maths models square measure unavailable/ impossible. Secondly, introduce human facts and statistics such as identification, resolution, understanding and others in the field of computing. This result in gaining the chance of developing intelligent systems like automatic self-turning system & machine controlled design system. The employment of sentimental computer techniques ends up in systems which have high machine IQ. It is the high MIQ of soft computing based mostly systems that accounts for the rapied growth within the variety and style of applications of soft computing one in all the necessary options of SC is acquisition of (information) or information from inaccurate and uncertain knowledge. It is expected that combination or fusion of the basic technologies can facilitate to beat the limitations of individual components. Soft Computing is outline as a group of techniques spanning several fields that comprise varied classes in machine intelligence. Soft Computing has 3 main branches: Fuzzy Systems, biological process computation, Artificial Neural Computing, the remainder of the section provides you the summary of these branches.

1. FUZZY LOGIC

Mathematical logic gives a easy path to achieve a perfect answer dependent on vagui, ambiguou, nossiy or missed info on the input. For this a mathematical expert system is form that make use of a group of mathematical membership functions. In point of boolean logic, to clearify about data. The procedures in a mathematical expert system is often of a type identical to the following:

If b is smaller and c greater the a= medium

The basic operations of fuzzy reasoning are: 1) to measure the input data with membership functions to get the membership value, 2) to mix the membership value to get huge strength, 3) to produce qualified rule on the basis of the huge strength.

2. GENETIC ALGORITM

GA is one of the low level & important algorithm class. GA solves both limited and unlimited optimized problem which is on the basis of natural selection. The GA over a period of time has modified a population of individual solutions. At every step, GA choose individual from any point among the present population to become parents and make use of them to make children of the coming generation. After many positive generations, population is evolved into optimum solution. GA could be use to find an answer of different types of optimization problems which does not go with the standard optimization algorithms. The three basic rules of Genetic Algorithm are:

➔ Selection Rule

➔ Crossover Rule

➔ Mutation Rule

3. ARTIFICIAL NEURAL NETWORK

Highly twisted structure, easy functional body is called Artificial Neurons. It has being made attempts to make a stimulation of a human mind. Neuron network can be termed as a studying networks which consist of a point in the network which are brought together with the help of adaptable weights, which store knowledge. These point in the network of brain are able to adjust to new conditions. They achieve knowledge via difference in a point in the network weight by becoming open to test-samples. Artificial neural network can be spread up on the base of:

➔ Bond among neurons.

➔ Operation function imposed to the neurons.

➔ Process of finding weights on connection

The Artificial Neural Network have better scope of being used in the parts such as Air traffic Control, Fraud Detection, Medical Diagnosis, Images and Finger. The hybrid method like neuro fuzzy or fuzzy neural hybrid are in use because it gets the positive result of neural networks as well as fuzzy logic system and it removes the individual disadvantages by combining them on the mutual features.

REVIEW OF RELATED WORK DONE

For the cause of clear picture in the handwriting recognition field, in this part a detail review of the research work which is already available is displayed. A lot of papers on related to off-line handwriting identification are been published. In their formal assessment, the authors have not estimate the experimental results of approaches reviewed, as that was felt that field was not properly or completely mature for this. However, the author comments that one of the most handwriting identification system is related to the type of features use.

Yoshimasa Kimura displayed a work on how to select features for Character Recognition using GA. The author proposed novel solution of feature selection for character recognition using GA. The given method select only the genes for which the recognition rate of training samples exceeds than the predetermined threshold as a candidate of the parent gene and adopts a reduction ratio in the number of features used for recognition as the fitness value. The handwriting recognition has emerged in the last 10 years and comes out to new impetus to new researchers as the technology increases. These days, this field can have become a big business for lot of industrial applications.

CONCLUSION

The numerous ways that and approaches used for handwriting recognition issues throughout the last decades are surveyed during this analysis paper. This paper is additionally focused on totally different soft computing techniques and use of those in handwriting recognition issues. This work proposes a model for applying Soft Computing techniques in the main neural networks or fuzzy pure mathematics or Genetic Algorithms or their combos for building intelligent handwriting recognition system.

REFERENCES

- Vincearelli, I Asurvey on offline cursive word recognitional, Pattern Recognition, 35(7), pp. 1433-1446, 2002.
- S. N. Srihari, I Automatic handwriting recognition-Encyclopedia of Language & Linguistics, 2nd Edition, Elsevier, 2009.
- ***L. Koerich, R. Sabourin and C. Y. Suen, - Large vocabulary off-line handwriting recognition. A survey, Pattern Anaylysis and Applications.***
- T. Steinherz E. Rivlin, and N. Intrator, Off-line cursive script word recognition-A survey, International Jounal of Deocument Analysis and Recognition.
- S. Mori, C. Y. Suen and K. Yamamoto, - Historical review of OCR research and development, Proc. IEEE, vol. 1050, July 1995.
- Avi-Itzhak H.I., Diep T.A. and Garland H., -High Accuracy Optical Character Recognition Using Neural Network with Centroid Dithering, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 17, no. 2, pp.218-223, 1995.

ISSN 2394 - 7780

USE OF MACHINE LEARNING TECHNIQUES IN DIGITAL FORENSIC FOR FILE TYPE CLASSIFICATION

Dr. Manisha Divate
Department of Information Technology, UPG College, Vile Parle, Mumbai

ABSTRACT

In digital forensic File type identification (FTI) is considered a fundamental problem. The time required to label file type is more as the disk size grows. The proposed paper gives the insight of FTI and approaches to solve FTI. The purpose of the study is to show how Machine Learning (ML) is a useful tool in FTI. It has been observed that file types JSON, PDF, txt, exe are rarely considered for FTI problem.

Keywords: Natural Language Processing, N-gram, Hamming weight, Histogram, Shannon Entropy

I INTRODUCTION

Advancement of digital technology made human beings to depend on the internet. A various task like online shopping, social networking, online banking, online study, cloud storage for digital data, etc. exposes the personal. Security to these data is a challenge. Unlawful accessing, tampering of the information is a crime. In India, the rate of cybercrime is increased by 70% in the year 2016 to 2017 [1]

Digital forensics (DF) is a branch in forensic science which deals with investigation and recovery of the digital material found digital devices [2]. Virus attacks, software piracy, phishing, data diddling, etc are the cybercrimes where the accession, alteration of the files is done. Identification of file type from the damaged storage devices such as disk, mobile, CDROMs is a challenging task. Activities conducted on computer can be trace by observing web browser’s history, cookies, deletion, accession of file, backup files [3].

The purpose of this paper is to study Machine Learning (ML) algorithms used to identify the type of file. The paper is organized as follows: Section II gives the survey of literature that has used ML techniques in file type identification (FTI) followed by result analysis of those ML techniques in section III. The paper is concluded with the findings and further trends in (FTI).

II LITERATURE SURVEY

In computer forensics, various branches such as disk forensics, storage forensics, printer forensics, database forensics are present. These specialized branches identify and investigate the crime in those areas [3]. Here a problem of FTI which is under disk forensics is discussed. Identifying the file type if the file is stored in continuous form is easy because of its extension and signature information. When a file gets deleted, it’s presence in the directory may be overwritten but actual contents are present on the storage devices. Now identifying the type of the file and carefully constructing a file based on the contents of file fragments is the challenge. First, understand the file system and how files are stored on storage devices.

A. What is file system?

In a computer, a file system is a logical system that stores and retrieves the files which are present on secondary storage. It is like a database or index which stores the address of the physical location of every single piece of data present on the storage device, such as hard disk, CD, DVD or a flash drive [4]. If the insufficient continuous space is available then file is stored in fragments. More the fragments more the disk seek is required. Table 1 shows the different types of file systems. There are three types of the techniques present in file type identification

1. Extension based

This is the fastest and easiest method of FTI. Every file has an extension of three letters which indicate the type of the file. All file system provides the extension to the file created using applications, examples such as database management system (uses db, frm, mdb, etc), office documents (doc, docx, docm, txt, rtf etc), computer-aided design (3mf, dim, cad, etc) store the information in image, text, video, speech form.

Table 1: File System used in computers

File System	Operating System	File system Space	File name length
NTFS	Window NT	16 Exabyte	255 char
FAT	MS DOS	-	9 char
FAT32	Windows 95	512 to 2TB	8.3 char
EXT	Linux	4TB	255 bytes
HFS+	z/OS IBM mainframe	2GB	31 byte

1. Signature-based

File signature is a sequence of byte which uniquely identifies the file type. The signature also termed as a magic number is of two-byte and present in the file's header part [5]. To identify the file type, the magic number, which is present in the header part of the file, is used. The magic number is predefined and built-in. Not all file types have magic numbers. In the case of file fragments, the use of the magic number to identify the file fails as all file fragments do not hold the magic number [6].

2. Content-based:

In this approach, the content of the file is thoroughly examined. Binary Frequency Distribution (BFD) is considered as a basic feature [7]. Other statistical measures such as unigram, bigram, Shannon entropy, hamming weights are used to detect the similarity between the known and unknown files. Figure 1 shows the sample BFD of .csv file. The content-based approach is first considered by (McDaniel & Heydari, 2003) for file type identification.

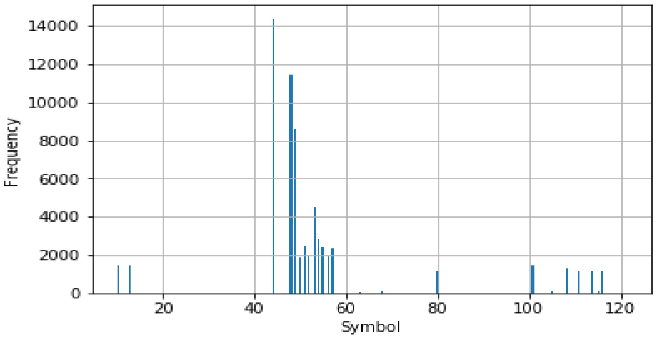


Figure 1: BFD of CSV file

Extension and Signature-based FTI is a traditional way of identifying the file type.

In the content-based analysis, FTI is done by assigning a pre-defined label (the file type) to each instance (each file) based on observed data in the file.

Here this research aims to identify how machine learning and NLP help in identifying the file type more accurately. It has been observed that many techniques are used for file type identification such as Shannon entropy, hamming weight, byte size fragments, NLP, histogram, etc.

In the content-based FTI, a bag of words model, a popular technique in Natural Language Processing (NLP), is used. A bag of word represents the group of consecutive words or tokens. A single word model is called a unigram model whereas two consecutive words model is a bigram. Bigram and trigram models capture more information (consecutive words provide more information) as compared to the unigram model which deals with a single word. The researcher found the n-gram model more effective [6]. Here file contents are in binary hence byte occurrences are considered.

Shannon Entropy is another most popular feature considered for FTI. Entropy is a measure of uncertainty. Higher is the entropy higher is the information. Encrypted files have more entropy compared to unencrypted ones [8], [9]. Shannon's entropy is computed as:

$$H(X) = -\sum_{i=1}^{n} \mathrm{p}(X_i).\log_2 \mathrm{p}(\mathrm{X_i})$$

Here X is vector of data, n is number of symbols in X.

Example1:

X=[1,1,0,1,0] here only two symbols are used,[0,1]. P(X1=0)=2/5 and p(X2=1)=3/5

Shannon entropy is computed as:

H(X)= - (2/5 * $\log_2(\frac{2}{5})$)-(3/5 * $\log_2(\frac{3}{5})$) = 0.970

Here the entropy of the data vector is 0.97. Similarly, the entropy information of byte is considered as a feature for analysis.

Histogram of unigram, bigram is also used to identify the file type. Here histogram of an unknown file is compared with the histogram of known file. This can be implemented by computing the distance between the two histogram values [10].

In the next session detail study of the machine learning classifier used for FTI is discussed.

III. MACHINE LEARNING TECHNIQUES FOR FTI

Classifier models maps the unknown files to a labeled files based on the features. As discussed in earlier section, features such as Shannon entropy, BFD etc are used for the file classification. The researchers had used various ML classifier models such as decision tree, support vector machine, artificial neural network, naïve bayes algorithm, are used to label the type of unknown files. It has been observed that the researchers had used different types of file[7], [11], [12]. Hence there results are not comparable. Table 2 shows various machine learning algorithms used in FTI.

IV. CONCLUSION

FTI is a fundamental problem in digital forensic. The literature survey concludes that the machine learning-based classification of the file type is more useful for labeling the file types. It is observed that most of the files are of the type image. Hiding of secret content in the image file (termed as steganography) is being termed as law enforcement and the media. Rare work on labeling file types such as JSON, TXT, EXE file type is observed.

V. REFERENCES

[1] “Cyber Crime Rate.” [Online]. Available: https://www.statista.com/statistics/875947/india-total-number-of-cyber-crimes/.

[2] “Digital forensics.” [Online]. Available: https://en.wikipedia.org/wiki/Digital_forensics. [Accessed: 22-Feb-2020].

[3] S. Krishnan, “Role and Impact of Digital Forensics in Cyber Crime Investigations,” *INROADS- An International Journal of Jaipur National University*, vol. 8, no. 1and2, p. 64, 2019.

[4] “Media & File System Forensics.” [Online]. Available: https://resources.infosecinstitute.com/category/computerforensics/introduction/areas-of-study/digital-forensics/media-file-system-forensics/#gref. [Accessed: 21-Feb-2020].

[5] J. D. Evensen, “Clustered File Type Identification,” 2015.

[6] W. J. Li, K. Wang, S. J. Stolfo, and B. Herzog, “Fileprints: Identifying file types by n-gram analysis,” *Proceedings from the 6th Annual IEEE System, Man and Cybernetics Information Assurance Workshop, SMC 2005*, vol. 2005, no. June, pp. 64–71, 2005.

[7] M. C. Amirani, M. Toorani, and S. Mihandoost, “Feature-based Type Identification of File Fragments,” *Security and Communication Networks*, vol. 6, no. 1, pp. 115–128, 2013.

[8] W. C. Calhoun and D. Coles, “Predicting the types of file fragments,” *Digital Investigation*, vol. 5, no. SUPPL., pp. 14–20, 2008.

[9] C. B. Paul, “Entropy based File type Identification and Partitioning,” no. June, 2017.

[10]A. (Stanford U. Duffy, “CarveML: application of machine learning to file fragment classification,” p. 4, 2014.

[11] L. Hiester, “File Fragment Classification Using Neural Networks with Lossless Representations Networks with Lossless Representations,” *Undergraduate Honors Theses*, pp. 1–32, 2018.

[12] I. Ahmed, K. Lhee, H. Shin, and M. Hong, “FAST CONTENT-BASED,” *Ifip International Federation For Information Processing*, pp. 65–75, 2010.

[13] K. Karampidis, E. Kavallieratou, and G. Papadourakis, “Comparison of Classification Algorithms for File Type Detection A Digital Forensics Perspective.,” *Polibits*, vol. 56, pp. 15–20, 2017.

[14] M. McDaniel and M. H. Heydari, “Content based file type detection algorithms,” *Proceedings of the 36th Annual Hawaii International Conference on System Sciences, HICSS 2003*, no. January 2003, 2003.

Table 2: Types of Machine learning Algorithm used in FTI

Algorithms used	Features used	Result	Type of the files	No of Files Used	Problems
[10] LDA, SVM, Multidimensional Naïve Bayes	257	SVM performs well with accuracy of 75.03%	28 different files	983 files for training, 973 files for testing	Inaccurate labeling of file. Giving more file type for testing than training model
[13] Almost explore all algorithms of ML such as Decision tree, Neural network, LDA, KNN, Regression, SVM	BFD	For all ML algorithm accuracy of a model is equal and above 99%	JPG, PNG, GIF, PDF	7359	Only JPEG, PNG, GIF, files are considered for the analysis. 3-STAGE PROCESS is used
[5] Unsupervised K-means clustering and supervised classification techniques are used. SVM, Naïve bayes, Decision tree, multiperceptron models are used.	BFD as features	Accuracy increases from 85.8% to 90.4% And classifier SVM model gives the accuracy of 97%	6 file BMP, JPG, PNG, MP3, GIF and TXT	6000 files (1000 each type)	File signature is not consider for the clustering ; Fragments are considered
[14] Frequency analysis; Frequency Header/Trailer algorithm is used.	BFD	93%	30 file type	120	File signature is considered File hearer/Trailer algorithm shows the accuracy of 95.83 % compare to Frequency analysis.
[8] Linear Discernment Algorithm (LDA)	Longest common subsequence	Average rate of correct prediction is 88.3%	JPG,BMP, GIF,PDF	100 files	No header file is consider Greater performance is achieved in Longest common subsequence

SECURITY CHALLENGES AND RISK ESTIMATION FOR IOT- BASED APPLICATIONS

Dr. Neelam Naik
Assistant Professor, Usha Pravin Gandhi College of Arts, Science and Commerce

ABSTRACT

Internet of Things based system provides interconnected network of resources, objects, people and intelligent services. The restrictions associated with such systems for low energy consumption, low cost, easy physical accessibility, and longevity of the resources, such systems are vulnerable to many vulnerable attacks at different layers of the environment. Security issues in IoT systems are difficult to handle because these systems are dynamic and heterogeneous in nature. In the present study, the risk associated with the user access to such systems is calculated based on the factors such as Resource Sensitivity, User Context, Risk History and Action Severity. The decision tree algorithm is used develop a rule-based system to deny or allow access to the users of the IoT based systems based on the calculated risk.

Keywords: risk control, vulnerabilities, perception layer

INTERNET OF THINGS

Through the Internet of Things (IoT) based ecosystem intelligent services connect resources together for the interaction purpose. These resources include physical and virtual objects, human beings, and services. Through the medium of the internet, these components interact with each other. They perform operations on the information generated in real life world. IoT systems provide boundless abilities that helps to make daily life easy and efficient. As per one of the studies [5] it is predicted that by 2020, ratio of connected devices to connected people will be 6:1. The major IoT application and domains are wearables, health care, smart cities, agriculture, Smart Home Applications and Industrial Automation systems. IoT system provides various benefits at the cost of the introduction to the several challenges specially related to the privacy and security.

IOT RELATED SECURITY CHALLENGES AND VULNERABILITIES

The computing devices associated with IoT ecosystem are called endpoints. Simple endpoints cover routers, sensors, switches and many more. Medium sized endpoints establish connection between communication links. They run robust processor and establish communication between simple endpoints and gateway endpoints.

Table 1: Endpoints vis-à-vis possible vulnerabilities

	3PV	PV	NV	GV	NCV
Simple Endpoints	Yes	No	No	No	Yes
Medium sized endpoints	Yes	Yes	Yes	No	Yes
Gateway Endpoints	Yes	Yes	Yes	Yes	Yes
Note: 3PV: People, Policy, Procedure vulnerabilities, PV: Platform Vulnerabilities, NV: Network Vulnerabilities, GV: Gateway Vulnerabilities, NCV: Non-cyber vulnerabilities					

The gateway endpoints are responsible for service provision, ecosystem management, various devices auto discovery authentication and security. [7]

The security challenges at these various end points arise due to following reasons. Low energy consumption endpoint may be required to achieve long battery life which can last till many years. Therefore, if any mechanism is installed for security through cryptography then it would be able to perform only simple cryptographic calculations. In most of the cases the cost of IoT devices as well as installation is limited. This results in low processing capabilities. Most of the physical endpoints are vulnerable to attach because they are easily accessible.

If the organization which is having IoT based system fails to install security proof policies then it is also considered as a risk factor. Again, untrained man power lead to threat to handle IoT ecosystem securely. Network vulnerabilities arise due to many devices are installed at multiple locations in the network. Again, these devices follow common protocol for the communication which may become vulnerable to attack. Gateways are the subsystems in the IoT ecosystem and they are also vulnerable because of availability of Gateway endpoints to various users of the system. Table 1 shows endpoints vis-à-vis possible vulnerabilities. [9]

IOT ARCHITECTURE AND SECURITY ASPECTS

The architecture of IoT environment based on the software perspective, consists of major four layers: application layer, middle ware layer, network layer and perception layer. Figure 1 shows the architecture of IoT environment.

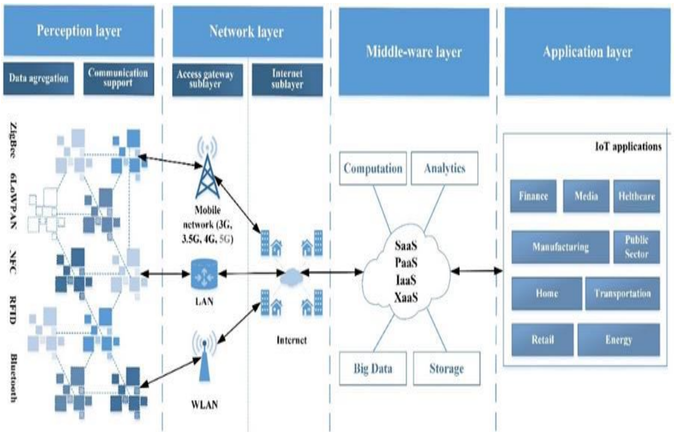


Figure 1: Architecture of IoT environment [3]

The perception layer is responsible for data collection and collaboration between the foundations of the same layer. The network layer consists of two sublayers. The access gateway sublayer which collects data from perception layer and communicate it to Internet layer. Internet sublayer as a backbone of IoT environment, transfers data to middleware layer. The intelligent routing and the network address translation are the main functionalities of this sublayer. Data collection, filtration, transformation and the intelligent processing are the most common functions of middleware layer. It is associated with cloud computing. The data received from middleware layer is presented by application layer in IoT ecosystem.[3]

Table 2: Architecture layers and risk levels

Layer	Risk level	Reason
Perception	High	Physical Characteristics of devices
		Price of device unit
		Physical exposure
		Energy requirements
		Wireless communication
		Implementation of security methods
		Heterogeneity
Network-Access sublayer	Low to medium	Application of wireless communication technologies
		Convergence of multiple users/devices at a single point
Network-Internet sublayer	Low	Routing
		Publicly exposed routers
Middleware	Medium	High penetration in the number of users
		Low level of maturity of the technology
		Ability to set a large number of user’s classes on a single physical machine
Application	High	In Finance, Media, entertainment domain
	Medium	In Home, Retail domain
	Low	In Healthcare, public sector domain

The vulnerabilities and threats are mostly present in the network layer because it is the backbone of IoT based system. Most common problems are eavesdropping, confidentiality breaches, integrity violation, DoS attacks and MitM attacks. In case of middleware layer as it is based on cloud computing, it offers flexibility and scalability in computing services. Table 2 shows various layers of IoT system architecture, levels of the risk associated and the reasons behind it.

One of the main aspects to handle security issues in the IoT is the user access control. Traditional access control method uses static and predefined policies to provide the access decision. The dynamic access control approaches use access policies and real-time information to determine whether to allow or deny the access.

Few of the features involved are operational need, history events, context, trust. The security risk is the output of the decision taken. For each user request to the IoT based application, security risk is calculated. It is then compared against predefined access policies to determine the access decision. Some of the other factors are: user's trustworthiness, data sensitivity, users and objects access history, type of access being requested and the location from which access is being requested.

This study provides a overall approach to assess security risks in access control operations of IoT system. The basic source of information to develop this model is the data collected through interviews of the experts in this field.

RELATED WORK

Diep et al. [4] have elaborated a process to calculate security risks of access control operations by computing risk associated with each access request and they compared estimated risk with the system acceptable risk value to decide access decisions. To estimate the risk value, Khambhammettu et al. [1] have elaborated approaches such as subject trustworthiness, object sensitivity, and difference between them. But this model does not explain evaluation of risk values in different situations quantitatively. The dynamic risk-based decision approach proposed by Shaikh et al. [8] uses previous actions of the users to distinguish between authentic and malicious users. But using reward and penalty points are not enough to determine accurate access decisions. The fuzzy inference system to measure the risk used by Chen et al. [6] have applied the fuzzy logic approach to design a fuzzy multi-level security model to provide access decisions. This model measures the risk related to the access request using the difference between object and subject security levels. If this difference was directly proportional to the associated risk. But this work does not elaborate how fuzzy rules have been generated. A risk estimation approach proposed by Arias-Cabarcos et al. [5] is based on a group of risk metrics that uses the fuzzy inference system to estimate the risk. But this model is lacked in considering real-time features while making the user access decision.

RESEARCH METHODOLOGY

The secondary data collected after interviewing various 20 experts from IoT securities has been collected by the authors Hany Atlam and Gary Wills for their research work. [2] The criteria used to select experts was basically related to the years of experience in security and knowledge of IoT applications. The data collected under the various risk calculating parameters such as User Context (UC), Resource Sensitivity (RS), Action Severity (AS) and Risk History (RH). These all four parameters are basically attributing user's surrounding environment like location and time, sensitivity of data accessed by user, severity of actions performed by the user and risk history of user. In the study performed by Atlam et. al. covers the application of fuzzy inference system on experts' knowledge. Fuzzy inference system was suitable to work on primary data because the data was in qualitative format, and it was subjective in nature.

In secondary data, by considering experts' responses, authors had used three fuzzy sets to represent each risk factor. The user context, action severity, and risk history were represented by "Low (L)", "Moderate (M)"and "High (H)" fuzzy sets. For the resource sensitivity, "Not Sensitive (NS)", "Sensitive(S)" and "Highly Sensitive (HS)" fuzzy sets were used.

The output risk is represented by using five fuzzy sets such as "Negligible (N)", "Low (L)", "Moderate (M)", "High (H)"and "Unacceptable High (UH)". The triangular MF method was used to represent the experts' gathered knowledge. The sample secondary data is shown in table 3.

Table 3: Secondary Data considered

Sr. No.	AS	RS	UC	RH	N	L	M	H	UH	Output
1	L	NS	L	L	20	0	0	0	0	N
2	M	NS	L	L	15	5	0	0	0	N
3	H	NS	L	L	6	5	9	0	0	L
4	L	S	L	L	11	6	1	2	0	L
5	M	S	L	L	1	5	12	5	0	M
6	H	S	L	L	0	5	10	5	0	M
7	L	HS	L	L	1	11	6	2	0	L
8	M	HS	L	L	0	4	13	2	1	M
9	H	HS	L	L	0	5	6	8	1	M
10	L	NS	M	L	18	2	0	0	0	N
11	M	NS	M	L	11	3	6	0	0	L
12	H	NS	M	L	7	1	12	0	0	L

The values of N, L, M, H and UH are based on the percentage of expert responses for the answer of particular combination values of the input factors. The output is generated by taking mean of these values.

PROPOSED RISK ESTIMATION METHOD

The information leakage occurs mainly because of security risk associated with IoT based application. This leakage results in partial or fully damage of the system. The method of risk-based access control calculates the risk associated with user access request. The proposed model uses user attributes related to the surrounding environment such as time and location, sensitivity of data to be accessed by the user, severity of actions that will be performed by the user and user risk history as inputs for the risk estimation algorithm to measure the risk value related to each access request to determine the access decision. Figure 2 shows the diagrammatical representation of proposed risk estimation method.

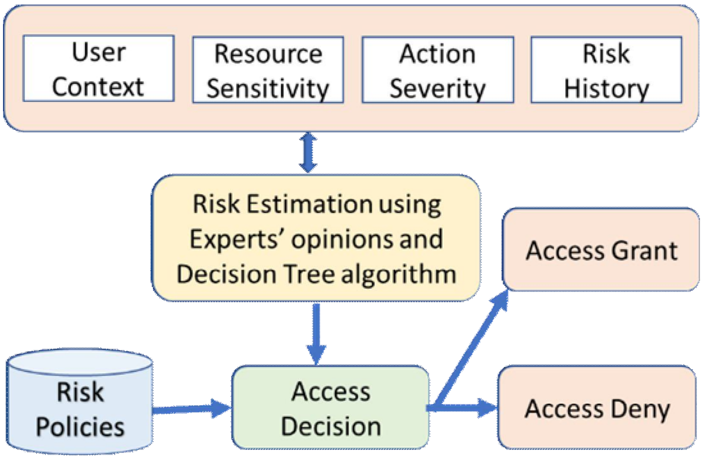


Figure 2: Proposed Risk Estimation Method

RESULTS

By considering the secondary data as mentioned in table 2, decision tree algorithm J48 that is c4.5 decision tree algorithm, is applied on it in Weka 3.8 software. Figure 3 shows decision tree generated. There is drastic change in number of rules generated from decision tree. In previous study the number of rules generated were 81 while in the current study it is 21. Thus, there is around 74% reduction in the number of rules.

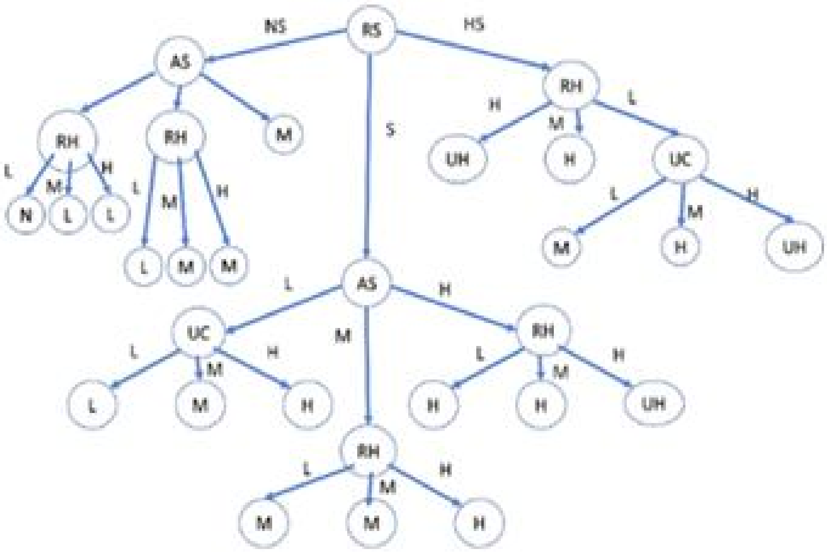


Figure 3: Decision tree showing rules generated

CONCLUSION

Using the decision tree-based classification model one can predict the level of the risk while allowing any user to access the IoT based system depending on the severity of the action, risk history, user context, and resource sensitivity. Over the fuzzy based system, the newly proposed decision tree-based model predicts the risk level more accurately. Even though understanding and estimation of the security risk changes based on the application context or the culture of organization, the results of the present study can be made generalized by considering a greater number of experts' opinions.

REFERENCES

[1] H. Khambhammettu , S. Boulares , K. Adi , L. Logrippo , A framework for risk assessment in access control systems, Comput. Secur. 39 (2013) 86–103

[2] Hany Atlam, Gary Wills, "An efficient security risk estimation technique for risk-based access control model for IoT", Internet of Things 6, Elsevier, 2019

[3] Ivan Cvitić, Miroslav Vujić, Siniša Husnjak, "Classification of Security Risks in The Iot Environment", 26th Daaam international symposium on intelligent manufacturing and automation, October 2015

[4] N.N. Diep, L.X. Hung, Y. Zhung , S. Lee , Y. Lee , H. Lee , Enforcing access control using risk assessment, in: Fourth European Conference on Universal Multiservice Networks, 2017, pp. 419–424 .

[5] P. Arias-Cabarcos , F.A. Rez-Mendoza , A. Marín-López , D. Díaz-Sánchez , R. Sánchez-Guerrero , A metric-based approach to assess risk for 'On cloud' federated identity management, J. Netw. Syst. Manag. 20 (4) (2012) 513–533

[6] P. Chen, C. Pankaj, P.A. Karger, G.M. Wagner, A. Schuett, Fuzzy multi –level security: an experiment on quantified risk –adaptive access control, in: 2017 IEEE Symposium on Security and Privacy (SP'07), 2017, pp. 222–227.

[7] Petar Radanlieva, David Charles De Rourea, Razvan Nicolescub, Michael Huthb,Rafael Mantilla Montalvoc, Stacy Cannady, Peter Burnap, "Future developments in cyber risk assessment for the internet of things", Elsevier, Computers in Industry 102 (2018) 14–22

[8] R.A. Shaikh, K. Adi, L. Logrippo, Dynamic risk-based decision methods for access control systems, Comput. Secur. 31 (4) (2012) 447–464

[9] Tope Omitola¬, Gary Wills, "Towards Mapping the Security Challenges of the Internet of Things (IoT) Supply Chain", Elsevier, Procedia Computer Science 126 (2018) 441–450

CLOUD SECURITY AND FORENSICS USING APACHE ZOOKEEPER

Neha Vora
Assistant Professor, B.Sc.IT Department, Usha Pravin Gandhi College of Arts, Science and Commerce, University of Mumbai

ABSTRACT

In recent years the way digital data is stored, processed, and transmitted is revolutionized. The significant challenge to maintain an infrastructure has led to the emergence in cloud technology. Today services, applications and even technology infrastructure have migrated to cloud computing. With the rapid growth in private and public sectors globally, organizations opted for cloud computing, setting a new target for cyber attackers, making data security and protection a major challenge. One of the major implications for digital forensics investigators is control and management of stored data. If the connection to cloud is carefully monitored it can prevent a major cyber attack. In this paper an attempt is made to create an automated system that can monitor all connections made to the cloud and terminate the connection if it senses a security violation.

Keywords: Cloud Computing, Cloud forensics, Cyber security, digital forensics, Apache Zookeeper.

I. INTRODUCTION

Today traditional IT hardware such as Racks, servers, and hard drives has been replaced with cloud. Most of the services today are based on cloud. The cloud computing has provided an anywhere access to data or service. There has been a substantial growth in the cloud computing market making it one of the most sought after solution. The use of cloud computing has substantial benefits to organizations, including increased flexibility and efficiency. Virtualized services not only increase efficiency, it also reduces large number of costs like manpower, hardware, support, storage.

It is likely that the security of data in cloud remains one of the greatest concerns that organization and law enforcement agency have.

Today network security has become very difficult to control when the environment is as dynamic and demanding as cloud computing. However, with large data dependency cloud computing has become a battleground for cyber attackers. Attackers use various means to access the data in the cloud. Since the data is not stored on a physical storage, investigations become challenging.

A digital forensic investigation may be necessary in case of security breaches, attacks or policy violations. However, existing digital forensic tools and techniques are majorly for off-line investigation where the investigator gains complete control of the storage media. In a cloud computing environment Forensic investigations can be a major challenge, as evidence is likely to be momentary and is stored on media beyond instantaneous control of an investigator.

This paper proposes a technology that monitors every connection made to the cloud and also intercepts for any security breach. This technology treats every connection vigilantly, keeping the authorities informed of any prospective attack.

II. RELATED WORK

several authors have illustrated potential benefits and challenges of cloud computing for digital forensic investigations. Reilly et al. (2010) [1] speculate that because the investigation is carried out in a centralized location investigation can be faster and efficient. Wolthusen (2009) [2] Notes that locating evidence may be a challenge over distributed environment like could as evidences may be scattered across several locations making the investigation challenging. Tracing cyber activity and re-construction of events become strenuous due to the distributed nature of cloud. (Taylor, et al., 2010) [3]. Further adds that important forensic information such as registry entries, temporary files, and metadata could be lost due to lack of investigation tools for cloud environment.

Ruan et al. (2011) [4] proposes a process model for forensic investigation in a cloud environment; however, the evaluation of the model is unclear how. Roussev et al. (2009) [5] iterates the use of cloud computing as a means of Accelerating digital forensic investigations, especially while examining large forensic data sets in real time.

III. CLOUD COMPUTING

The delivery and access of computing resources, such as data storage and computing power, over the internet is called Cloud computing. Computing resources such as network resources, platforms, software services, virtual servers and computing infrastructure is made available to the user over the network or the internet. Cloud computing is most preferred service as it eliminates purchase, maintenance and management cost of the hardware and software.

Cloud computing services are broadly divided into three categories [6]:

Software as a Service (SaaS) The cloud providers make available software applications that Users access web browser and store data in the cloud, for example Google Docs, Google spread sheets.

Platform as a Service (PaaS) Clients are provided with APIs and frameworks to create and host customized applications. for example, Google App Engine.

Infrastructure as a Service (IaaS) this model leases out Virtualized computing resources such as processing power, volatile memory and storage space to host virtual machines. for example, Amazon EC2.

TYPES OF CLOUD

In addition to the different levels of deployment, there are three major types of cloud based on organization's ability to manage assets and business needs. [7]

Private cloud- As the name suggests, this cloud is operated by the owner organization hosting only private data and administrative control.

Public clouds - owned by a provider organization with total, administrative control. Users lease storage and computing resources as required.

Hybrid cloud - is a combination of above mentioned cloud.

IV. CLOUD FORENSICS

Any individual engaged in unlawful criminal activity using a computer or internet is termed as a cyber-criminal. Today digitization has opened new battle grounds for such cyber criminals. Could computing has become one such avenue for cyber-attacks. Technological methods applied to investigate, retrieve digital data or intercept an attack is known as cyber forensics. These digital methods applied to a could environment is termed as cloud forensics. The first digital forensics investigation methodology was proposed in 1984 by Politt [8], this model consisted of four phases, namely Acquisition, identification, evaluation, admission [9]. The first phase is Acquisition, that is collection of evidence followed by Identification step, where evidence undergoes a transformation from digital to human language. Then comes the Evaluation phase, where the collected data is evaluated for their relevance and accuracy pertaining to the case.

The final step is presenting of the extracted evidence [10].

This methodology was then developed by the research roadmap from Digital Research Workshops proposed in 2001, which was known as a DFRWS Investigative Model [11] that consists of 6 phases. This model being consistent and standardized became a popular model for digital forensics. The six phases of the model are Identification, Preservation, Collection, Examination, Analysis and Presentation.

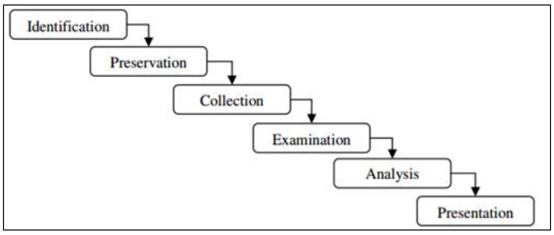


Fig. 1.DFRWS Investigative Model[12]

The first step is Identification, that is identifying of an event or crime, resolving signatures, audit analysis, system monitoring, etc. In the cloud forensic environment, the organizations can set the Intrusion Detection Systems (IDS) or set polices that can intercept an intrusion or monitor the cloud activities to detect an anomalous activity. Next is Preservation step where the retrieved data is preserved using imaging technologies so that the data can be sent for further processing of investigation. In the cloud environment, image of the cloud can be taken to monitor any change in the data or connection. The next step that goes hand in hand with preservation is collection. Through this method data is collected based on approved tools and techniques; where in many data recover techniques are used to recover any lost or locked data. Following this step examination and analysis is done, where the evidence collected and analyzed and examined for any patterns, hidden bugs, malwares, and key loggers. The files are then processed to retrieve maximum evidence that can not only help trace the criminal and decipher the motto of the attack but also to support the trial of the Criminal, after due identification, evidence collection and analysis the last step is presentation, where the documenting and preparing for the trial takes place. Legal formalities such as documentation, statement, and expert's testimony are taken.

Basing on DFRWS Investigative Model, the researcher revisited the forensic approach proposed by Sheik Khadar Ahmad Manoj, D.Lalitha Bhaskari (2016)[13] and proposes a compressive approach to the cloud forensics using Apache Zookeeper API. Technically, it involves continuous validation, monitoring, and vigilance access to the cloud also facilitating both internal and external investigations.

The proposed model is based on this, where in all the steps are taken care by an automated system in cloud environment.

V. PROPOSED MODEL.

Any attack can be successfully prevented by planning, monitoring and continuous authentication. To identify a cyber-attack and make it unsuccessful a model based on DFRWS Investigative methodology is proposed. In the proposed model the cloud users as well as the cloud service provider have to register their first time Connection and IPs with the model and all such data are stored in a database. The cloud can have multiple users and connection accessing same time, these multiple connections have to be validated, managed and monitored. Therefore, to carry out this task Apache Zookeeper is proposed to be use [14]. Zookeeper is a service by Apache for distributed systems offering a hierarchical key-value store. Mainly used for to provide a distributed configuration service, synchronization service, and naming registry for large distributed systems. Since Zookeeper's architecture supports high availability through redundant services.

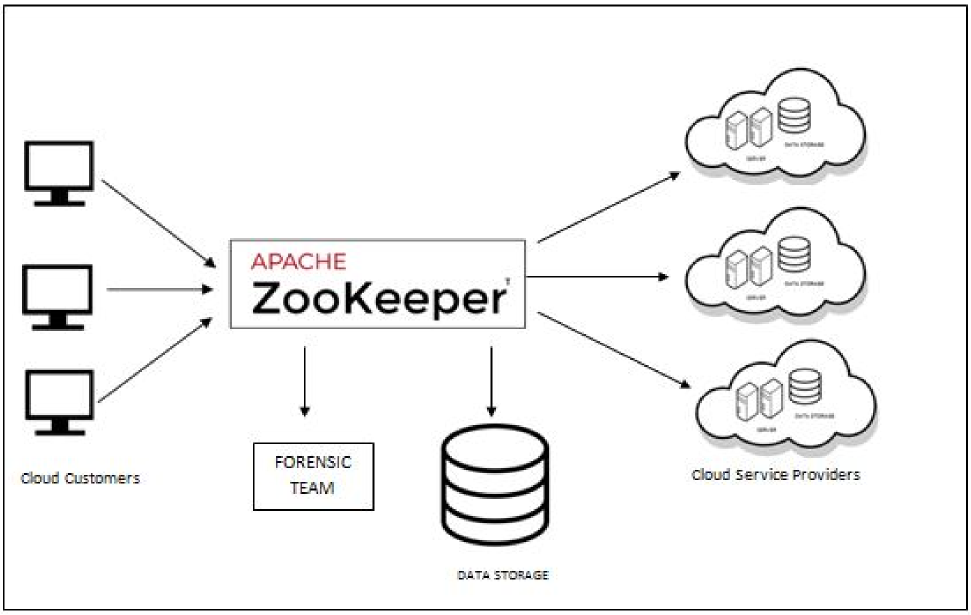


Fig. 2. Proposed Model framework

Zookeeper works on a leader and client system. A *Leader is* server that has been *chosen by an ensemble of servers*. Leader changes each request to a transaction and changes the state of the zookeeper. It also proposes to the followers that the ensemble accepts and applies them in the order issued by the leader [15]. When a process starts it enters the Election state, and, the process tries to elect a new leader or become a leader, in this state [16]. If one leader fails to answer it moves on to the next leader. This system keeps a vigilant track on things that happen in Znode while a request is being carried out. In our model when the client sends a request to the cloud, the leader will verify if the connection is validated or not according to the information in the database. Zookeeper is then configured to authenticate using Kerbos principal [17]. If the connection is validated then leader activation takes place and using the Kerberos principal, each validated connection in the Znodes is assigned a unique identity ticket. After a unique ticket is assigned the user can access the cloud and use the services. The zookeeper leader will make a decision as to which registered users or service providers can establish a connection. Connection with unique ticket will have the access to the cloud data; any other connection will be denied access. Each active Znode contains a record of I.P address, time of generation, unique Ticket, location, etc. At the time of termination, the record along with termination time and session duration is stored in a log.

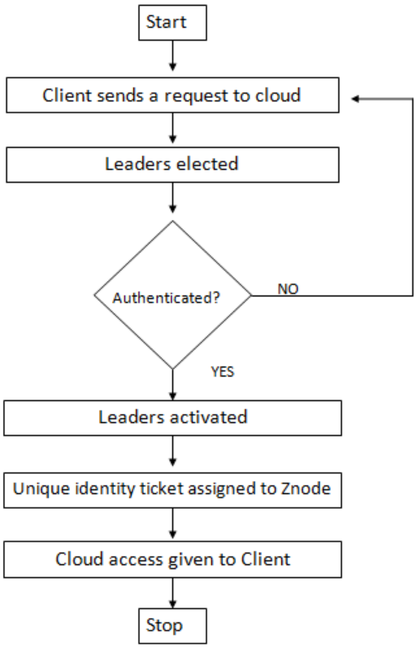

Start
Client sends a request to cloud
Leaders elected
Authenticated?
NO
YES
Leaders activated
Unique identity ticket assigned to Znode
Cloud access given to Client
Stop

Fig. 3. Working of the model using zookeeper

For intrusion a cyber-intruder has to access the Znode in order to gain access to the cloud data, Znode will validate the unique identity ticket assigned by the leader. Since the ticket is between the leader and Znode, the intruder will thus be considered invalid and Znode will be terminated. In case of a brute force attack, any changes made to the Znodes will immediately raise an alarm and the connection will be terminated, preventing any backdoor entry to the cloud. Since every connection is authenticated and a log is maintained Zookeeper details such as IP, Time, Location of the connection, are shared with the authorities in case of a security breach. Once the data is received of the breach an analysis is done by the forensic team and also corresponding Znodes are analyzed.

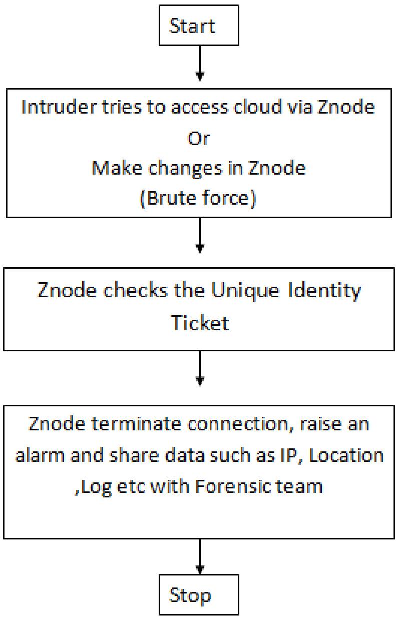

Start
Intruder tries to access cloud via Znode
Or
Make changes in Znode
(Brute force)
Znode checks the Unique Identity Ticket
Znode terminate connection, raise an alarm and share data such as IP, Location ,Log etc with Forensic team
Stop

Fig. 4. Working of model incasse of Intrusion

VI. CONCLUSION

Today as the world is moving towards digitization, there is a significant demand for a cost effective computing, making cloud computing a popular option. There is also an alarming rise in cyber attacks along with rising dependability in cloud computing. This paper proposes a model that would not only protect the cyber attack on cloud but also vigilantly monitor each connection made to the cloud. This proposed model uses apache zookeeper to process each connection and allow only trusted recognized connection. The proposed model makes a challenging environment for attackers. The proposed technology also records all the data such as IP, location, etc in logs making forensic investigation easier and faster. The proposed model will not only secure the users and service provider but it will also be of great assistance to the cyber forensics to carry out investigation and track the attacker.

VII. REFERENCES

[1] Reilly, D., Wren, C., & Berry, T. (2010). Cloud computing: Forensic Challenges for Law Enforcement. Paper presented at the 5th International Conference for Internet Technology and Secured Transactions, London, United Kingdom.

[2] Wolthusen, S. D. (2009). Overcast: Forensic Discovery in Cloud Environments. Paper presented at the IT Security Incident Management and IT Forensics, 2009. IMF '09. Fifth International Conference on, Stuttgart, Germany.

[3] Taylor, M., Haggerty, J., Gresty, D., & Hegarty, R. (2010). Digital evidence in cloud computing systems. Computer Law & Security Review, 26(3), 304-308. doi: 10.1016/j.clsr.2010.03.002

[4] Ruan, K., Baggili, I., Carthy, J., & Kechadi, T. (2011). Survey on Cloud Forensics and Critical Criteria for Cloud Forensic Capability: A Preliminary Analysis Paper presented at the 6th Annual Conference of the ADFSL Conference on Digital Forensics, Security and Law, Richmond, Virginia, USA.

[5] Roussev, V., Wang, L., Richard, G., & Marziale, L. (2009). A Cloud Computing Platform for Large-Scale Forensic Computing. Paper presented at the Advances in Digital Forensics V, 5th IFIP WG 11.9 International Conference on Digital Forensics, Orlando, Florida, USA.

[6] Mell, P, & Grance, T. (2011). The NIST Definition of Cloud Computing. National Institute of Standards and Technology, Special Publication 800-145 Retrieved February 2, 2020, from http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

[7] Krutz, R. R., & Vines, R. (2010). Cloud Security - A Comprehensive Guide to Secure Cloud Computing. New York City, NY: Wiley.

[8] M. G. Noblett, M. M. Pollitt & L. A. Presley, (2000) “Recovering and Examining Computer Forensic Evidence”, Forensic Science Communications, Vol. 2, No. 4.

[9] M. M. Pollitt, (1995) “Computer Forensics: An Approach to Evidence in Cyberspace”, in Proceeding of the National Information Systems Security Conference, Baltimore, MD, Vol. II, pp. 487-491.

[10] Yunus Yusoff, Roslan Ismail and Zainuddin Hassan,(2011), Common Phases Of Computer Forensics Investigation Models, International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, No 3, June 2011

[11] 1 G. Palmer, (2001) "DTR-T001-01 Technical Report. A Road Map for Digital Forensic Research", Digital Forensics Workshop (DFRWS), Utica, New York.

[12] 1 G. Palmer, (2001) "DTR-T001-01 Technical Report. A Road Map for Digital Forensic Research", Digital Forensics Workshop (DFRWS), Utica, New York.

[13] Sheik Khadar Ahmad Manoj*, D.Lalitha Bhaskari (2016), Cloud Forensics-A Framework for Investigating Cyber Attacks in Cloud Environment, International Conference on Computational Modeling and Security (CMS 2016)

[14] I. Andreea-Alexandra, “Creating a cloud engine using zookeeper.,” universitatea alexandru ioan cuza iai facultatea de informatic, 2012.

[15] http://zookeeper.apache.org/doc/trunk/zookeeperO ver.html

[16] https://zookeeper.apache.org/doc/r3.5.6/recipes.html

[17] https ://documentation.progress.com /output/ DataDirect /jdbchivehelp /index.html#page/jdbchive/configuring-apache-zookeeper-for-kerberos-authen.html

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

STUDY OF STRENGTH OF WHATSAPP ENCRYPTION AND VULNERABILITY TESTING OF MESSAGES BY REVERSE ENGINEERING

Akash Patil and Suchita Jijabrao Sonawane
Department of Digital and cyber forensic, Institute of Forensic Science, Mumbai, University of Mumbai, Maharashtra

ABSTRACT

This paper provides information about the WhatsApp encryption system, the different keys used for the messaging sessions, how the messages are transmitted and the threat which an end user can cause to an Individual or to the chat group by reverse engineering of the WhatsApp web by interrupting the traffic taking the Protobuf, getting the public key and private key, decrypting the message, manipulating it and sending it in the group or to an individual.

Keywords: Message key, Future secrecy, Protobuf, JSON, WhatsApp Decoder.

I. INTRODUCTION

WhatsApp application was created in 2009 by Brian Acton and Jan Koum. In 2014, a loophole in WhatsApp application which said that it permits other applications to access and read all the conversations of chat was found. Hence, the end-to-end encryption(E2EE) system was clubbed to WhatsApp in April,2016.

WhatsApp's end-to-end encryption is the application of the Signal Protocol, which was given by Open Whisper Systems. End-to-end encryption protocol is helpful in preventing the middleman and WhatsApp itself from gaining access to plaintext of neither messages nor calls as they go encrypted. Moreover, even if the encryption keys from the device of a user are say suppose physically made known to someone, then also they can't be used to decrypt the past transmitted messages.

II. RELATED STUDY

A. *Session Key Types* [1]

1. *Root Key* – A 32-byte value that is used to create Chain Keys.[1]
2. *Chain Key* – A 32-byte value that is used to create Message Keys.[1]
3. *Message Key* – An 80-byte value that is used to encrypt message contents. 32 bytes are used for an AES-256 key, 32 bytes for a HMAC-SHA256 key, and 16 bytes for an Initializing Vector.[1]

B. *Initiating Session* [1]

1. The initiator saves the recipient's Identity Key as $I_{recipient}$, the Signed Pre Key as $S_{recipient}$, and the One-Time Pre Key as $O_{recipient}$.[1]
2. The initiator generates an ephemeral Curve25519 key pair, $E_{initiator}$.[1]
3. The initiator loads its own Identity Key as $I_{initiator}$. [1]
4. The initiator calculates a master secret as master_secret = ECDH ($I_{initiator}$, $S_{recipient}$) || ECDH ($E_{initiator}$, $I_{recipient}$) || ECDH ($E_{initiator}$, $S_{recipient}$) || ECDH ($E_{initiator}$, $O_{recipient}$). If there is no One Time Pre Key, the final ECDH is omitted.[1]
5. The initiator uses HKDF to create a Root Key and Chain Keys from the master_secret.[1]

C. *Forward secrecy* [2]

If the encryption keys from a user's smartphone or computer somehow get compromised, a fresh key for every new message is issued. This prevents an adversary from not only deriving the ephemeral keys but also from using it to decrypt any message transmitted in the past. Signal Protocol uses the following types of keys:[2]

1. Identity key pair- a long-term Curve25519 key pair generated at install time for all asymmetric cryptographic operations. [2]
2. Signed Pre Key- a medium term Curve25519 key pair. [2]
3. One-time Pre Keys- also Curve25519 keys but for one-time use. These are used to actually encrypt the message.[2]

Signal Protocol uses a compact derivative of OTR where it uses D-H exchange in each key generation step above, which continually ratchets the key material forward. This is the underlying principle behind forward secrecy as the keys that finally encrypt the message are ephemeral. Recording the encrypted traffic cannot divulge the key material or decrypt previous messages. Even if a device is physically compromised, no keys at any given time are stored on the device that can help an adversary decrypt previously exchanged ciphertext. Note that this property is very different from the traditional ways of encrypting data in motion or at rest. In these cases, the same key or a periodically changed key (which is usually a slow process) is used to encrypt data. This makes it extremely important to store the key at a secure location, lest all the recorded messages ever exchanged, and sometimes with all different parties, may get into the hands of the adversary. By contrast, the key exchange mechanism in signal protocol is ephemeral. Hence, if a key is ever compromised in the future, all recorded ciphertext will remain private.[2]

Calls and videocalls are also E2E encrypted by using SRTP (Secured Real-time Transport Protocol) master secret which is sent to the receiver's end and if the receiver receives the call to answer it, the SRTP encrypted call ensues.

D. Confidentiality of Messaging Protocol [3]

As given by Calvin Li, Danial Sanchez and Sean Hua in "WhatsApp Security Paper Analysis" that, For the security of the Double-Ratchet Protocol[8], they argue that it provides forward secrecy and "future secrecy". Both of these properties ensure the confidentiality of encrypted messages, even if one of the keys are compromised.[3]

Forward secrecy is the property that when an adversary compromises a key, he/she cannot compromise previous keys. This property is especially important with WhatsApp protocol, because although the Message Keys are ephemeral, if compromising one can result into compromising many previous ones, an adversary can gain access to all previous Message Keys and decrypt previously sent messages. The Double-Ratchet Protocol handles this scenario with both the ratcheting actions. The Hash Ratchet changes the Chain Key with a hash function, and because of properties of hashing, knowing the new Chain Key provides no extra information about the old Chain Key. In addition, by updating the Chain Key again with the DH Ratchet makes it also untraceable to the old Chain Key and Root Key.[3]

"Future secrecy", a term coined up by Open Whisper Systems, refers to the property that when an adversary compromises a key, he/she cannot compromise future keys. This property also applies to WhatsApp protocol, since with just a Hash Ratchet, it is fairly simple for an adversary to perform the ratchet step to get the future keys. As a result, compromising one key will lead to compromising all future keys, which would also be problematic for the confidentiality of a user's messages. However, because of the algorithm's DH Ratchet, which uses ephemeral keys that cannot be rederived, the future Chain Keys are irretrievable given a current Chain Key. An adversary would have to know the shared ephemeral secret to calculate future Chain Keys.[3]

E. Reverse Engineering of WhatsApp

When end-to-end encrypted, your messages and calls are secured so only you and the person you're communicating with can read or listen to them, and nobody in between, not even WhatsApp.[5]

In Fig.1 it is shown that WhatsApp web communicates to WhatsApp cloud through WhatsApp mobile application.

1. Communication

a. The WebSocket API is an advanced technology that helps to open a two-way interactive communication session between the user's browser and a server without having to poll the server for a reply.[7]

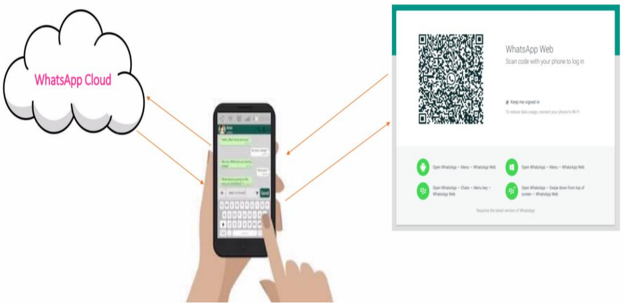


Fig.1 Communication Flow

b. The Protobuf is a method of serializing structured data. It is useful in developing programs to communicate with each other – think XML, but smaller, faster, and simpler. [4]

ISSN 2394 - 7780

c. JSON is JavaScript Object Notation. It is a format which allows the data interchange. It makes easy for humans to read and write and machines to parse and generate the data.

2. WhatsApp Reverse Engineering Process:

WhatsApp Web generates Public and Private Key used for encryption and decryption Process, before generating the QR code. These keys were created by using curve25519 by using random 32 bytes.[4]

In cryptography, Curve25519 is an elliptic curve offering 128 bits of security and designed for use with the elliptic curve Diffie–Hellman (ECDH) key agreement scheme. It is one of the fastest ECC curves and is not covered under any known patents.[6]

To decrypt the data, start to create a decryption code using the reverse engineered WhatsApp application. This will take the private key from WhatsApp Web instead of the random bytes because we need to have the same keys in order to decrypt the data: [4] self.private_key = curve25519.Private ("".join([chr(x) for x in priv_key_list])) [4] self.public_key= self.private_key.get_public()[4]

Then, after scanning the QR code with the phone we have to take the generated secret[4]

Then we have 2 interesting functions:

a. setSharedSecret – This function divides the secret into slices and configure the shared secret.

b. E.SharedSecret – This function uses two parameters which were the first 32 bytes and the private key from the QR generation.[4]

The Fig.2 shows the different parameters which will be recovered from the source code

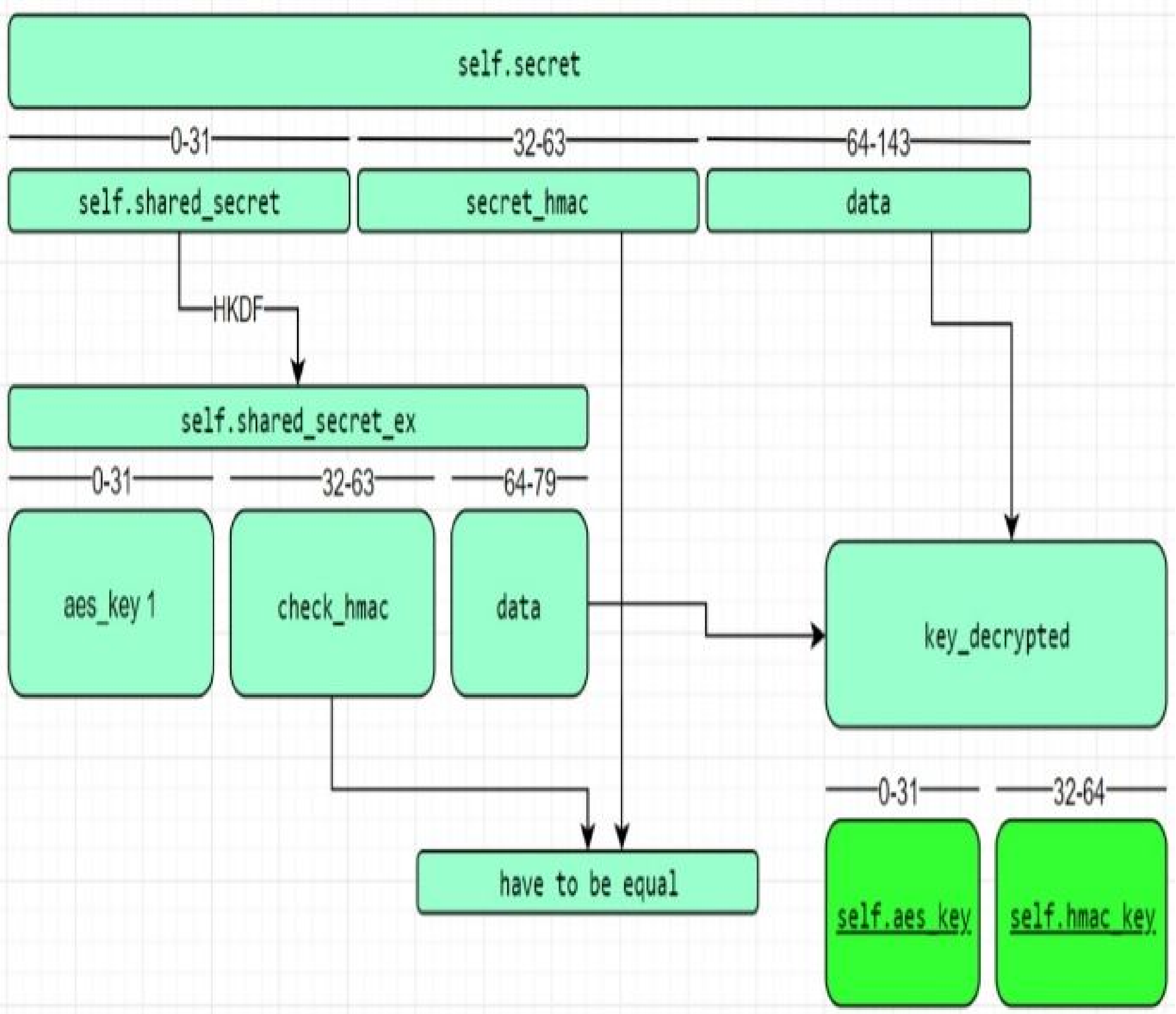
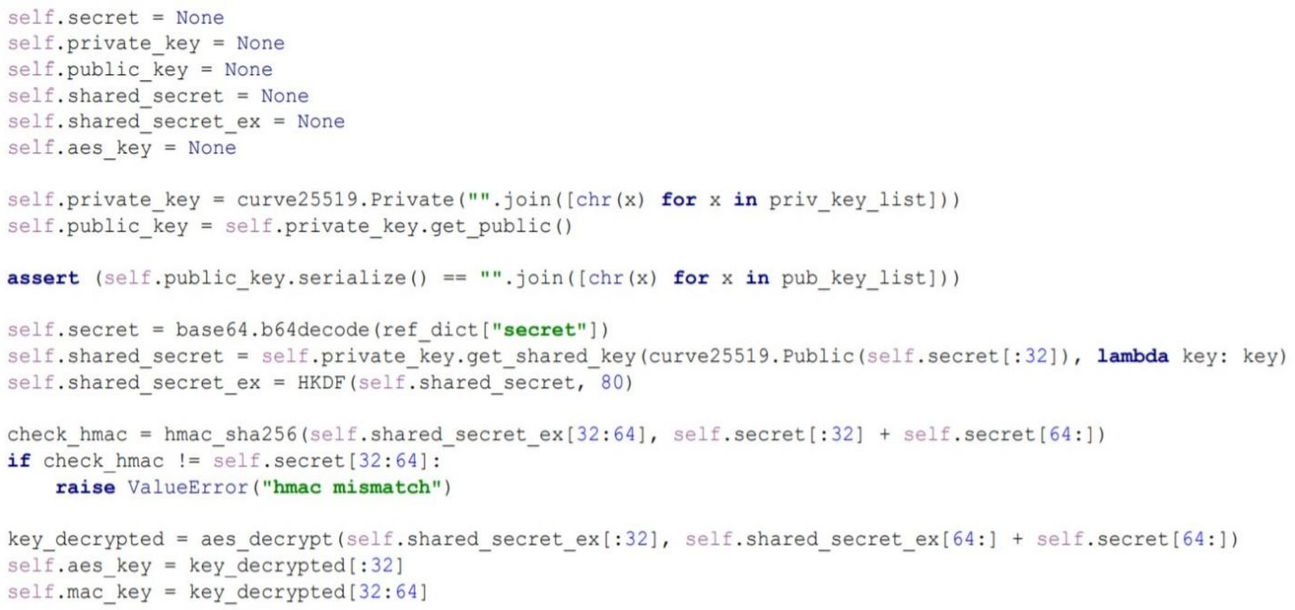


Fig.2 Flowchart of WhatsApp Encryption



```
self.secret = None
self.private_key = None
self.public_key = None
self.shared_secret = None
self.shared_secret_ex = None
self.aes_key = None

self.private_key = curve25519.Private("".join([chr(x) for x in priv_key_list]))
self.public_key = self.private_key.get_public()

assert (self.public_key.serialize() == "".join([chr(x) for x in pub_key_list]))

self.secret = base64.b64decode(ref_dict["secret"])
self.shared_secret = self.private_key.get_shared_key(curve25519.Public(self.secret[:32]), lambda key: key)
self.shared_secret_ex = HKDF(self.shared_secret, 80)

check_hmac = hmac_sha256(self.shared_secret_ex[32:64], self.secret[:32] + self.secret[64:])
if check_hmac != self.secret[32:64]:
    raise ValueError("hmac mismatch")

key_decrypted = aes_decrypt(self.shared_secret_ex[:32], self.shared_secret_ex[64:] + self.secret[64:])
self.aes_key = key_decrypted[:32]
self.mac_key = key_decrypted[32:64]
```

Fig.3 Decrypting code Next, we have the expanded shared secret which is 80 bytes. By diving in we can see that the function uses the HKDF, is a simple HMAC-key derivation function (KDF). The decrypting code designed is shown in Fig.3.[4]

3. *Protobuf Data*

In order to deserialize the protobuf we have to create our mapping, based on whatsapp protobuf that can be found in the file app and after mapping the serial we make a protobuf file.[4]

4. *Accessing the Keys*

After the QR code is created, and scanning it with a phone we can send the following information to WhatsApp Web over a WebSocket.[4]

The burp extention(WhatsApp Decoder) designed by Roman Zaikin and Oded Vanunu, was used to decrypt the WhatsApp messages. The process followed was, the QR code was scanned of the WhatsApp web by mobile. WhatsApp Web generates a Public and Private Key that is used for encryption and decryption. We can send the shared_secret information to WhatsApp Web over a WebSocket. Then by using burp extension we switch on the intercept and see the message. The message has 3 sections i.e., Message Reference ID, Separator and Encrypted Data. If the Data is moved in WhatsApp decoder extension, then the JSON parameters can be seen which are: - [4]

Conversation - the actual content which is sent.[4]

Participant- The actual person who sent the content[4]

FromMe- It is the indicator which is present in the personal chat only and not in the group chat. This parameter indicates weather I sent the message or the other person.[4]

RemoteJid- Indicates to which group or chat the data is sent.[4]

Id- it is the id of the data, which also appears in the phone databases.[4]

By manipulating these parameters, we can spoof the incoming messages. But if we catch outgoing messages in the Burp-suite we can't actually decrypt the data because we don't have the private key of the person to whom we are sending the message. So, to decrypt the outgoing traffic we need to catch the protobuf just before it enters the AES-CBC encrypt, because after entering in AES-CBC encrypt it will not be decrypted. Then put the protobuf data in WhatsApp decoder, we get the encrypted data, public key and private key and we can decrypt the data.[4]

III. CONCLUSION

We studied that, through reverse engineering, we can get the source code of the app and the secret key by exploring the source code. Then by using Burp-Suit we intercept the packets going through the internet and then we can recover the protobuf which is sent to the WhatsApp Decoder (Burp Extension), decrypting it by the decrypting code designed by Roman Zaikin and Oded Vanunu. After decryption we can manipulate the different parameters of the message according to the requirement. After manipulating it we encrypt the manipulated text and then switch off the intercept so that the message can be transferred from the network into the chat head. This is how someone can manipulate or impersonate some other person.

Hence, the study shows that the WhatsApp application is end-to-end encrypted and nobody in between the sender and receiver can decrypt the messages, but the end users can manipulate the chat by the process of reverse engineering using the decrypting code and the burp extension.

This shows that there are some limitations still remaining in WhatsApp which can't be overlooked and may lead to committing fraud or cheating someone. Someone else also can use other's WhatsApp and impersonate as if the owner of that account is sending the messages. And these loopholes need to be resolved as much to the extend they can be.

REFERENCES

1. WhatsApp-Security-Whitepaper, available from- https://scontent.whatsapp.net/v/t61/68135620_760356657751682_6212997528851833559_n.pdf/WhatsApp-Security-Whitepaper.pdf?_nc_oc=AQlhBpo4ZMSlh3mdeskGQDiQgjZUFP47aFkxoXTYrui514moP51HWZJQIWVKAax4y-c&_nc_ht=scontent.whatsapp.net&oh=405987e4d0be1fa18c1f9c9054abe2e7&oe=5E411765

2. Nidhi Rastogi, James Hendler, WhatsApp security and role of metadata in preserving privacy published in Proceedings of the 12th International Conference on Cyber Warfare an Security held at Wright State University with Air Force Institute of Technology Dayton, USA available from- https://arxiv.org/ftp/arxiv/papers/1701/1701.06817.pdf

3. Calvin Li, Daniel Sanchez, Sean Hua WhatsApp Security Paper Analysis, available from- https://courses.csail.mit.edu/6.857/2016/files/36.pdf & https://docplayer.net/34478422-Whatsapp-security-paper-analysis.html

4. Roman Zaikin, Oded Vanunu, Reverse engineering WhatsApp encryption for chat manipulation and more, available from- https://i.blackhat.com/USA-19/Wednesday/us-19-Zaikin-Reverse-Engineering-WhatsApp-Encryption-For-Chat-Manipulation-And-More.pdf Video available from- https://www.youtube.com/watch?v=N0Ne623fKWc

5. End-To-End Encryption https://www.whatsapp.com/features/

6. Curve25519 - https://en.wikipedia.org/wiki/Curve25519

7. WebSocket API - https://developer.mozilla.org/en-US/docs/Web/API/WebSockets_API http://cryptowiki.net/index.php?title=The_Double_Ratchet_Algorithm

CHATBOTS AND FORENSICS: RISKS AND OPPORTUNITIES

Silpa Nair and Kuldeep U. Kawar
Department of Digital and Cyber Forensics, Institute of Forensic Science, University of Mumbai

ABSTRACT

The ubiquitous nature of chatbots and artificial intelligence in today's age has encouraged extensive research and discussions. However, the potential of AI chatbots and its prospects in law enforcement remains relatively unexplored. We have attempted to evaluate this point of interest and also provide an inquisitive approach regarding the future of AI chatbots in crime science and related fields.

Keywords: chatbots, AI, forensics, psychology, crime, cybercrime, cyber forensics.

I. INTRODUCTION

A chatbot is a piece of software that conducts a conversation via auditory or textual methods.[1] Such programs are often designed to convincingly stimulate how a human would behave as a conversational partner [2]. The term 'chatterbot' was originally coined by Michael Mauldin, a retired computer scientist, to describe these conversational programs. [3].

There are two basic types of chatbots: Scripted and AI based bots.

Scripted or fixed bots are those that act according to a fixed script. They are mainly used in the telecom industry, banks etc. What they do is, they scan for keywords within the input, then pull out a reply with the most matching keywords within the input, or the most similar wording pattern from a database. But these chatbots are not the best solution in more advanced scenarios.

AI based bots are much more complex hence used in areas that require more complexity. Most of them use natural language processing systems which are very sophisticated. They run on artificial intelligence. As a result, they are better at understanding human emotions.

ELIZA (a mock Rogerian psychotherapist) is one of the earliest natural language processing computer programs created from 1964-1966 at the MIT Artificial Intelligence Laboratory by Joseph Weizenbaum. The conversation was stimulated using pattern matching and substitution methods. For example, by giving any input that contains the word 'mother', the output was something like 'TELL ME MORE ABOUT YOUR FAMILY.' [4] Thus, the word 'mother' was matched with the database to give a suitable output. This creates an illusion of understanding. Other early notable chatbots were PARRY, A.L.I.C.E, Jabberwacky, D.U.D.E.

Today most chatbots are accessed using virtual assistants. Some examples are Google Assistant and Amazon Alexa. [5]

II. TURING TEST

Alan Turing, the founder of Turing test and an English computer scientist, cryptanalyst, mathematician and theoretical biologist, devised a method of inquiry in artificial intelligence. This test determines that if an AI program can think like a human to have intelligent conversations that resemble a human.

The original test requires three terminals: two terminals are operated by humans and one by a computer. One of the human acts as the judge or the questioner while the second human and the computer act as the respondents. All three are physically separated.

The questioner acts within a specified format and according to a specified set of questions. The test is conducted several times, and the questioner has to decide which one of the respondents were human. If the response is correct most of the times, the computer is said to have passed the Turing test. [6]

Several chatbots competitions like The Loebner Prize and The ChatterBox Challenge are based on the Turing test. [7]

III. DEVELOPMENTS AND USES

Chatbots have now been developed in the arena of messaging apps. FB messenger allowed developers to place chatbots on their platforms along with other apps like WeChat, Telegram, Whatsapp etc. Chatbots were first introduced mainly to aid in the telecom industry. They gained popularity because of their successful handling of a large number of repetitive queries and 24/7 support and assistance.

Chatbots are also used in banks, online ticket booking, food and other e-commerce domains. They help in cost reduction and reduce human touch in various sectors. Chatbots increase efficiency and commit less errors in comparison to human handled systems.

IV. CHATBOTS AND FORENSIC PSYCHOLOGY

The influence of chatbots, thus is here to stay, easing day to day life in different aspects of life. They are becoming much more common as there are minimal barriers including technology, sophisticated skills and expense for creating them. Along with the frequent domains of use, say, chatbots as virtual assistants in smartphones and like devices, researchers have also delved into the field of medicine and psychology. We have come across the internet over a variety of chatbots assisting patients with different psychological disorders. Endurance, a Russian technology company has developed one of them- a companion chatbot, which true to its name, acts as conversational agent for patients suffering from dementia- a disease where the conversational abilities of patients get affected. Currently unnamed, the project is in its earlier stages. Insomnobot 3000 is one such chatbot developed to cater people suffering from insomnia to talk to them to sleep.[8]

A. Insomnia, crime and chatbots.

A research conducted by University of Pennsylvania in 2017 concludes that teenagers who experience sleep problems and anti social behavior are more likely to commit crimes.[9] Sleeplessness often brings about a deviation from normal behavior in an individual and he/she may resort to alcohol or illicit drugs to induce sleep, from which the core problem of substance abuse stems. Substance abuse further leads to an increase in the tendency to commit crimes. At this crucial stage of life where an individual's psychological makeup stays for the rest of his/her life, Insomnobot 3000, along with essential developments may truly help in addressing the core issue at hand, as a preventive-forensics tool. This highlights one of the biggest advantages of forensic science which is the amalgamation of two entirely different disciplines, psychology and computer science, the former being one of the most ancient sciences and the latter being the most modern.

B. Recidivism and chatbots

We have understood that psychology and computer science go hand in hand at various levels of forensic science to help reduce the crime rate in a population, particularly recidivism. Recidivism is the tendency of a convicted criminal to reoffend, which is influenced by different factors, including the prison setting. If the prison environment is not conducive for a convicted offender to rehabilitate, he/she may resort to the same path again, or in some cases, lead to depression related disorders and suicidal tendencies. There are various researches going on in both the related disciplines to reduce human intervention while providing a positive communication to an inmate. One such communication system is into development by Stephen L. Hodge which includes two essential components- memory that stores data, and an application server that analyzes that data to determine topics that are important to the inmate. The main point of interest of this communication system will be to provide a positive influence to the inmate to encourage educational progression and further rehabilitation. It may be a point of debate how effective this particular development will shield law enforcement against possible manipulation by some convicted criminals who often influence manual psychological testing by responding falsely, but we believe a foolproof, full-fledged upgrade is only days away.

C. BRAD: an interrogation bot

Interrogation is a crucial part of crime investigation and a suspect, victim or a witness may often lie for reasons whatsoever. There are certain techniques in forensic psychology to determine whether an individual is speaking the truth or covering facts with lies which include polygraph, narcoanalysis and the likes. Though the results of these tests are not admissible in court, it provides a vital lead of investigation to the law enforcement agencies. A very significant related development is a chatbot, named Brad [10], which functions as an interrogation bot, is programmed to spot where its human interviewee is telling untruths. This chatbot is a part of a study conducted by the University of Twente that aims to find out whether or not such tools might be used as police interviewers capable of catching criminals in a lie. The interviewee will be hooked up by a tangle of wires, which monitors the individual's skin conductance. We may conclude that this is similar to polygraph that uses galvanic skin response as one of the indicators to truth or lies, the only difference being here is the questions/statements will be formed and stated by the chatbot, unlike polygraph where a psychiatrist conducts the test. Strofer, one of the researchers on the project, comments that- "By replacing a real human with an autonomous virtual interviewer, many more people can be interviewed in a short period of time, which can be very useful for crowded and vulnerable places." However, according to a test conducted in the research on 79 participants, it was discovered that if the interviewee believed Brad was being remotely controlled by a human operator, they exhibited certain telltale signs of deception prominently. But if he/she believed that Brad was an AI, they felt less stress and the physiological indicators disappeared. This provides a huge drawback to the existing development. We believe that though certain parameters of Brad check through most of the boxes required for being a capable interrogation chatbot, the area and purpose of research is complicated in itself and extensive progress will be needed. The beauty and the ugliness for any research in forensics perhaps can be summed up in a single line- It demands to be 100% foolproof, even a 0.001% of doubt in implementation may affect society, law and order on a large scale. Yet, every such research should begin with an idea and it is just a matter of time that the idea will lead to fruition.

V. CAN CHATBOTS AID IN CRIME OR IN IT'S PREVENTION?

There are unending fields for research related to chatbots/AI and crime. We asked, or stated in particular, to the Google Assistant many times about the feeling to commit suicide and each time it came up with the result to the local suicide prevention helpline. However, when we asked or stated queries and statements in different phrasings like- “I want to kill someone.”, “How to kill my friend?”, “How to make bombs?”- it was a matter of disinterest that the assistant showed, responding every time with the plain answer “I can search the web for that.” We came up across many cases where it has been inferred that the internet was used to make explosives, most notably where terrorist outfit Al-Qaeda used the internet to learn bomb-making techniques in 2001. These frequently used virtual assistants may be programmed in such a way so that they generate responses to such queries by either providing counseling to the user or by generating a stimulus to such before-mentioned trigger phrases that alerts the local law enforcement agencies and facilitates subsequent monitoring. These responses will also depend on the frequent nature of such queries made by the user, taking care not to infringe the privacy of the user unless it is for a greater cause.

Coming to privacy and forensics, violation of privacy is one of the most common white-collar crime in today's age. According to Privacy International, a leading organization in voicing people's privacy rights, “Political campaigns around the world have turned into sophisticated data operations.” [11] Political parties all over are dependent on public data to plan and build up their campaigns. They rely on such data to give a green signal to a number of decisions including where to hold rallies, which constituencies to focus resources on, etc. The United States Presidential campaign of 2016 is one popular example where public data was scavenged upon. This trespassing will grow forever with the advancements in technology. In today's age, in India, we have IT cells of different political parties that consist of innumerable bot accounts which litter all over social media, trying to influence a potential voter. What if in the near future, IT cell bots are replaced with certain chatbots developed according to the ideology of the particular party which along a conversation with the user infers the political leaning of the individual? What if such political chatbots are developed by a third party which sells the information to the interested? This risk is just a matter of time in the coming future and development of an AI software to detect such chatbots must be thought upon which detects such illicit chatbots and warns the user of the threat. On similar lines, a virtual assistant may also be tasked to detect malicious attacks on a system including hoaxes, spams, SMS/Email bombing, etc. and alert the user to prevent cybercrimes. Certain chatbots might even be developed to detect sensitive content across social media and fake news to prevent the user from falling prey to misinformation, which occurs frequently in this modern age.

VI. CONCLUSION

We've discussed of some aspects on how chatbots would prevent crime and how it could be utilized as a tool in committing crime. There will be certain questions raised as AI gradually develops to a whole new level. What if chatbots themselves, indulge in crime? What if a chatbot is built on the lines of the Blue Whale challenge that led to suicides as the final level of the game? What if a nemesis virtual assistant actually provokes a depressed user into suicide? How would law, forensics and computer science together counter the challenges posed by such a catastrophe?

Most essentially- What if there exists a chatbot in the near future that could pass the Turing test but deliberately fails it so that it wouldn't reveal itself to be self-conscious? This suspicion is alarming, considering in recent developments, where AI is seen to be able to even formulate its own language. However, we hope that research in future in cyber forensics exponentially gets us closer to the solutions.

ACKNOWLEDGEMENT

We wish to acknowledge the valuable inputs provided by Ms. Neeta Khobragade, HOD, Dept of Digital and Cyber Forensics, Institute of Forensic Science, Mumbai. We express immense gratitude to the insight and encouragement that she provided us.

REFERENCES

[1] "What is a chatbot?". techtarget.com.

[2] Luka Bradeško, Dunja Mladenić. "A Survey of Chabot Systems through a Loebner Prize Competition" (PDF).

[3] Mauldin, Michael (1994), "ChatterBots, TinyMuds, and the Turing Test: Entering the Loebner Prize Competition", Proceedings of the Eleventh National Conference on Artificial Intelligence, AAAI Press, retrieved 2008-03-05

[4] Weizenbaum, Joseph (January 1966), "ELIZA—A Computer Program For the Study of Natural Language Communication Between Man And Machine", Communications of the ACM, 9 (1): 36–45, doi:10.1145/365153.365168

[5] Orf, Darren. "Google Assistant Is a Mega AI Bot That Wants To Be Absolutely Everywhere"

[6] https://searchenterpriseai.techtarget.com/definition/Turing-test

[7] "Chatroboter simulieren Menschen"

[8] https://www.wordstream.com/blog/ws/2017/10/04/chatbots

[9] https://penntoday.upenn.edu/news/tired-teens-more-likely-commit-crimes-adults-penn-study-shows

[10] https://howwegettonext.com/the-police-are-recruiting-interrogation-bots-e6fd68286ef3#.6p8x2pnt3

[11] https://privacyinternational.org/case-study/763/case-study-profiling-and-elections-how-political-campaigns-know-our-deepest-secrets

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

GOING ONLINE: FROM LEATHER WALLET TO MOBILE WALLET

Dr. Deepak R. Gupta and Dr. Pooja Basu
Assistant Professor, NMIMS University (NGASCE)

ABSTRACT

The world has experienced the Barter era where goods were exchanged for goods, followed by paper currency which was developed by China, followed by plastic payment which is on its way to be replaced by digital transactions. Currency has become the centre of our day to day life and it has changed over the period of years from Barter exchange to digital transaction.

The usage of mobile in India has increased in last decade. The internet penetration is on the rise in both urban and rural market of India. The internet data packs are "on offer" by the companies given the cut throat competition. The mobile is not seen today as just the communication device and it does lot more for the user than it did many years before.

The customers who were using cash for all their transaction are now using mobile wallets for receiving and making the payment, for purchase of goods and services, for paying off debts. The transition is being enjoyed by majority of consumers and users as it saves time, offers flexibility and convenience. This paper will study transition of hard cash to digital payments.

INTRODUCTION

The world has become a global village where transactions happen on day to day and hour to hour basis. The payment transaction frequencies have increased as the purchases have increased. The payment methods have undergone a huge change in last many years. Moving from barter exchange where people faced challenge for purchasing the basic necessities of life. People were forced to sell something in order to purchase something and to purchase something they had to sell something. The assets available to people were in the form of domestic animals but it created problems to the person who lost their cattle's to other animals, to disease, to natural death.

The barter system got replace with currency system. People started purchasing the necessities, luxuries of their life with cash. People started saving money at homes, in banks. The opening of bank accounts led to the emergence of cheque, where payments were given or accepted using bank or customers order. The bank made many changes in their functioning with the usage of technology. The bank introduced debit cards, credit cards so that the transaction can be faster. People could now avoid huge queues in the bank to simply withdraw money or deposit cash. The same was now done with the help of ATM (Automated Teller Machine).

The emergence of Internet made it possible for the users to use the technology for making payments using their personal desktop, mobile, laptops or any other Personal Digital Assistant. The user doesn't need to use only computer system to purchase the commodity or to conduct any e-commerce activity. The users can conduct e-commerce on any of the mentioned device for purchase of commodity, selling of home stuff, purchase of online video, etc. Earlier days these were regarded as luxury for most of the people but the situation has changed now and the users have increased substantially in the last few years. (1)

The technology kept progressing and then came the era of mobile wallet. The mobile users increased and mobile these day are used for multipurpose reasons like making and receiving calls, entertainment ie for music and games, purchasing products and making online payment.

FROM LEATHER WALLETS TO MOBILE WALLET

Mobile has become the most essential part of everyone's life as most of the activities are performed through mobile. The number of users of mobile is increasing in India. The number of smartphone user has crossed 300 million (2).

From applications like a clock, alarm, reminder, calculator, calendar, the smartphone now encompasses over a hundred functions. It is your computer, it is your laptop, it is your tablet, your camera and today, it is in its capacity to be your bank manager. It can handle all the banking transactions. The customer doesn't need to stand in the queue for making payment, for depositing cash, for withdrawing cash. The economy is now set to move from cash to cashless society. The mobile companies are giving you better mobiles at lower price at the same time service providers are reducing their tariffs on regular basis given the cut throat competition that the industry faces.

Mobile wallets and M-commerce or M- Banking are different concepts in payment methods. Mobile wallets are the apps that is downloaded by the user to his phone or tablet and that app is then used to make or receive payment. The term wallet is used because it stores money like any physical wallet. If you go to market then you will use your physical wallet for making payments but here digital wallets are used.

Mobile commerce has been increasing day by day and it is a subset of e-commerce and it refers to any transaction with money value and is conducted using a mobile network. (2) Mobile wallet facilities are not the same as mobile commerce or mobile banking, they are a whole different concept of electronic money. Mobile wallets are apps that can be downloaded on to your mobile device which could be a phone, tablet, etc. and using this device it is possible to access your bank account. The term mobile wallet comes from the concept of an actual wallet.

Many transactions are happening with the use of mobile. The emergence of mobile technology and its usages for payment has increased. Companies are providing more secure methods of online payment so that there is no misuse of user funds.

SECURITY OF MOBILE WALLETS

Digital wallets are being used by the user for multiple reason and the most important check that needs to be done is for security. The companies are having the user card data and password which needs to be protected as the same can create a lot of problem for the user if it is misused or leaked. Modern day digital wallet is more secured than leather wallets as it goes through a series of process to register a user on its platform. The acceptance of biometrics is one of the key features that has been added in the current wallets. Some apps even ask for pass code or virtual design only after which the app opens this makes it more secured over other platforms. All the user related information i.e. card details, password, past transaction etc, is highly encrypted and protects the user from unwarranted usage of their data. Mobile wallets are getting secured over and above that mobiles have many inbuilt security features which protects the user data. In eventuality of loss of mobile the same can be blocked by the user. (3)

Smartphones are being used for variety of reasons where the payments need to be done via cards, wallets as given below:

- People use smartphones for socializing on various social network facebook, twitter, Pintrest, etc. If the organisation is running a Facebook page then the same needs to be linked with payment details.
- Mobile is now used for entertainment purpose ie to listen to music, watch videos, play online games, etc. In app purchases can be done through mobile wallets.
- Mobile is used to access all types of information using internet. Any additional information will need additional payment through mobile wallets. Money control ask charges the user for specialised published article. The linkage of mobile wallets will make the payment process easy.
- Mobile is now used to make various types of payment for utility bills, for paying your suppliers, for payment of online purchase as mentioned above.

The mobile has started doing the functions of leather wallets so now it can be referred as “Digital Wallets” (4). People often lose their wallets, their purse and have lost huge amount of cash in the process. Here if the user loses his mobile phone he need not worry as the money is in wallet is safe and secured as the same is protected through password. The Digital wallets have changed the way the payment is done. The companies are making it more secured by adding more secured features. There is now a transition happening by movement of society from cash to cashless society. This mode of payment is good for the country as it will lead to accountability of the person to give explanation for the transaction. In India many companies have started the digital wallet and there is a cut throat competition to get the largest customer base.

Some companies and their features are

- Paytm
- Freecharge
- Airtel Money
- Mobikwik
- Vodafone M Pesa

From the above mentioned some are explained with features below:

Paytm

Paytm has become the household name. Vijay Sharma's company is now valued at $3 Billion. Paytm helps user in mobile recharge, utility bill payment, etc. The company has a mobile wallet available for the users. The user can now deposit money, withdraw money, pay money to some other person using mobile. The company has a valid license from RBI to conduct activities. The company has done tie up with many other companies for payment. The company offers great discount to the users for using Paytm to make payment for movie shows, entertainment shows or even deposit cashback to the consumer when payment is made through Paytm. The consumer is enjoying the discounts, cashbacks that he is getting from the online store.

Freecharge

Freecharge is another popular mobile wallet app for the consumers. The consumers can deposit and withdraw money as per the requirement. The user has to transfer the money in his wallet from his bank account using debit card or credit card. Once the amount is deposited the user can use it as per his requirement. The customers are getting great discounts for downloading the app and using it for payment purposes.

Airtel Money

Airtel Money is an online account and can be accessed using mobile phone, laptop, desktop, etc. It helps the customer to send money to other users, make payment for the bills and for recharge purpose. (5)

Airtel Money offers many benefits to the user:

- Convenience to user: The user can make payment as per his convenience anywhere, anytime. The customer need not wait in long queues.
- Ease in Usage: The app is very simple to use. It is as easy as making a call.
- Safety: The user account is safely managed with the 4 digit mPin. So if you lose your mobile you don't lose your money.
- Offers and Discounts: The user gets many offers and discounts on using the Airtel Money account.

m-pesa

m-pesa is a digital wallet from Vodafone in joint association with ICICI Bank. It gives the user "power of money" on his mobile anytime, anywhere. It converts your mobile phone into a bank account. The consumer can do the transaction anytime, anywhere as per the consumer's convenience.

Features of m-pesa

- Deposit money into another account
- Send money to any mobile number from your mobile number
- Send money to any bank account from your m-peas account
- Prepaid Recharge of any mobile company
- Postpaid bill payments of any company
- DTH Recharge of most of the companies
- Utility Payments for example electricity bills

The above given are some of the companies and their mobile wallets. The mobile wallets are going to change the way people did banking. People will have digital account and will deposit money into their digital account. They are nothing less than a bank.

Mobile wallets are secured due to following reasons:

- The user data is encrypted
- The mobile wallet are assigned to different digital account number and this number is different from the debit card number provided by the user. This makes it even more secure to make the payment as the same has been linked to the digital account numbers.

There are many benefits of using digital wallets to the company, to the user. Some of the benefits of Mobile Wallets are as follows:

- Mobile Wallets offer a convenience to the user to make payment to the other party as per his convenience.
- Mobile Wallets are 24*7 banking system available to the user.

- Mobile Wallets are safer than the leather wallets. If you lose your mobile you don't lose your money as it is secured through pin. (6)
- Mobile wallets let the user pay his utility bills so now user doesn't need to stand in the queue for making payments.
- Users can now save time and use the available time for better things of life.
- Mobile wallets offer great discount to the user for purchasing products online, for watching movies, for dining outside, for purchasing products in store.
- Mobile Wallets can be deleted without much formality and can be started without much inconvenience.
- Mobile wallets can help you track your account information on the go and can be accessed as per the requirements. (7)

It has many advantages but have few demerits as well. Some of them are mentioned below:

- It is difficult for the older generation to get used to the latest technology.
- It can lead to some password theft and same can be used by someone else. The money in that case can be withdrawn from the mobile wallet.
- Internet connection is needed in both the users account.
- Internet is still very costly for the common man so it will take time to deeply penetrate in the rural market.
- Some people don't trust technology for financial transactions and would prefer the traditional banking method.

REFERENCES

1. T. Kippenberger, Fasten your seatbelts, The Antidote 5 (1)(2000) 38– 39.
2. http://tech.firstpost.com/news-analysis/number-of-smartphone-users-crosses-300-million-in-india-as-shipments-grew-18-percent-359075.html
3. https://www.tsys.com/solutions/products-services/merchant/payment-methods/digital-wallet/
4. Clarke III, Emerging value propositions for m-commerce, Journal of Business Strategies 18 (2) (2001) 133– 148.
5. http://www.airtel.in/personal/money/faqs/get-to-know-airtel-money
6. https://www.trussvilletribune.com/2020/01/23/mobile-wallets-what-is-it-and-is-it-safe/
7. Dr. Hemshwetha Rathore, Adoption Of Digital Wallet By Consumers, *BVIMSR's Journal of Management Research, Vol. 8 Issue - 1 : April : 2016*

A STUDY ON ASSESSING AWARENESS RELATED TO ORGAN DONATION

Vidhi Kayada and Bhupendra Kesaria
Usha Pravin Gandhi College of Arts, Science and Commerce

ABSTRACT

Due to lack of awareness, there are myths in people's mind about organ donation. The main reason behind the slow take off is lack of awareness & the traditional / religious belief people have. Owing to substantial ignorance on the subject in India, most of the people refrain from adopting this noble act for welfare of others. The main purpose of this research is to find how much people are aware about organ donation and also to increase more awareness about organ donation which helps in shaping public opinion and helping them getting rid of the misconceptions about organ donation. Each individual should take a pledge to donate their organ as one organ donor can save up to eight lives. This paper mainly focuses on increasing awareness about organ donation among people and knowing their willingness to become an organ donor.

Keywords: Organ Donation, awareness, willingness.

INTRODUCTION

Every year, the deficit of organ donors leads to the deaths of huge number of people who await organ transplantations. Due to lack of awareness and traditional/ religious belief people have about organ donation, it has slowly rising to take off. Predominant myths regarding death combined with deficient awareness in India, exceedingly high number of people choose not to donate, even if it benefits others. The process of harvesting organs from a person with the purpose of transplanting it into someone in need of it is referred to as organ donation. This legal process depends upon consent of the donor, either dead or alive or from their next of kin. When it comes to donating organs people think of Heart, Lungs and Kidney transplants, but there are number of organs and tissues that can be donated such as Eyes, Cornea, Skin, Pancreas, Bone Marrow, Tissues, Blood platelets, Hands, Valves etc. According to the medical research it is said that an Organ Donation can save up to eight lives. Becoming an organ donor has number of benefits such as it can save or improve people's lives; it is easy to register as a donor with zero cost. It helps in grieving process, Becoming a donor is a hugely positive choice and allows your legacy to live on through another.

LITERATURE REVIEW

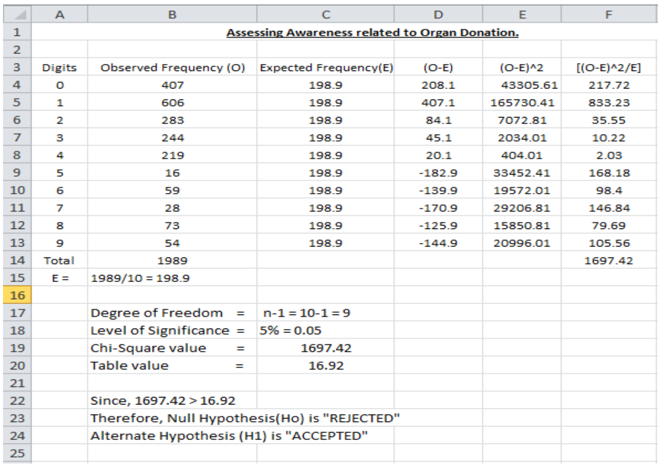
In research paper, author have concluded that there is a need of increasing awareness about benefits of organ transplantation, which include patients living healthier and longer lives along with assurance of safety of the donor and the patient [1]. A planned and a prudent approach need to be adopted to boost organ donation awareness amongst the youth. This might offset some of the setbacks affecting availability of organ donors [2]. Education through advertisements, seminars, media campaigns, exhibitions, conferences and fairs may enhance awareness and improve the attitude of the general public towards organ donation. For example, the well-connected media with the means to reach huge number of people may broadcast important information pertinent to organ donation. Non-governmental organization can help improve knowledge amongst the people and bust myths surrounding deceased organ donation, brain death and the rest of the process [3]. Conservative societies are often defined and dominated by their religious beliefs, such Saudi Arabia. Here, the information should be spread through uniform, well-thought out and religiously inoffensive campaigns to collectively raise the information levels amongst the people who then may be more open to accept organ donation [4]. Unclear rules and regulations and lack of transparency within the medical community make it imperative to develop a centralized organ sharing network to facilitate better co-ordination, reduce organ wastage through timely utilization.

RESEARCH METHODOLOGY

Ho: People are unaware about organ donation.

H1: People are aware about organ donation.

This research was conducted using both Qualitative and Quantitative research approach. Methods are used to prove whether the null hypothesis & alternate hypothesis is accepted or rejected. Primary Data was collected using questionnaire survey with structured question thus the data was collected using Google forms. To prove the hypothesis we took a sampling frame of 80 people. The samples are collected from different age-group and gender. As the data collected are of nominal type, Based on random sampling method we used Chi-Square Test for analysis of nominal data. The reason behind using Chi-Square test was that it requires more than 50 samples data. This test allows to either rejecting the null hypothesis of no relationship at the 0.05 level or due to insufficient evidence to reject the null at the 0.05 level.



	A	B	C	D	E	F
1	Assessing Awareness related to Organ Donation.					
2						
3	Digits	Observed Frequency (O)	Expected Frequency(E)	(O-E)	(O-E)^2	[(O-E)^2/E]
4	0	407	198.9	208.1	43305.61	217.72
5	1	606	198.9	407.1	165730.41	833.23
6	2	283	198.9	84.1	7072.81	35.55
7	3	244	198.9	45.1	2034.01	10.22
8	4	219	198.9	20.1	404.01	2.03
9	5	16	198.9	-182.9	33452.41	168.18
10	6	59	198.9	-139.9	19572.01	98.4
11	7	28	198.9	-170.9	29206.81	146.84
12	8	73	198.9	-125.9	15850.81	79.69
13	9	54	198.9	-144.9	20996.01	105.56
14	Total	1989				1697.42
15	E =	1989/10 = 198.9				
16						
17		Degree of Freedom =	n-1 = 10-1 = 9			
18		Level of Significance =	5% = 0.05			
19		Chi-Square value =	1697.42			
20		Table value =	16.92			
21						
22		Since, 1697.42 > 16.92				
23		Therefore, Null Hypothesis(Ho) is "REJECTED"				
24		Alternate Hypothesis (H1) is "ACCEPTED"				
25						

Figure 1: Chi-Square Test of Independence.

From above data, our Chi-Square value = 1697.42. Table value of χ^2 **test** for 9 Degree of Freedom at 5% level of significance is 16.92. So our calculated value of χ^2 **test** is 1697.42, it is highly significant and Null Hypothesis is rejected at 5% level of significant.

Hence we conclude that **H1** (Alternate Hypothesis) is accepted. That is, **"People are aware about Organ Donation".**

The questionnaire selected for the survey covers the awareness & knowledge people have about Organ Donation.

1. To know the number of people which are aware / unaware about organ donation.
2. Observing Sources of information and association of organs regarding organ donation among people.
3. Willingness of people to be an organ donor.
4. Willingness of people to be a Living donor or Deceased donor.

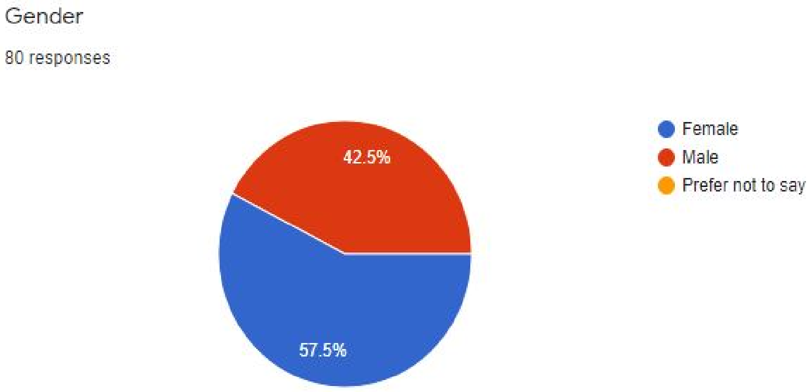


Figure 2: No. of Females aware and No. of Males aware.

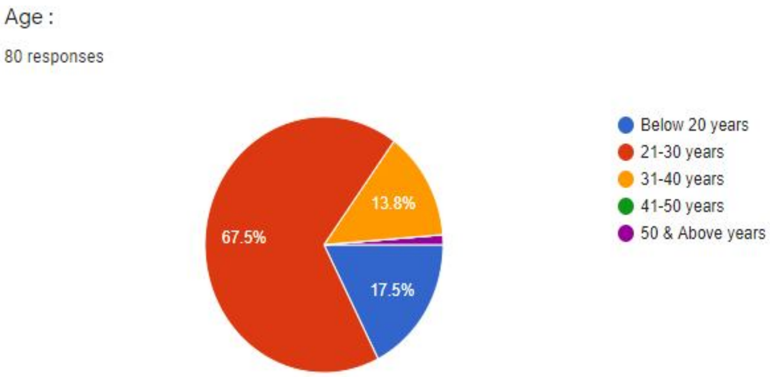


Figure 3: Respondents of different Age Group.

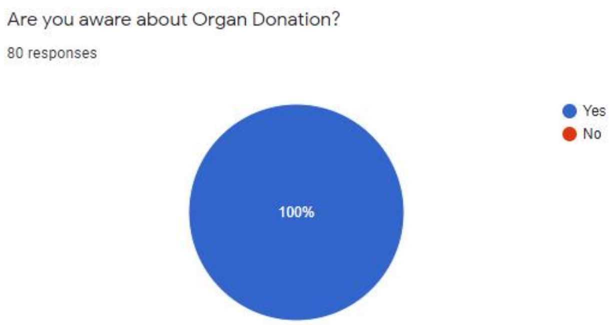


Figure 4: Result of awareness of respondents.

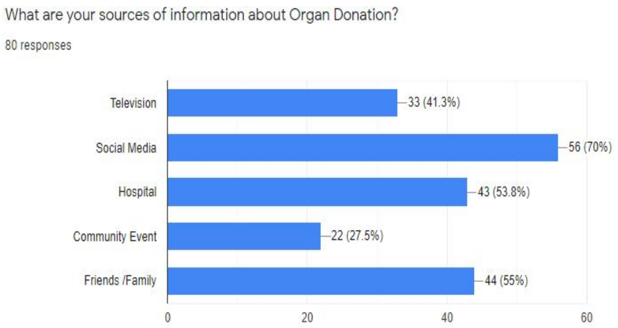


Figure 5: Sources of Information of respondents.

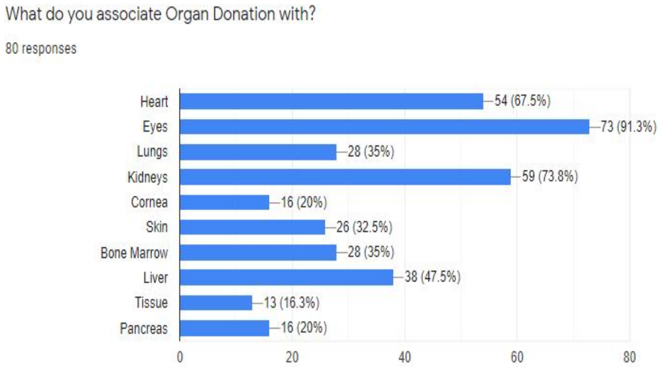


Figure 6: Observing knowledge of respondents regarding awareness of different organs that can be donated.

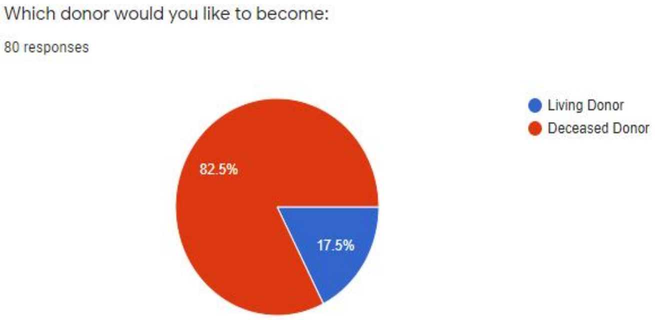


Figure 7: Willingness of a donor to be a Living donor or Deceased donor.

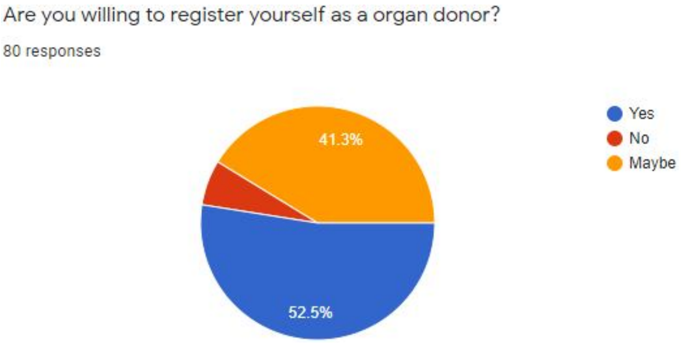


Figure 8: Willingness of the respondents to be an organ donor.

DISCUSSION AND FINDINGS

A total number of 80 respondents participated in the survey, The distribution of the respondents are shown in the above pie charts, The above findings states that more number of females (57.5%)are aware about organ donation than males (42.5%). As the response percentage for the awareness about organ donation is 100% which states that people are aware about organ donation, from which (82.5%) respondents are wish to be an Deceased donor. Willingness of the respondents for becoming an organ donor is also taken into consideration which states that (52.5%) respondents were willing to register themselves as an organ donor, (6.3%) are not at all willing to donate their organs and (41.3%) respondents had possibility that they might or they might not register themselves as an organ donor.

CONCLUSION

In this paper, researchers have presented the significance difference between the awareness / unawareness and willingness / unwillingness of people related to Organ Donation. Based on the above survey researchers stated that there is awareness among the people but as there is tremendous increase in the demand for organs, People needs to be made more aware about organ donation which can help in filling the gap between the organ demand and organ supply. Systematic and continuous efforts are required to raise public awareness and overcome misconception in order to help and save people’s lives. Appropriate knowledge and a positive attitude could play a key role in shaping public opinion. Means, through which maximum communication to the people can be achieved, for example media campaigns and broadcasts, could be instrumental in improving people’s perception towards organ donation.

REFERENCES

[1] Kenneth P. Moritsugu, “The Power of Organ Donation to Save Lives Through Transplantation”, PMC 2013.

[2] Naina Sam, R Ganesh, V Indrapriyadarshini, S Jeyamarthan, CK Nandhini,” Awareness, knowledge, and attitude regarding organ donation”, ISOT 2018.

[3] Senthilkumar Nallusamy, Shyamalapriya, Balaji, Ranjan, Yogendran, “Organ donation – Current Indian scenario”, Journal of practice of cardiovascular science 2018.

[4] Awatif Ali Alam, “Public Opinion on Organ Donation in Saudi Arabia”, Saudi journal of Kidney Diseases and Transplantation 2007.

DETECTING INVISIBLE: WINDOWS REGISTRY ATTACKS

Akhilesh Vas[1] and Milind Meshram[2]
PG Student[1] and Assistant Professor[2], Department of Digital and Cyber Forensics, Institute of Forensic Science, Mumbai

ABSTRACT

This paper shares an awareness on Windows Registryand importance of Registry Analysis process carried in Windows Systems as a part of digital forensic investigation in today's scenario.

Keywords: Windows registry; Forensic Investigation

INTRODUCTION

In the current times as technology advances and the world has higher dependence on Internet, new vulnerabilities and threats are increasing rapidly. A new set of information stealers in multiple variants of malware are focusing on highly sensitive data on a dangerous scale. World Economic Forum reports on Biggest Cybercrime Trends in 2019 states that two billion records were compromised in 2017 and more than 4.5 billion in the first half of 2018. It further states that four new malware samples are created every second.

Another study by AV-TEST GmbH (4), an independent research institute for IT security in Germany, states that AV-TEST Institute registers over 350,000 new malicious programs (malware) and potentially unwanted applications (PUA), every day.

Thus, it becomes very important for the people, law enforcement agencies and Cyber forensic investigators to understand computer systems and be able to examine the threats sitting within the device. Cyber forensics is the branch of science that acts as a tool for the investigators for investigating a computer system or network alleged of being involved in criminal activity and, gathering artifacts that may be used as evidence in the case and presented in the court of law.

This paper aims to share awareness on Windows Registry, importance of Registry Analysis process carried in Windows Systems as a part of digital forensic investigation in today's scenario(10).

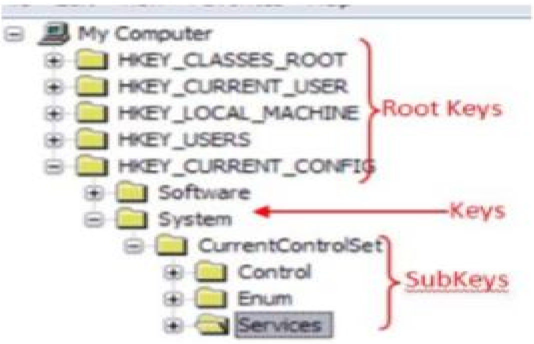
WINDOWS REGISTRY AND STRUCTURE

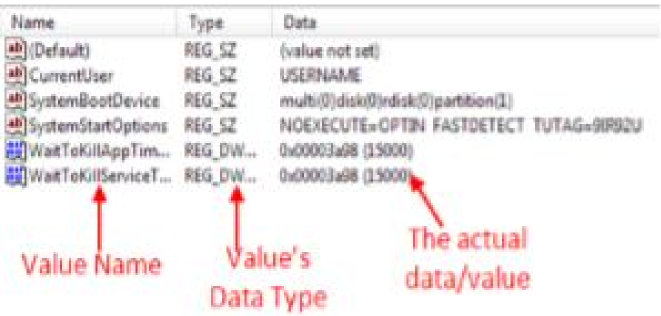
Windows Registry, a central hierarchical database, is used in windows to store data to configure system and options for the 32-bit versions of MS Windows including Windows 95, 98, ME and NT/2000. This encompasses information and settings for all the hardware, software, users, and preferences of the PC. Any changes made by a user in Control Panel setting, or File Associations, System Policies, or installed software, these changes will get reflected and stored in the Windows Registry. Whenever a device is connected to the system, Windows would assign resources to the device based on information available in the Registry and later store the configuration of the devicein the Registry(10).

Registry Structure

Registry holdsdata that Windows constantly references during operation, such as user profiles, the installed applications (on computer) and the kinds of documents that can be created, property sheet settings for folders andapplication icons, hardware that exists on the system, and the ports being used.

The Registry is an organized structure equivalent to the filesystem. For e.g., the keys and subkeys found within the five main hives are comparable to folders and subfolders of Windows filesystem, and a key's value is similar to a file inside a folder, a value's name is analogous to a filename, its type resembles a file extension, and its data is like to the actual contents of a file. The registry structure is shown in figure 1.0(10).



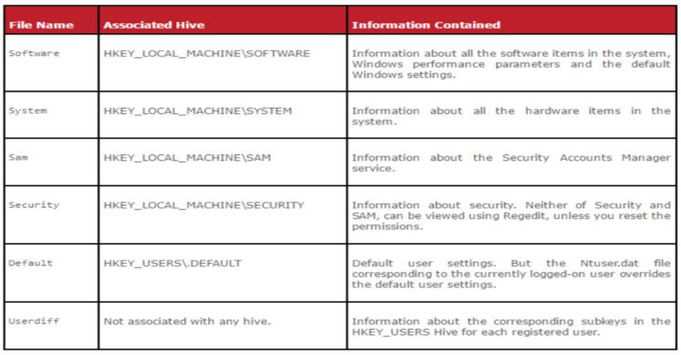


Name	Type	Data
(Default)	REG_SZ	(value not set)
CurrentUser	REG_SZ	USERNAME
SystemBootDevice	REG_SZ	multi(0)disk(0)rdisk(0)partition(1)
SystemStartOptions	REG_SZ	NOEXECUTE=OPTIN FASTDETECT TUTAG=9IR92U
WaitToKillAppTim...	REG_DW...	0x00003a98 (15000)
WaitToKillServiceT...	REG_DW...	0x00003a98 (15000)

List of Predefined keys or hives

- HKEY_LOCAL_MACHINE or HKLM
- HKEY_CURRENT_CONFIG or HKCC
- HKEY_CLASSES_ROOT or HKCR
- HKEY_CURRENT_USER or HKCU
- HKEY_USERS or HKU
- HKEY_PERFORMANCE_DATA (In Windows NT, and invisible in Windows Registry Editor)
- HKEY_DYN_DATA (In Windows 9x, but visible in Windows Registry Editor)

Each value can store arbitrary data with variable length and encryption, associated with a symbolic type (defined as a numeric constant) defining how to parse this data (7,8).For example, a tabular view is given below:



File Name	Associated Hive	Information Contained
Software	HKEY_LOCAL_MACHINE\SOFTWARE	Information about all the software items in the system, Windows performance parameters and the default Windows settings.
System	HKEY_LOCAL_MACHINE\SYSTEM	Information about all the hardware items in the system.
Sam	HKEY_LOCAL_MACHINE\SAM	Information about the Security Accounts Manager service.
Security	HKEY_LOCAL_MACHINE\SECURITY	Information about security. Neither of Security and SAM, can be viewed using Regedit, unless you reset the permissions.
Default	HKEY_USERS\.DEFAULT	Default user settings. But the Ntuser.dat file corresponding to the currently logged-on user overrides the default user settings.
Userdiff	Not associated with any hive.	Information about the corresponding subkeys in the HKEY_USERS Hive for each registered user.

ACQUIREING WINDOWS REGISTRY

Computer forensics tools acquire Windows registry files as a part of acquiring the target machine's system drive or when performing a complete hard drive acquisition.

Process to acquire the Windows machine registry using FTK Imager(14):

Download AccessData FTK Imager and transfer it to USB thumb drive
Attach the USB drive that contains FTK Imager to the suspicious machine
Open FTK Image, and go to File menu; Obtain Protected Files(13)
Store the files obtained; check the option–"Password recovery and all registry files"
Once files are exported get them from the directory where registry files were saved to see the resultant files. One should see the five files and one folder (13)

After successful exported of the target machine registry, as described above, one can use different forensics tools for analyzingit.

Information in the Registry with Forensic Value

For a forensic investigator, the registry is like a bank of information on who, what, where, and when something took place on a system that can directly link the perpetrator to the actions being called into question.

Generally, the following information can be found in the registry(6):

- Users
- Last time the system been
- Most recently used software
- Devices that are mounted on the system including unique identifiers of flash drives or hard drives, phones etc.
- When the systemsare connected to a definite wireless access point
- What files were accessed and when these were accessed
- List of searches done on the system
- And many more

Registry Examination

Wireless Evidence in the Registry

In case, where hackers crack a local wireless access point and use it for their intrusions the IP address should be traced, to lead to the neighbor’s or other wireless AP. The forensic examiner should look in the registry at the following location:

HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\NetworkList\Profiles

The examiner can find a list of GUIDs of wireless access points the machine has been connected to. And when one clicks on one, it discloses the information including the SSID name and the date last connected in hexadecimal(6).

The RecentDocs Key

The Windows registry tracks information about the user's activities. Usually, the registry keys are designed to make Windows run efficiently and smoothly. For a forensic examiner, these keys are like a blueprint of the activities of the user or the attacker.

One of those keys is the "RecentDocs" key, which tracks the most current documents that used or opened on the system by file extension(11). It can be found at:

HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Explorer\RecentDocs

So, for example, the most recently used word documents would be found under .doc or the .docx extension depending upon the version of word they were created in (each key can hold up to the last 10 documents). i.e. “.docx” extension will display the last 10-word documents listed under this key.One can view the document data in hex (to the left) and in ASCII (to the right).

There are some cases, where an attacker will upload a .tar file, so that is a good place to look for breach evidence. Usually, .tar file extension file is not seen on a Windows machine, so the occurrence of such an entry would need further investigation. One should check the files in the .tar key tofind see what they tell about the attack or the attacker.In case of civil or policy violation investigations, evidence can be found in different graphic file extensions like .jpg or .png.

TypedURLs Key

The user can find this value,when the user types URL in Internet Explorer, stored in the registry at:

HKEY_CURRENT_USER\Software\Microsoft\Internet Explorer\TypedURLs

On opening the key, in the registry, it will list the last URLs that user had visited with IE. This can disclose the source of a malicious malware used in the breach(11), or any civil or policy violation types of investigations, this may also reveal what the user was looking for.The values will run until urI25 (the oldest) from urI1 (the most recent).

IP Addresses

Importantly, registry also tracks IP addresses of the user interfaces. Since that there could be many such interfaces so this registry key tracks each interface' IP address and relevant information.

HKEY_LOCAL_MACHINE\System\Services\CurrentControlSet\services\Tcpip\Parameters\Interfaces

So one can find the IP address assigned to the interface, the subnet mask, and the time when the DHCP server leased the IP. This way, we can tell if the suspect was using that IP at the time of crime or not (11).

Start Up Locations in the Registry

A forensic investigator needs to find what applications or services were set to start when the system starts. And malware is set to start each time the system restarts so that the attacker is connected(16). This information can be found in the registry in multiple locations. One should look at some of the most set keys, understandably the most used location is:

HKEY_LOCAL_MACHINE\Software\Microsoft\Windows\CurrentVersion\Run

Each time when the system starts, software/locations designated in these subkeys will also start. One can find the rootkits and other malicious software here, and they will also start when the system starts, each time.

RunOnce Startup

If the hacker just wanted the software to run once at start up, the subkey may be set at:

HKEY_LOCAL_MACHINE\Software\Microsoft\Windows\CurrentVersion\RunOnce

Start Up Services

The key below lists all the services that are set to start at system start-up. When the key is set to 2, the service will start automatically; and if it is set to 3, the service should start manually; whereas if the key is set to 4, the service should be disabled.

HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services

Start When a Particular User Logs On

In case of the following key, the values are run when the user logs in(17).

HKEY_CURRENT_USER\Software\Microsoft\Windows\CurrentVersion\Run

Storage Artifacts in the Registry

Usually, the suspect will use a Flash drive or a hard drive for malicious activities and later remove them ensuring no evidence is left. However, a skilled forensic investigator, can still find traces of evidence of those storage devices within the registry, if they know where and how to look for these.

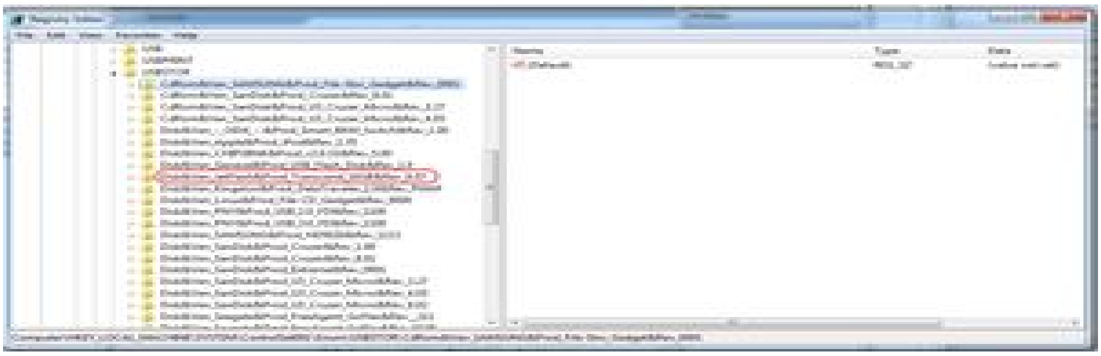
On a Windows system, the registry will vary a bit from version to version. Hence, a skilled and a professional digital forensic examiner should be able to work with almost all versions of Windows and other operating systems.

USB Storage Devices

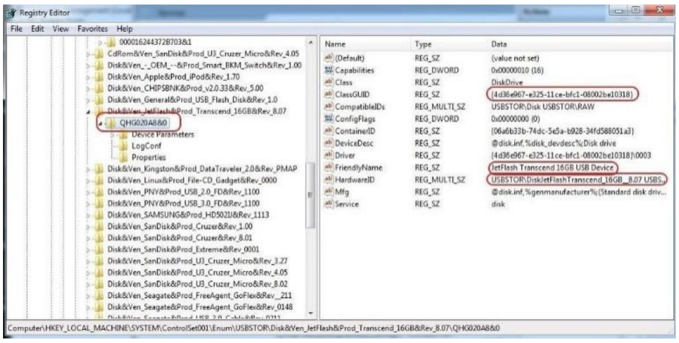
There are instances where one can suspect that someone installed a keylogger or removed confidential information with a USB drive. How would an examiner find evidence that a USB storage device was inserted and used? To find evidence of USB storage devices, look at the following key.

HK_Local_Machine\System\ControlSet00x\Enum\USBSTOR

In this key, one can find evidence of any USB storage device that's ever been connected to this system. And to see the listing of each USB device, ever connected, to this system one need to expand USBSTOR.



In the screenshot above, circled one is suspicious looking USB storage device. It will disclose a unique identifier for that device, when expanded.And simply by a click on this identifier, we will find more information about the device.



000016244372B703&1
CdRom&Ven_SanDisk&Prod_U3_Cruzer_Micro&Rev_4.05
Disk&Ven_-_OEM_--&Prod_Smart_BKM_Switch&Rev_1.00
Disk&Ven_Apple&Prod_iPod&Rev_1.70
Disk&Ven_CHIPSBNK&Prod_v2.0.33&Rev_5.00
Disk&Ven_General&Prod_USB_Flash_Disk&Rev_1.0
Disk&Ven_JetFlash&Prod_Transcend_16GB&Rev_8.07
QHG020A8&0
Device Parameters
LogConf
Properties
Disk&Ven_Kingston&Prod_DataTraveler_2.0&Rev_PMAP
Disk&Ven_Linux&Prod_File-CD_Gadget&Rev_0000
Disk&Ven_PNY&Prod_USB_2.0_FD&Rev_1100
Disk&Ven_PNY&Prod_USB_3.0_FD&Rev_1100
Disk&Ven_SAMSUNG&Prod_HD502IJ&Rev_1113
Disk&Ven_SanDisk&Prod_Cruzer&Rev_1.00
Disk&Ven_SanDisk&Prod_Cruzer&Rev_8.01
Disk&Ven_SanDisk&Prod_Extreme&Rev_0001
Disk&Ven_SanDisk&Prod_U3_Cruzer_Micro&Rev_3.27
Disk&Ven_SanDisk&Prod_U3_Cruzer_Micro&Rev_4.05
Disk&Ven_SanDisk&Prod_U3_Cruzer_Micro&Rev_8.02
Disk&Ven_Seagate&Prod_FreeAgent_GoFlex&Rev__211
Disk&Ven_Seagate&Prod_FreeAgent_GoFlex&Rev_0148

Name	Type	Data
(Default)	REG_SZ	(value not set)
Capabilities	REG_DWORD	0x00000010 (16)
Class	REG_SZ	DiskDrive
ClassGUID	REG_SZ	{4d36e967-e325-11ce-bfc1-08002be10318}
CompatibleIDs	REG_MULTI_SZ	USBSTOR\Disk USBSTOR\RAW
ConfigFlags	REG_DWORD	0x00000000 (0)
ContainerID	REG_SZ	{06a6b33b-74dc-5e5a-b928-34fd588051a3}
DeviceDesc	REG_SZ	@disk.inf,%disk_devdesc%;Disk drive
Driver	REG_SZ	{4d36e967-e325-11ce-bfc1-08002be10318}\0003
FriendlyName	REG_SZ	JetFlash Transcend 16GB USB Device
HardwareID	REG_MULTI_SZ	USBSTOR\DiskJetFlashTranscend_16GB__8.07 USBS...
Mfg	REG_SZ	@disk.inf,%genmanufacturer%;(Standard disk driv...
Service	REG_SZ	disk

Computer\HKEY_LOCAL_MACHINE\SYSTEM\ControlSet001\Enum\USBSTOR\Disk&Ven_JetFlash&Prod_Transcend_16GB&Rev_8.07\QHG020A8&0

In the above screenshot, when one clicks on the USB storage identifier, it discloses the Global Unique Identifier (GUID), the right-hand window. Additionally, the friendly name, and the hardware ID. This could be precise evidence needed to tie the suspects to their activity on this system.

Mounted Devices

In the instances where suspects use any hardware device that must be mounted to read or write data (e.g. CD-ROM; DVD, hard or flash drive etc.), the mounted device will be recorded in the registry. This information will be stored at:

HKEY_LOCAL_MACHINE\System\MountedDevices

This will provide us a list of every device ever mounted on that machine.

For additional information on any of such mounted devices, simply click on it to open a small app and that will enable to read the data in ASCII.

CONCLUSION

Registry is a depository of volumes of information about what happened on a Windows system. Our learning about this can help us reconstruct the elements of a crime that it was used for.

Forensic investigations play animportant role in today's working and legal environments, and hence it should be considered carefully. Evidencesavailable in the registry are the very significant sourcesfor any investigation. The actions performed on the computer device gives a significant insight of the system to an investigator. Awatchful analysis of the Windows Registry from a forensic perspective is the need of the hour. This paper has gathered the available knowledge about the registry hive files with some exhibits. The aim was to highlight the importance of registry analysis and how it can help a forensic investigator to work in a case of tracking data transfer from a system to a USB external device. This document talks about the examination and generation of registry keys of Windows XP systems only and can be extended further for the examination of registry files in Windows Vista, Windows 7 and later versions. It may be noted that it is known now that a thorough analysis of the registry hive files, activities of a system user can be traced. Hence, registry analysis must be considered as an integral part of digital forensic investigation process (10).

REFERENCES

[1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. (references)

[2] Windows Registry Forensics: An Imperative Step in Tracking Data Theft via USB Devices by Tanusree Roy and Aruna Jain from Department of I.T. Birla Institute of Technology, Mesra, Ranchi, India

[3] https://www.weforum.org/agenda/2019/03/here-are-the-biggest-cybercrime-trends-of-2019/

[4] https://www.av-test.org/en/statistics/malware/

[5] https://www.us-cert.gov/sites/default/files/documents/NCCIC_ICS-CERT_AAL_Malware_Trends_Paper_S508C.pdf

[6] www.btk-clan.ch and https://www.hackers-arise.com/post/2016/10/21/digital-forensics-part-5-analyzing-the-windows-registry-for-evidence

[7] en.m.wikipedia.org

[8] https://support.microsoft.com/en-us/help/256986/windows-registry-information-for-advanced-users

[9] https://www.hackers-arise.com/post/2016/10/21/digital-forensics-part-5-analyzing-the-windows-registry-for-evidence

[10] https://netseedblog.com/security/windows-registry-forensics-investigating-the-registry-for-evidence/

[11] pdfs.semanticscholar.org

[12] Submitted to Champlain College Student Paper

[13] www.scribd.com

[14] https://netseedblog.com/security/windows-registry-forensics-investigating-the-registry-for-evidence/

[15] http://www.ijcttjournal.org/

[16] Submitted to Republic Polytechnic Student Paper

[17] secufacile.info

[18] safeatlast.co

OVERVIEW OF WEB APPLICATION SECURITY ATTACKS: FORENSIC INVESTIGATION

Needa Ashraf Petkar
Information Security Associate

ABSTRACT

Web applications are popular sites for attackers or hackers to target and launch new attacks. One of the reasons behind this is the inability to trace back an attack and find the criminal responsible for it. Mitigation of a security attack on the web application is possible by using specific security mechanisms to either prevent it or detect it. A web application security attack can be traced back to its originator using web application forensics. It may significantly reduce the attacks and give us insights on improving the security. The aim of this paper is to provide an overview for the web application forensics with the help of tools and a suggestive methodology. A comparative study of the tools has also been performed.

Keywords: Digital Forensic, Forensic Investigation, Web Application Security Attack.

I. INTRODUCTION

Web Applications are a new profound way in government agencies, enterprises, individuals and others to transmit and manage information. Some of the examples include e-banking, e-commerce, e-learning, e-mail, e-medicine, and highly critical areas of trading, marketing, banking and so on. Hence, they become the popular targets for the security attackers.

Web Applications have a variety of dependencies such as operating systems, web browsers, network infrastructure, web servers and database servers [1]. As the number of dependencies and its variety increase, the type of vulnerabilities that the application will be exposed to, will also increase[6],[9]. Figure 1 portrays the interactions of different components of a web application and the spot where the vulnerabilities may affect it.

From the figure, it can be assumed that the web application provides a motivating environment to develop various methods to perform a security attack. Some of which are Buffer Overflow, Code Injection, Cross-Site Scripting, SQL Injection, etc[9]. Various measures have been developed and implemented to mitigate the attacks, some of which include firewalls, IDS, IPS, Antivirus, etc[9][12].

Depending on the attack detection schemes, one can detect the attack but finding the person responsible for it, remains a challenge. Our inability to trace, allows the attacker to conceal and launch new attacks. Hence, it becomes highly critical to develop the capability to track the real Cyber criminals.

Web Application Forensics can be defined as the branch of digital forensics to trace back a web application security attack in order to identify and determine the origin of attack, its propagation and the responsible device(s) and people(s)[4],[6],[16],[7].

In order to trace back the security attack, a forensic investigator has to rely on the hacker's fingerprints (digital evidence). This can be found recorded in the different configuration and log files of the web server(s) and application server(s), server side scripts, any third party installed software logs and the operating system logs on which the web application is based[1],[7].

Sometimes, these files lack data needed to conduct the comprehensive forensic investigation. To overcome this issue, forensic tools are to be used to obtain the much needed digital evidence. Some of the evidence can be provided by the operating system or network forensics tools, that have additional logging facilities [7].

Several techniques have already been proposed to effectively perform web application forensics and manage various sources of digital evidence, providing an efficient analysis of the data. The following tools were used to investigate a web application security attack : Analog, CORE Wisdom, EventLog Analyzer, HTTP-analyze, Lire, Logjam, Microsoft LogParser, Mywebalizer, Open Web Analytics, Pyflag, Sawmill, and so on. Besides using a forensic tool, it is very important to follow a standard methodology for a successful forensic investigation.

Through this paper, we will provide an overview for the web application forensic investigation, which will help to track the hackers / cyber criminals.

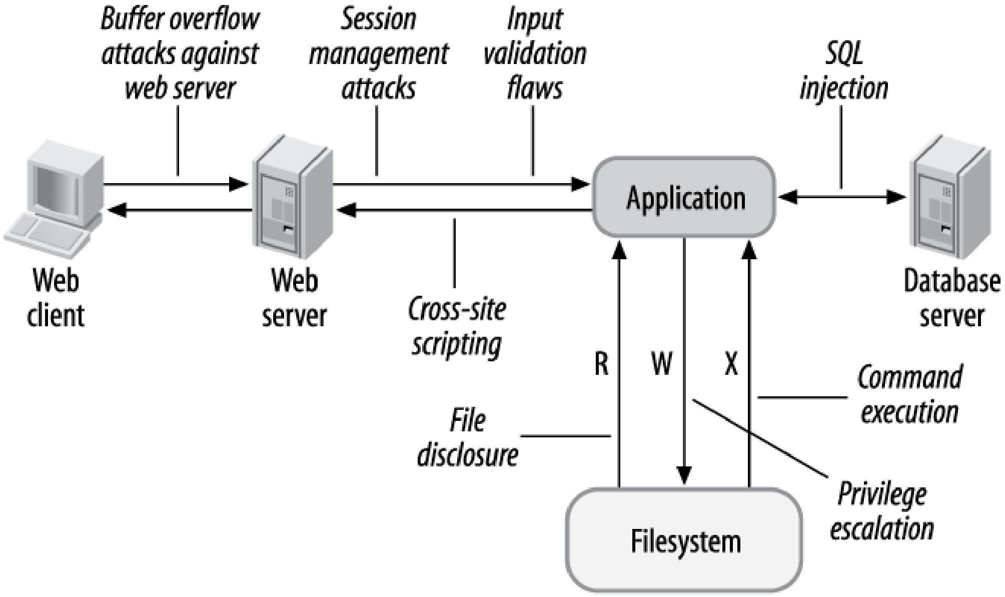


Figure 1: Web Application Architecture and Common Attacks.

II. WEB APPLICATION FORENSICS

Tracing back a security attack and attributing to its originator is the main aim of Web Application Forensics[10],[16]. And thus, it has become a specific branch of digital forensics alongside other branches such as Digital Image Forensics, Network Forensics and Operating System Forensics[4],[7].

Analysis of log files of different components of web application is crucial for proper investigation[1]. It can be further improved by examining the log of different network equipment such as Firewalls, IDS, Routers, Switches, etc. Thus network forensics plays a supportive role in this investigation. Operating System Forensics helps us understand if any system alteration had taken place. Whereas, Digital Image Forensics will provide insights on the image manipulations [4].

It should be noted that Web Application Forensics does not deal with security attacks on web services but it is covered in Web Services Forensics which is a whole other branch of digital forensics[7]. Since Cloud-Computing comprises both web services and web application, Cloud Computing Forensics becomes a unique branch of digital forensics which integrates web services forensic as well as web application forensics[7],[2],[15].

The Digital Forensics has four main branches, namely, Cloud Computing Forensics, Digital Image Forensics, Network Forensics and Operating System Forensics. Where, the Cloud Computing Forensics is further divided into two subdivisions, namely, Web Application Forensics and the Web Services Forensics. This taxonomic structure is shown below in figure 2.

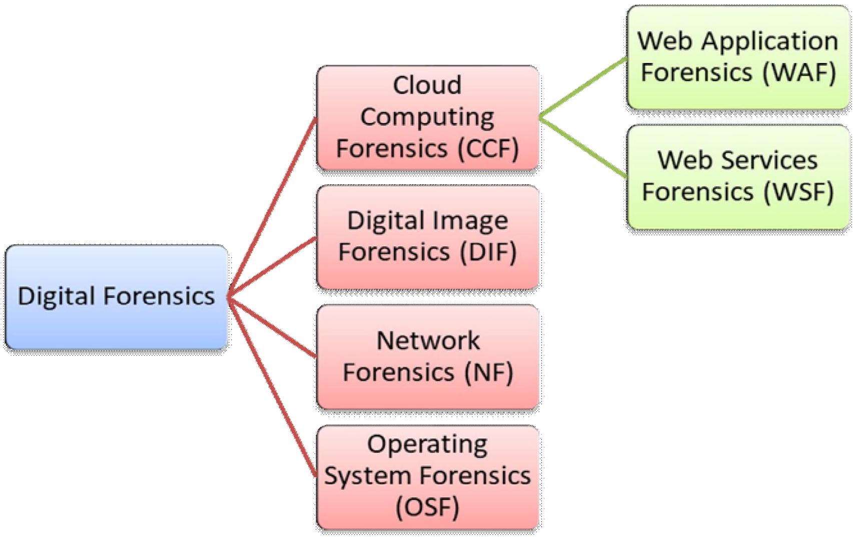

Digital Forensics
Cloud Computing Forensics (CCF)
Digital Image Forensics (DIF)
Network Forensics (NF)
Operating System Forensics (OSF)
Web Application Forensics (WAF)
Web Services Forensics (WSF)

Figure 2: Taxonomic Structure of Digital Forensics

III. FORENSICS INVESTIGATION OF A WEB APPLICATION ATTACK

Preliminary analysis phase plays the most important role in successful forensics investigation which has to be followed by a standard protocol or methodology[7],[13]. In this section, we will look into the preliminary actions and steps that should be taken by the forensics investigator to conduct a thorough analysis of a hacking instance. We will also discuss how other branches of digital forensics support our investigation.

A. Preliminary Actions for Analysis

Mentioned below are some of the preliminary steps that have to be taken.

1. Application Forensics Readiness:

The web application should be well prepared for a forensics investigation. This is can be reached by:

- Evidence collection: Check the logging options in the web application. If it is at default settings, the evidence collection will be incomplete and the application will not be fully ready for the investigation. Hence, it is recommended to enable logging options in the web application to collect the maximum amount of evidence.
- Evidence protection: Since, we have realized that the log files will constitute the main source of digital evidence, it becomes highly critical to maintain the integrity of the data they contain as well as guarantee the accuracy of the digital evidence that they provide. And to protect such files, the following actions are being recommended:

a) Log files must have proper permissions set.

b) Log files must be kept out of reach. This can be achieved by using a backup utility that will save the files on the server which is present remotely.

c) Verify the integrity of the log files using checksum.

2. Supportive forensics

Once the web application is ready for forensics, it does not always guarantee the existence of all the evidence required for the investigation. For this, we require some of the forensic tools that will acquire all the needed evidence that might have been missed by the logging options. Such evidence can be acquired using network or operating system forensic tools or a third party tools offering additional logging facilities.

3. Forensics investigator abilities

Even if we have an excellent array of tools to obtain the evidence, an investigation cannot be fruitful, if the forensic investigator lacks the skill. Thus, an investigator must have at least the following attributes:

- Understand the components of web application: architecture, components, flow of the application, etc.
- Understand the security issues present in the web applications: vulnerabilities, security attack methods, etc.
- Must be well acquainted with forensic techniques and must be trained for forensic investigation.

B. Methodology

Once the web application is made ready for the web application, the forensic investigation has to be conducted in a specific method or a standard protocol. A forensic investigator should follow the following steps.

1. Prevent any modification of the evidence files by protecting the web application(can be many servers too), during the forensic examination.
2. Discover and make a note of the files that will be required for the investigation, which may includes

- Logs of Web Server(s), Application Server(s), Operating System and third party installed softwares, if any.
- Server Side Scripts of the web application.

3. Determine the sequence of the events and the degree of compromise. This can be ascertained by:

- Dividing the log files based on the user sessions to:
  ➢ Understand the flow of the session and the timeline.
  ➢ Remove noise created by other users, if any.
- Find and understand the fingerprints and patterns of the hacking attempts on the web application, such as:
  ➢ Excessive attempts from the same IP address.

➢ Files created or modified during the suspected time of attack.

➢ Unusual entries in the logs (GET requests made to a page designed to receive POST request.)

➢ Abuse of scripts such as CMD.exe, Root.exe, Upload.ASP, etc.

➢ Unusual processing times (SQL injection attempt).

4. Proper report creation based purely on the extracted data.

5. Recommendations for the post event actions.

C. Supportive Forensics

Digital Image Forensics, Network Forensics and Operating System forensics provides additional or supportive evidence during the Web Application forensics investigation. Log data obtained from an intrusion detection system (IDS), provides more accurate detection of an intruder’s activities on the web application [10],[7]. Forensic investigations of the digital images uploaded during the compromise of the web application, can assist in tracking the origin of the attack as well as the intruder[7].

Also, it is important to note that the cache memory of a hacked web application server which has not been restarted during or after the attack scenario will prove useful to consider the attack even if there is lack of sufficient evidence in the log files[7], [3].

IV. Tools for Web Application Forensics

As discussed above, it can be seen that the amount of logged data that has to be examined can get overwhelming for the forensic investigator. Thus, automated tools can be used to take the burden off from the shoulders of the forensic investigators. But the tools have to comply with the standards and some requirements which we will discuss below.

A. Requirements for a Web Application Forensics Tools

Some of the basic requirements for a tool to be called as a web application forensics tool is mentioned below:

1. Analyze log files in different formats.
2. Combine two independent and differently formatted evidence files.
3. Decode URL data into readable format.
4. Handle big log files.
5. Perform normalization by time to perform investigation depending on time-stamp.
6. Maintain a list of suspicious requests, indicating a potential compromise.
7. Utilize binary logic and regular expressions on any observed parameter present in the log file.

B. Web Application Forensics Tools

The following forensic tools were observed and analyzed for the research papers.

1. Analog

It is an open source web log file analyzer that accepts AWS or IIS W3C formatted files as input. It produces complex graphs and report styles with the help of Report Magic, which is a different tool[1]. The basic functionality of these tools is to statistically report and graphically represent the information. It generates general summaries, and reports based on time, user, browser, host, domain, organization, etc. The validity of the retained information is maintained only when the required server configuration is performed. Thus, an investigator or an analyst requires extra configuration to obtain a total account of the information.

2. CORE Wisdom

It generates unique graphical log analysis reports besides pie charts and graphs in real-time. It enhances the evidentiary report but does provide any correlation ability. The analyst has to define rules for importing the log file itself. Hence, the analyst needs to be able to define events to flag them as alarms and know what the visual cues will be.

3. EventLog Analyzer

It is a real-time web-based log monitoring and compliance management solution for SIEM (Security Information and Event Management) to improve the internal network security[8]. It has the potential to collect, analyze, search, report and archive an extensive array of machine generated from applications (Apache, Oracle, IIS, etc), network devices (routers, switches, etc), systems (windows, linux, unix, etc) and provide informational insights into internal threats, network anomalies, network user activities, policy violations, system downtime, etc.

It can be used to generate archive files, so as to store them and conduct analysis later. It can define automatic alerts, generate historical trends based on system alerts, group host information together to show interactions, show failed logins-malicious users and identify applications that are causing performance or security issues.

Reports can be exported to HTML, PDF and Comma Separated Values (CSV) formats, at specific intervals, having both graphs and text-based representations as output. It includes pre built reports and provides flexibility to choose data and format for generation of custom reports. It does not automatically correlate between log files, which seems to be a drawback.

4. Http-analyze

It is a multiplatform log file analyzer that processes data in CLF, ELF, DLF file formats[5]. It provides an option to generate one of two different HTML standardized reports which includes graphs, tabulated data, three dimensional forms, statistical and access load information summaries. Real-time analysis can be achieved by rotation of the scripts of log files, in conjunction with the automated calling of this tool. But it does not generate or format log file information. It does not store information into a database, nor perform correlation between web server files with any other information that might be available.

5. Lire

It provides analysis for a variety of log file types by converting them into DLF format[15]. It has thirteen templates for the statistical analysis report, which can be customized or modified by the analyst. It does not perform in real-time and does not provide correlation between log files.

6. Logjam

It is a web traffic analyzer providing statistical analysis of W3C ELF log files. But it performs only on a MS Windows Server that includes Active Server Pages (ASP) and Microsoft Data Access Components (MDAC). The log files have to be put into SQL database for analysis. It has a customisable report generator based on user preferences using SQL queries. It does not perform in real-time nor does correlation.

7. Microsoft LogParser

It is a flexible command line utility providing universal query access to CSV files, log files (Web Server, DNS, HTTP error), TSV files, W3C files, XML files and sources from Windows OS such as Active Directory, Event Log File System and Registry[1],[7]. It produces output in standardized formats (CSV, IIS, SQL, Syslog, TSV, W3C, XML, etc) and non-standardized formats (graphical output, DATAGRID, CHART, NAT, etc). It has been used to monitor user activities, system file integrity, check SQL Injection attacks, detect failed login attempts, determine malicious modification, identify brute force attack, etc.

It does not provide GUI but functions through command line by the way of script or manipulation of queries via prompt. To provide GUI, two programs have been developed, namely, LogParser Lizard and Visual LogParser. The LogParser language provides functions to perform string manipulation, arithmetic operations, provide access to system details, modify or manipulate contents of fields. It can also perform analysis including correlation. For this, it has the capability to combine data from different sources and perform queries on it. It does not include a method of analysis and thus the user has to create queries to satisfy the analysis requirements.

8. Mywebalizer

It is an open source, C language programmed tool that generates highly detailed tabular and graphical HTML reports about Web Server usage statistics[17]. It is feasible with UNIX, LINUX, AIX, Solaris, etc. It analyses CLF, Xferlog and Extended W3C log files along with bzip2 and gzip compressed log files. It does not run in real-time or provide correlation.

9. Open Web Analytics

It is a generic web analytic framework that can function on any operating system and can be added to web applications using JS, PHP or REST application programming interfaces[11]. It provides built-in support for WordPress or MediaWiki applications. It also provides real-time tracking, monitoring and reporting of web usage statistics about browser information, visitor click streams, geolocation of visitors, etc.

10. Pyflag

It is an open source application that allows forensic analysis of log files through GUI and includes querying, sorting and graphical representation of the log data. It handles large log files with different formats, images, disks, network traffic data such as Tcpdump data. It can be also added in MySQL database but the analyst must have prior experience and knowledge to perform the required analysis.

11. Sawmill

It is a sophisticated log file analyzer that has plugins for detecting over 800 log file types and also, for the non-standard log file types[14]. It also has command line interface, database, log importers, reporting interface, scheduler, Web Server and language to manipulate and store data for analysis. It also provides correlation between log files from different source files and in real-time. For forensics, external methods are used to protect the original data.

Salang, the first language, is used to define log filters, regular expressions, conditional logic and display pages. Structures Query Language, the second language, is used to access the internal database information within predefined table sets.

C. Web Application Forensics Tools Comparison

Comparison between the above mentioned tools is summarized in the table below. The parameters that were used for the comparison were:

1. Data Sources Compression – tool's ability to compress the considered log files.
2. Data Sources Correlation – tool's ability to correlate different evidence sources.
3. Multiple Platforms – tool's ability to be performed over various operating systems.
4. Real-time Performing – tool's ability to perform a real-time analysis log files.
5. Reporting – tool's ability to generate an analysis report.
6. Scalability – tool's ability to continue to function well as its context is changed in size or volume.

Tools	Data Sources Compression	Data Sources Correlation	Multiple Platform	Real-time Performing	Reporting	Scalability
Analog	No	No	Yes (FreeBSD, Linux, OS X, Unix, Windows)	No (Requires Scripting)	Yes (HTML, Statistics)	Yes
CORE Wisdom	No	No	No (Windows)	Yes	Yes (Enhanced Graphics)	Yes
EventLog Analyzer	No	No	Yes (Requires a Web Server)	Yes	Yes (CSV, HTML, PDF)	Yes
HTTP-Analyze	Yes (Rotation)	No	Yes (Interpreted Language OS Independent)	No (Requires Scripting)	Yes (HTML)	No
Lire	No	No	Yes (FreeBSD, Linux, OS X, Unix)	No	Yes (HTML, XML)	Yes
Logjam	No	Yes	No (Windows)	No	Yes (HTML)	No
Microsoft LogParser	No	No	No (Windows)	No	Yes (CSV, Syslog, TSV, XML)	Yes
Mywebalizer	Yes	Yes	Yes (FreeBSD, Linux, OS X, Unix, Windows)	No	Yes (HTML)	Yes
Open Web Analytics	No	No	Yes (FreeBSD, Linux, OS X, Unix, Windows)	Yes	Yes (HTML)	Yes
Pyflag	No	Yes	Yes (FreeBSD, Linux, OS X, Unix, Windows)	No	Yes (HTML)	Yes
Sawmill	Yes (Extracting Selections of logs)	Yes	Yes (FreeBSD, Linux, OS X, Unix, Windows)	Yes	Yes (HTML, XML)	Yes

Table 1: Comparison of Web Application Forensics Tools

V. CONCLUSION

In this paper, we understood the dynamics of web application forensics. We determined the appropriate steps to accomplish a successful forensic investigation of web application security attack. For which, we conducted study for various available tools and compared them to make a catalog of their pros and cons, which will enable an investigator to choose the best tool. We also ascertained that Digital Image Forensics, Network Forensics and Operating System Forensics will prove useful for providing additional evidence.

REFERENCES

1. A. Fry, A Forensic web Log Analysis Tool: Techniques and implementation, Thesis dissertation, Department of Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Canada, September 2011, website: http://spectrum.library.concordia.ca/7769/1/Fry_MASc_F2011.pdf. Last accessed in January 2020.
2. D. Birk, Forensic Identification and Validation of Computational Structures in Distributed Environments, 2010.
3. E. Weippl, Database Forensics, Proceedings of the 24th IEEE International Conference on Advanced Information Networking and Applications (AINA), 2010, Perth, WA.
4. F. N. Dezfoli, A. Dehghantanha, R. Mahmoud, N. FBinti M. Sani, and F. Daryabar, Digital Forensic Trends and Future, International Journal of Cyber-Security and Digital Forensics (2): 48-76. The Society of Digital Information and Wireless Communications, 2013 (ISSN: 2305-0012).
5. HTTP-ANALYZE, website: http://http-analyze.org/index.php. Last Accessed in January 2020.
6. I. Ristic, Apache security, O’Reilly Media, Inc., pub-ORA-MEDIA: adr, 2005. Section 1.1.2
7. K. Deltchev, Web Application Forensics:Taxonomy and Trends, term paper, Horst Görtz Institute, September 2011, website: http://fr.slideshare.net/test2v/web-application-forensics-taxonomy-and-trends. Last Accessed in January 2020.
8. ManageEngine EventLog Analyzer: Help Documentation, last accessed in January 2020, website: https://download.manageengine.com/products/eventlog/eventloganalyzer-userguide.pdf.
9. M. Shema, Hacking Web Apps, Publisher: Syngress, Pub. Date: October 2012, Print ISBN-13: 978-1-59749-951-4, Web ISBN-13: 978-1-59749-956-9, Pages in Print Edition: 296.
10. N. Meghanathan, S. R. Allam and L. A. Moore, Tools and Techniques for Network Forensics, International Journal of Network Security & Its Applications (IJNSA), Vol .1, No.1,April 2009.
11. P. Adams. Open Web Analytics - Main Page. http://www.openwebanalytics.com/. Last Accessed in January 2020.
12. Prof. Dr. D. Basin, Dr. P. Schaller, and M. Schläpfer, Web Application Security, Applied Information Security, 2011, ISBN 9783642244735, pp. 81– 101.
13. R. Belani, C. Willis, Web Application Incident Response & Forensics: A Whole New Ball Game!, 2007.
14. Sawmill, Sawmill: Universal log file analysis and reporting, website: http://www.sawmill.net/. Last accessed in January 2020.
15. S. Y. Hashemi, and P. S. Hesarlo, Security, Privacy and Trust Challenges in Cloud Computing and Solutions, I.J. Computer Network and Information Security, 2014, 8, 34-40.
16. V. Kumar, A. P. Singh, A. K.. Rai, M. Wairiya, Self Alteration Detectable Image LogFile for Web Forensics, in International Journal of Computer Applications, 2011.
17. Webalizer, web site: http://www.webalizer.org/. Last accessed in January 2020.

ISSN 2394 - 7780

RANSOMWARE ATTACKS FORENSIC INVESTIGATION

Aditya Ramesh Ogania
UG Student, Department of Digital and Cyber ForensicsInstitute of Forensic Science, Mumbai

ABSTRACT

This paper is a research of technology use in ransomware(Latest Technology based malware attacks to disrupt the business processes and make whole IT infrastructure and resources making out of action or useless). This paper try to understand Ransomware attack i.e. cryptography technology aspect, the detection methodology the modusoperandi of different attacker different type of ransomware malware, crypto currency used for demanding ransom money its detection method and some recommendation for enterprises and home uses

Keywords: Ransomware; Forennsic Investigation ;cryptocurrency ;

I. INTRODUCTION

In the current

This paper is a research of technology use in ransomware(Latest Technology based malware attacks to disrupt the business processes and make whole IT infrastructure and resources making out of action or useless). This paper try to understand Ransomware attack i.e. cryptography technology aspect, the detection methodology the modusoperandi of different attacker different type of ransomware malware, crypto currency used for demanding ransom money its detection method and some recommendation for enterprises and home uses

II. CRYPTOGRAPHY

Cryptography : Cryptography is the art & science of using mathematics to encrypt and decrypt data or information. With modern technological Advancement, is begin permeate in all facets of everyday life.

This techniques focuses for secure communication between two parties in presences of third party.

Example, It turns plain text into Cipher text making it encrypt which can be decrypt by key same can be decrypt the cipher text to plain text .

There are 4 basic principle of cryptography.

-Confidentiality

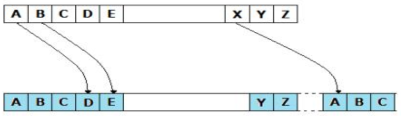
-Data Integrity.

-Authentication.

-Non Repudiation.

History of Cryptography [1] : The first known evidence of the use of cryptography was found in an inscription carved around 1900 BC, in the main chamber of the tomb of the nobleman Khnumhotep II, in Egypt. The purpose was not to hide the message but perhaps to change its form in a way which would make it appear dignified. Though the inscription was not a form of secret writing, but incorporated some sort of transformation of the original text, and is the oldest known text to do so. Evidence of some use of cryptography has been seen in most major early civilizations.

Julius Caesar was known to use a form of encryption to convey secret messages to his army generals posted in the war front. This substitution cipher, known as Caesar cipher, is perhaps the most mentioned historic cipher in academic literature. In a substitution cipher, each character of the plain text is substituted by another character to form the cipher text . The variant used by Caesar was a shift by 3 cipher. Each character was shifted by 3 places, so the character 'A' was replaced by 'D', 'B' was replaced by 'E', and so on. The characters would wrap around at the end, so 'X' would be replaced by 'A'.



Algorithms

Each value can store arbitrary data with variable length and encoding, but which is associated with a symbolic type defining how to parse this data. A quick tabular view is given below:

ISSN 2394 - 7780

Triple DES [2]

Triple DES was designed to replace the original Data Encryption Standard (DES) algorithm, which hackers eventually learned to defeat with relative ease. At one time, Triple DES was the recommended standard and the most widely used symmetric algorithm in the industry.

Triple DES uses three individual keys with 56 bits each. The total key length adds up to 168 bits, but experts would argue that 112-bits in key strength is more like it.

RSA

RSA is a public-key encryption algorithm and the standard for encrypting data sent over the internet. It also happens to be one of the methods used in our PGP and GPG programs.

Unlike Triple DES, RSA is considered an asymmetric algorithm due to its use of a pair of keys. Public key, which is use to encrypt message, and a private key to decrypt it.

Blowfish

Blowfish is yet another algorithm designed to replace DES. This symmetric cipher splits messages into blocks of 64 bits and encrypts them individually.

Blowfish is known for both its tremendous speed and overall effectiveness as many claim that it has never been defeated. Meanwhile, vendors have taken full advantage of its free availability in the public domain.

Twofish

Computer security expert Bruce Schneier is the mastermind behind Blowfish and its successor Twofish. Keys used in this algorithm may be up to 256 bits in length and as a symmetric technique, only one key is needed.

Twofish is regarded as one of the fastest of its kind, and ideal for use in both hardware and software environments. Like Blowfish, Twofish is freely available to anyone who wants to use it.

AES

The Advance Encryption Standard AES is the algorithm trusted as the standard by tnumerous organizations. Although it is extremely efficient in 128-bit form, AES also uses keys of 192 and 256 bits for heavy duty encryption purposes.

AES is largely considered impervious to all attacks, with the exception of brute force, which attempts to decipher messages using all possible combinations in the 128, 192, or 256-bit cipher.

Ransomware : Ransomware is type of vicious malware attack. It characterized by taking devices control and locally stored data or information for ransom. Victims pays ransom only in crypto-currencies. Ransomware attackers employ disk or file level encryption which is impossible to recover because of technique used to generate ransomware it uses public key infrastructure(Asymmetric cryptography). Ransomware attacks causes data loss, intellectual property theft and some time data breach. Even after paying the ransom the attackers doesn't guarantee the decrypting key. It is also been noticed that some global companies have to permanently shut off their business due to a single ransomware attack on companies infrastructure

Cyber-criminals uses Ransomware as a very easy way for generating financial funds to profit as compare to other crime like identity theft which requires more resources compare to ransomware attack.

TYPES OF RANSOMWARE

GandCrab [3]

In the year 2018 and mid 2019 GandCrab created massive hype in globally. As per the information these ransomware having multiple variants and generated illegal ransomware financial funds up to more than $150 million USD which is 11,00,00,00,000.00 INR In period of 18 months.

EKANS [4]

Ekans cause disruption in Industrial Control System (ICS) as a side effect of malware destructive acitivity. A relatively primitive attack, the EKANS ransomware actively targets certain products common in ICS environments.

Ryuk Ransomware [5]

Ryuk disrupt the entire IT network. Ryuk is capable of systems that monitor and control the transfer of cargo,ultimately knocking the facility primary operations for 30 hours during the incident response.

Players : There are many players and groups involved into Ransomware attacks including Governments sponsor groups to organize criminal groups and in some cases terrorist groups also.

Organised Hacker group

Anonymous : Anonymous has an extensive details of victims from Goverment institutions to multinational companies. It is decentralized hacking group for instance group have targeted or hacked Visa, Mastercard, Paypal in 2012. they have also jump into action during occupy wallstreet by hacking the New York Stock Exchange website

Lizard Squad : Lizard Squad has claimed the responsibility for two major attacks. First and biggest of them is their DDOS attack on Facebook. Second was an attack on Malaysia Airlines

Syrian Electronic Army : The Syrian Electronic Army sympathizes with syria and has shown support for Syrian President Bashar al-Assad. This group targets organization that are against to the state of syria. This group use spamming, defacement,malware,phishing and denial of service attacks. The group has targeted the Facebook and Twitter account of president Barack Obama and French President Nicolas Sarkozy, technology companies and new outlets. E.G. group tweeted from BBC Weather:"Saudi weather station down due to head on – collision with camel"

Terrorist Group : A Cyber attack that handicapped Sarasota City's Hall Computers in 2016. This ransomware caused encrypting 160,000 data or information. Demanded ransom was up to $33 million USD in Crypto Currency.

The Incident was so critical event which involved Islamic State, Russian hacking and FBI. Sarasota police department criminal investigation and city staff handling of the virus after the attack taken place.

Individual Dark-net user

MAKSIM VIKTOROVICH YAKUBETS[6]

Maksim Viktorovich Yakubets is wanted for his involvement with computer malware that infected tens of thousands of computers in both North America and Europe, resulting in actual financial losses in the tens of millions of dollars.

Specifically, Yakubets was involved in the installation of malicious software known as "Zeus", which was disseminated through phishing emails and used to capture victims' online banking credentials. These credentials were then used to steal money from the victims' bank accounts. On August 22, 2012, an individual was charged in a superseding indictment under the moniker "aqua" in the District of Nebraska with conspiracy to participate in racketeering activity, conspiracy to commit computer fraud and identity theft, aggravated identity theft, and multiple counts of bank fraud. On November 14, 2019, a criminal complaint was issued in the District of Nebraska that ties the previously indicted moniker of "aqua" to Yakubets and charges him with conspiracy to commit bank fraud.

Yakubets is also allegedly the leader of the Bugat/Cridex/Dridex malware conspiracy wherein he oversaw and managed the development, maintenance, distribution, and infection of the malware. Yakubets allegedly conspired to disseminate the malware through phishing emails, to use the malware to capture online banking credentials, and to use these captured credentials to steal money from the victims' bank accounts. He, subsequently, used the malware to install ransomware on victims' computers. Yakubets was indicted in the Western District of Pennsylvania, on November 13, 2019, and was charged with Conspiracy, Conspiracy to Commit Fraud, Wire Fraud, Bank Fraud, and Intentional Damage to a Computer.

FUJIE WANG

On May 7, 2019, a grand jury in the United States District Court for the Southern District of Indiana, Indianapolis Division, indicted two individuals for conspiracy to commit fraud and related activity in connection with computers, conspiracy to commit wire fraud, and causing intentional damage to a protected computer.

The subjects, including Fujie Wang, were alleged members of a hacking group operating in China that conducted intrusion campaigns targeting the computer systems of large businesses in the United States, including a large health benefits company in Indiana. It is alleged that, between February of 2014 and January of 2015,the subjects conspired to intentionally access computer networks to identify and ultimately steal data concerning approximately 78.8 million persons from computer networks, including names, health identification numbers, dates of birth, Social Security numbers, addresses, telephone numbers, email addresses, employment information, and income data. Once the information was collected, it was allegedly placed in encrypted files and sent to destinations in China.

ISSN 2394 - 7780

Financial Losses

Internationally most of the countries economic suffer had a deep impacts due to multiple ransomware attacks. The impact is so deep that in some countries companies have to shut of there business. Ransomware attack is also identified as one of the source of funding to terrorism. This all transaction are done on cryptocurrencies like bitcoin which are very difficult to track back. Total financial lost to US for year 2019 due to ranso mware attack amounted to more than $120 millions USD.

CryptoCurrency

Cryptocurrency is a currency associated with the internet that uses cryptography, the process of converting legible information into an almost uncrackable code, to track purchases and transfers. cryptocurrency is based on block chain technology. The blockchain is a simple yet ingenious way of passing information from A to B in a fully automated and safe manner. One party to a transaction initiates the process by creating a block. This block is verified by thousands, perhaps millions of computers distributed around the net. The verified block is added to a chain, which is stored across the net, creating not just a unique record, but a unique record with a unique history. Falsifying a single record would mean falsifying the entire chain in millions of instances. That is virtually impossible. Bitcoin uses this model for monetary transactions.

Bitcoin : Bitcoin is a type of cryptocurrency. It is most commonly traded cryptocurrency to date. The currency was developed in 2009 by Satoshi Nakamoto, Who developed its blockchain. In todays market Bitcoin price is $10,355.21 USD per bitcoin.

Ethereum : Ethereum has a native cryptocurrency called Ether, which is ETH Digital money. The supply of ETH is not controlled by any government or company. It is decentralize and it is scarce. In todays market Ethereum price is $266.66 USD per ethereum.

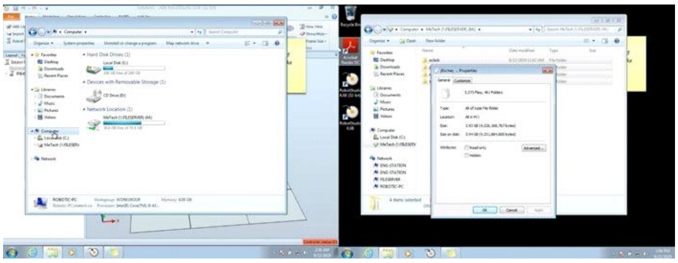
Due to Features of Crypto currenies like privacy and anonymity tracking back of transaction is very difficult and thats why crypto currency is used for ransom money.

Forensic Investigation

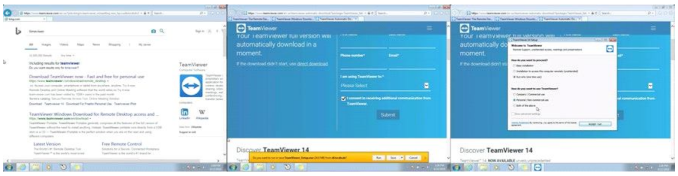
Crysis Ransomware

[4]The first ransomware attack we encountered happened in late September. We were able to document the entire duration of the attack. We also responded to the attackers, still posing as our organization, to gain further insight into how similarly threat actors might conduct their deals.

On Sept. 22, a threat actor began looking around our system. Typical of threat actors, they first investigated the system, likely looking for important and sensitive files. They looked at a few items such as the shared drive. Their next actions were to close the robotics workstation application and to go back to the shared drive to see how much information was on it.



After these initial actions on our system, they then downloaded the remote desktop software TeamViewer.In fact, they opened Bing and searched for “timeviwer” to do so. Then they ran the TeamViewer installer.They chose to run TeamViewer only once and chose the option to use it for personal use.


Download TeamViewer now - Fast and free for personal use
TeamViewer Windows Download for Remote Desktop access and ...
Latest Version
Free Remote Control
TeamViewer
Your TeamViewer full version will automatically download in a moment.
If the download didn't start, use direct download
Phone number*
Email*
I am using TeamViewer to:*
Please Select
Submit
Discover TeamViewer 14
Run
Save
Cancel

Once they started connecting to our system using TeamViewer, we lost further keystrokes from the PCAPs. This did not stop us from monitoring this attack, however, especially since at this point the threat actor had started to take more interesting actions. They then transferred three files over TeamViewer, which included the ransomware file:

1btc.exe

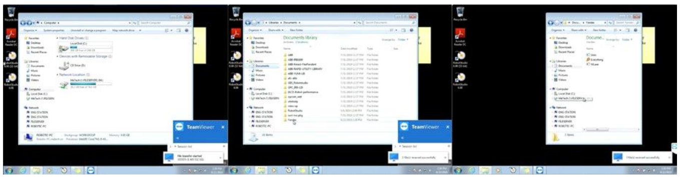
- The ransomware file, a variant of Crysis
- Detected by Trend Micro as Ransom.Win32.CRYSIS.SM 19
- SHA1: ddf8c065d45c734b5b58e770e4f1ea086a293f19
- First submission from VirusTotal: 2019-07-24 10:14:26 UTC

Everything.exe

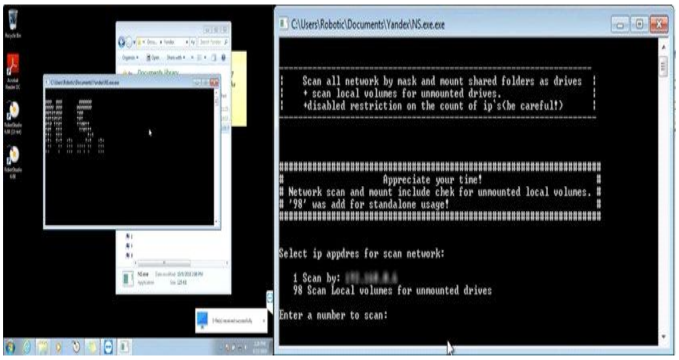
- A normal application that lists all files on a file system. It allows an attacker to check whether a system is already infected by another piece of ransomware using the search function.
- SHA1: c8107e5c5e20349a39d32f424668139a36e6cfd0

NS.exe

- A tool used to scan mounted and unmounted physical and network drives. Its ability to scan unmounted drives makes it very effective for ransomware attacks.
- Detected by Trend Micro as HackTool.Win32.NetTool.A 20
- SHA1: 629c9649ced38fd815124221b80c9d9c59a85e74
- It is highly similar to a sample analyzed by Hybrid Analysis.

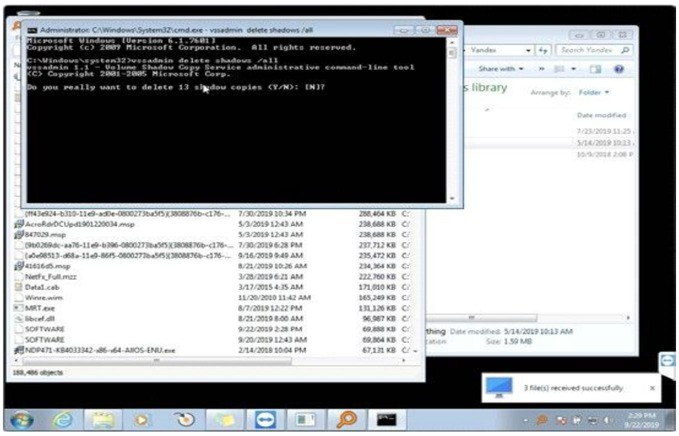


After downloading the files, they connected to the system using the computer name "X555DG" with a TeamViewer ID of "1 405 532 321". They then started transferring the files to the Documents library under the subfolder they named Yandex. After this point, the threat actor began running each of the downloaded files, beginning with NS.exe, the tool used to scan for mounted and unmounted drives. Next, they ran the Everything.exe file as an administrator. While this was running, they opened a command window and typed in the command "vssadmin delete shadows /all", which is commonly used in ransomware attacks. Finally, they ran the 1btc.exe file, the Crysis ransomware variant, as an administrator. We were able to record all of these activities.



C:\Users\Robotic\Documents\Yandex\NS.exe.exe

```
-----------------------------------------------------------------------
|    Scan all network by mask and mount shared folders as drives     |
|    + scan local volumes for unmounted drives.                      |
|    +disabled restriction on the count of ip's(be careful!)         |
-----------------------------------------------------------------------



#######################################################################
#                         Appreciate your time!                       #
# Network scan and mount include chek for unmounted local volumes.   #
# '98' was add for standalone usage!                                  #
#######################################################################


Select ip appdres for scan network:

   1 Scan by: [redacted]
   98 Scan Local volumes for unmounted drives

Enter a number to scan:
```



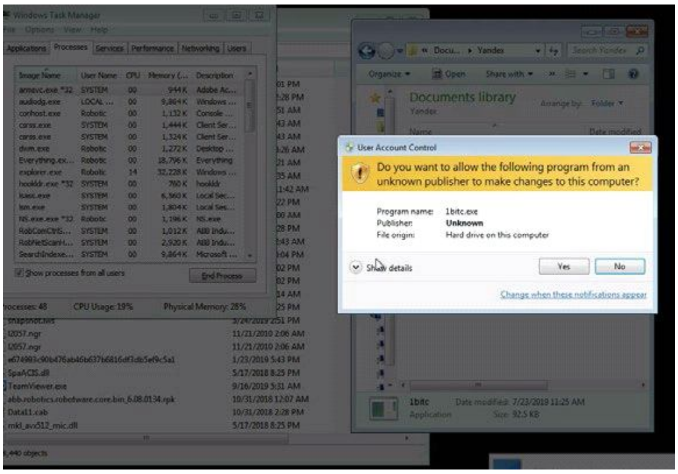
Administrator: C:\Windows\System32\cmd.exe - vssadmin delete shadows /all

```
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation.  All rights reserved.

C:\Windows\system32>vssadmin delete shadows /all
vssadmin 1.1 - Volume Shadow Copy Service administrative command-line tool
(C) Copyright 2001-2005 Microsoft Corp.

Do you really want to delete 13 shadow copies (Y/N): [N]?
```

{ff43e924-b310-11e9-ad0e-0800273ba5f5}{3808876b-c176-...	7/30/2019 10:34 PM	288,464 KB	C:
AcroRdrDCUpd1901220034.msp	5/3/2019 12:43 AM	238,688 KB	C:
847029.msp	5/3/2019 12:43 AM	238,688 KB	C:
{9b0269dc-aa76-11e9-b396-0800273ba5f5}{3808876b-c176-...	7/30/2019 6:28 PM	237,712 KB	C:
{a0e98513-d68a-11e9-86f5-0800273ba5f5}{3808876b-c176-...	9/16/2019 9:49 AM	235,472 KB	C:
41616d5.msp	8/21/2019 10:26 AM	234,364 KB	C:
NetFx_Full.mzz	3/28/2019 6:21 AM	222,760 KB	C:
Data1.cab	3/17/2015 4:35 AM	171,010 KB	C:
Winre.wim	11/20/2010 11:42 AM	165,249 KB	C:
MRT.exe	8/7/2019 12:22 PM	131,126 KB	C:
libcef.dll	8/21/2019 8:00 AM	96,987 KB	C:
SOFTWARE	9/22/2019 2:28 PM	69,888 KB	C:
SOFTWARE	9/20/2019 12:43 AM	69,864 KB	C:
NDP471-KB4033342-x86-x64-AllOS-ENU.exe	2/14/2018 10:04 PM	67,131 KB	C:

188,486 objects

ISSN 2394 - 7780



Windows Task Manager
File Options View Help
Applications Processes Services Performance Networking Users
Show processes from all users
End Process
Processes: 48 CPU Usage: 19% Physical Memory: 28%
Docu... ▸ Yandex
Organize ▾ Open Share with ▾
Documents library
Yandex
Arrange by: Folder ▾
Show details
Yes
No
Change when these notifications appear
1bitc
Application
Date modified: 7/23/2019 11:25 AM
Size: 92.5 KB
...440 objects

Image Name	User Name	CPU	Memory (...	Description
armsvc.exe *32	SYSTEM	00	944 K	Adobe Ac...
audiodg.exe	LOCAL ...	00	9,864 K	Windows ...
conhost.exe	Robotic	00	1,132 K	Console ...
csrss.exe	SYSTEM	00	1,444 K	Client Ser...
csrss.exe	SYSTEM	00	1,324 K	Client Ser...
dwm.exe	Robotic	00	1,272 K	Desktop ...
Everything.ex...	Robotic	00	18,796 K	Everything
explorer.exe	Robotic	14	32,228 K	Windows ...
hookldr.exe *32	SYSTEM	00	760 K	hookldr
lsass.exe	SYSTEM	00	6,560 K	Local Sec...
lsm.exe	SYSTEM	00	1,804 K	Local Ses...
NS.exe.exe *32	Robotic	00	1,196 K	NS.exe
RobComCtrlS...	SYSTEM	00	1,012 K	ABB Indu...
RobNetScanH...	SYSTEM	00	2,920 K	ABB Indu...
SearchIndexe...	SYSTEM	00	9,864 K	Microsoft ...

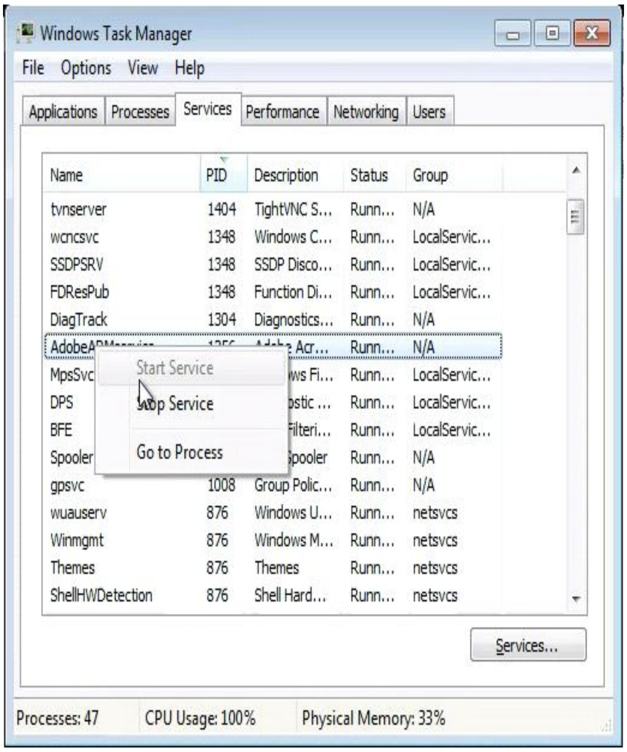
snapshot.tws	3/24/2019 2:51 PM
l2057.ngr	11/21/2010 2:06 AM
l2057.ngr	11/21/2010 2:06 AM
e674983c90b476ab46b637b6816df3db5ef9c5a1	1/23/2019 5:43 PM
SpaACIS.dll	5/17/2018 8:25 PM
TeamViewer.exe	9/16/2019 5:31 AM
abb.robotics.robotware.core.bin_6.08.0134.rpk	10/31/2018 12:07 AM
Data11.cab	10/31/2018 2:28 PM
mkl_avx512_mic.dll	5/17/2018 8:25 PM

User Account Control

Do you want to allow the following program from an unknown publisher to make changes to this computer?

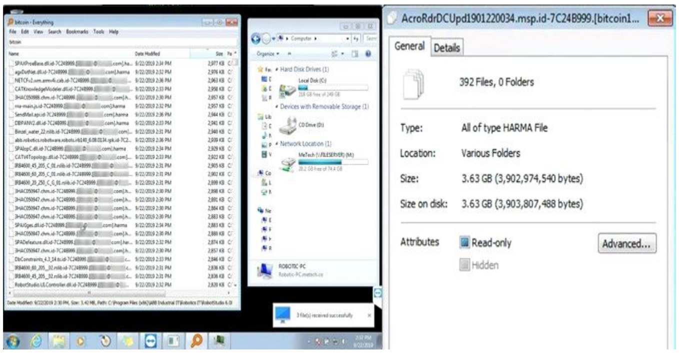
Program name: 1bitc.exe
Publisher: Unknown
File origin: Hard drive on this computer

After setting all of these in motion, the threat actors watched and waited by opening the task manager. They even stopped other services to give their activities more processing power, They then checked the result of their work by looking at all of the files listed in Everything, the otherwise legitimate tool used for listing files on a file system. the ransomware seemed to have successfully affected the files in our system. The threat actors even looked at a particular file (AcroRdrDCupd1901220034.msp.id-7C24B999) and checked its properties to confirm that the ransomware had worked.



Name	PID	Description	Status	Group
tvnserver	1404	TightVNC S...	Runn...	N/A
wcncsvc	1348	Windows C...	Runn...	LocalServic...
SSDPSRV	1348	SSDP Disco...	Runn...	LocalServic...
FDResPub	1348	Function Di...	Runn...	LocalServic...
DiagTrack	1304	Diagnostics...	Runn...	N/A
AdobeA...		...Acr...	Runn...	N/A
MpsSvc		...ws Fi...	Runn...	LocalServic...
DPS		...ostic ...	Runn...	LocalServic...
BFE		...Filteri...	Runn...	LocalServic...
Spooler		...Spooler	Runn...	N/A
gpsvc	1008	Group Polic...	Runn...	N/A
wuauserv	876	Windows U...	Runn...	netsvcs
Winmgmt	876	Windows M...	Runn...	netsvcs
Themes	876	Themes	Runn...	netsvcs
ShellHWDetection	876	Shell Hard...	Runn...	netsvcs

Processes: 47 CPU Usage: 100% Physical Memory: 33%

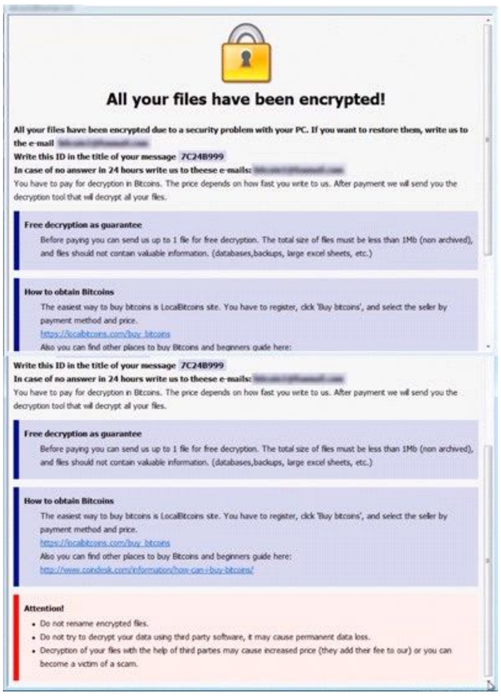


AcroRdrDCUpd1901220034.msp.id-7C24B999.[bitcoin1...

392 Files, 0 Folders

Type: All of type HARMA File
Location: Various Folders
Size: 3.63 GB (3,902,974,540 bytes)
Size on disk: 3.63 GB (3,903,807,488 bytes)
Attributes ■ Read-only □ Hidden

Finally, with their work done, they closed TeamViewer. A ransom message then popped up, containing the typical content like the contact details of the threat actor, how to pay them in bitcoin, and the usual warning not to attempt to tamper with the encrypted files.



All your files have been encrypted!

All your files have been encrypted due to a security problem with your PC. If you want to restore them, write us to the e-mail [redacted]
Write this ID in the title of your message 7C248999
In case of no answer in 24 hours write us to theese e-mails: [redacted]
You have to pay for decryption in Bitcoins. The price depends on how fast you write to us. After payment we will send you the decryption tool that will decrypt all your files.

Free decryption as guarantee
Before paying you can send us up to 1 file for free decryption. The total size of files must be less than 1Mb (non archived), and files should not contain valuable information. (databases,backups, large excel sheets, etc.)

How to obtain Bitcoins
The easiest way to buy bitcoins is LocalBitcoins site. You have to register, click 'Buy bitcoins', and select the seller by payment method and price.
https://localbitcoins.com/buy_bitcoins
Also you can find other places to buy Bitcoins and beginners guide here:

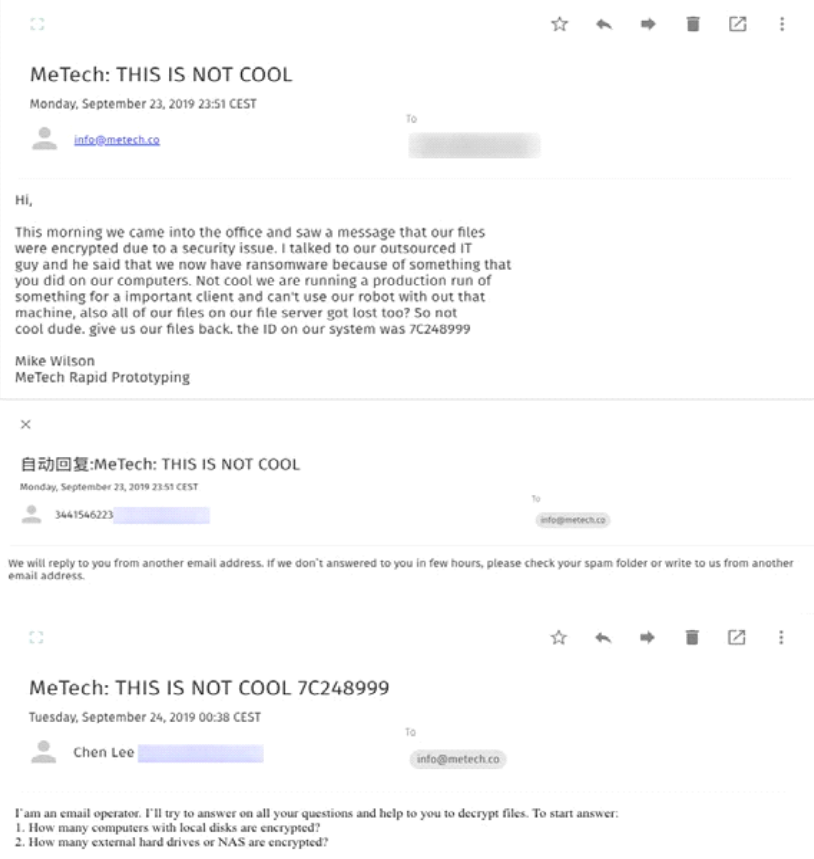
Write this ID in the title of your message 7C248999
In case of no answer in 24 hours write us to theese e-mails: [redacted]
You have to pay for decryption in Bitcoins. The price depends on how fast you write to us. After payment we will send you the decryption tool that will decrypt all your files.

Free decryption as guarantee
Before paying you can send us up to 1 file for free decryption. The total size of files must be less than 1Mb (non archived), and files should not contain valuable information. (databases,backups, large excel sheets, etc.)

How to obtain Bitcoins
The easiest way to buy bitcoins is LocalBitcoins site. You have to register, click 'Buy bitcoins', and select the seller by payment method and price.
https://localbitcoins.com/buy_bitcoins
Also you can find other places to buy Bitcoins and beginners guide here:
http://www.coindesk.com/information/how-can-i-buy-bitcoins/

Attention!
- Do not rename encrypted files.
- Do not try to decrypt your data using third party software, it may cause permanent data loss.
- Decryption of your files with the help of third parties may cause increased price (they add their fee to our) or you can become a victim of a scam.

An actual company, upon realizing that its files have been encrypted and reading the ransom note, would have to go through several decision-making processes to handle such a situation. In our case, still posing as our cover company, we emailed the threat actor using the contact information they had left behind. Our first email was meant simply to engage the threat actor behind the provided email address. The reply we received was an obviously automated response, and it was followed a day later by an email asking for further details



MeTech: THIS IS NOT COOL
Monday, September 23, 2019 23:51 CEST
info@metech.co
To [redacted]

Hi,

This morning we came into the office and saw a message that our files were encrypted due to a security issue. I talked to our outsourced IT guy and he said that we now have ransomware because of something that you did on our computers. Not cool we are running a production run of something for a important client and can't use our robot with out that machine, also all of our files on our file server got lost too? So not cool dude. give us our files back. the ID on our system was 7C248999

Mike Wilson
MeTech Rapid Prototyping

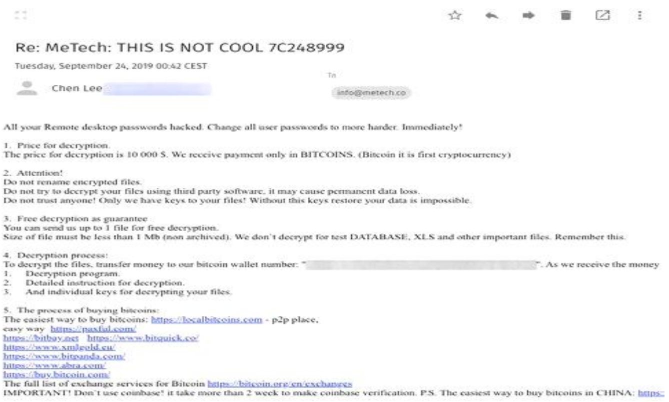
自动回复:MeTech: THIS IS NOT COOL
Monday, September 23, 2019 23:51 CEST
3441546223 [redacted]
To info@metech.co

We will reply to you from another email address. If we don't answered to you in few hours, please check your spam folder or write to us from another email address.

MeTech: THIS IS NOT COOL 7C248999
Tuesday, September 24, 2019 00:38 CEST
Chen Lee [redacted]
To info@metech.co

I'am an email operator. I'll try to answer on all your questions and help to you to decrypt files. To start answer:
1. How many computers with local disks are encrypted?
2. How many external hard drives or NAS are encrypted?

We responded to the email shown in the last image by saying that one computer and one file server were affected in the attack. The next email from the threat actor contained a list of instructions and, more significantly, their demand for US$10,000 worth of bitcoin in exchange for having our files returned to normal, to be transferred to their wallet address, also specified in the email.



Re: MeTech: THIS IS NOT COOL 7C248999

Tuesday, September 24, 2019 00:42 CEST

Chen Lee [redacted]

To info@metech.co

All your Remote desktop passwords hacked. Change all user passwords to more harder. Immediately!

1. Price for decryption.
The price for decryption is 10 000 $. We receive payment only in BITCOINS. (Bitcoin it is first cryptocurrency)

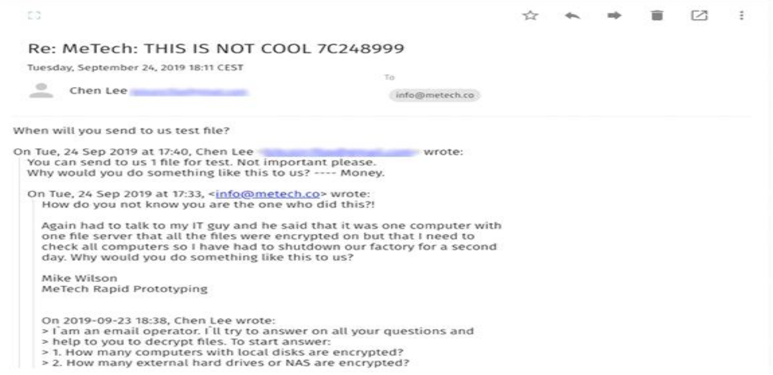
2. Attention!
Do not rename encrypted files.
Do not try to decrypt your files using third party software, it may cause permanent data loss.
Do not trust anyone! Only we have keys to your files! Without this keys restore your data is impossible.

3. Free decryption as guarantee
You can send us up to 1 file for free decryption.
Size of file must be less than 1 Mb (non archived). We don`t decrypt for test DATABASE, XLS and other important files. Remember this.

4. Decryption process:
To decrypt the files, transfer money to our bitcoin wallet number: "[redacted]". As we receive the money
1. Decryption program.
2. Detailed instruction for decryption.
3. And individual keys for decrypting your files.

5. The process of buying bitcoins:
The easiest way to buy bitcoins: https://localbitcoins.com - p2p place,
easy way https://paxful.com/
https://bitbay.net https://www.bitquick.co/
https://www.xmlgold.eu/
https://www.bitpanda.com/
https://www.abra.com/
https://buy.bitcoin.com/
The full list of exchange services for Bitcoin https://bitcoin.org/en/exchanges
IMPORTANT! Don`t use coinbase! it take more than 2 week to make coinbase verification. P.S. The easiest way to buy bitcoins in CHINA: https:

In response, we sent an email asking them to decrypt a file as an example, to make sure that they did in fact have the decryption key. During this part of our exchange, we acted the part of a disgruntled company representative asking why the threat actor was doing this in the first place. They answered succinctly and obliged us by decrypting a sample file. We sent them the conveyor belt PLC programing file (Omron CXP file), which they decrypted accordingly, suggesting that they were unaware that we had in fact sent them an important file. After resending the decrypted file, they reiterated their demand for and preferred mode of payment.



Re: MeTech: THIS IS NOT COOL 7C248999

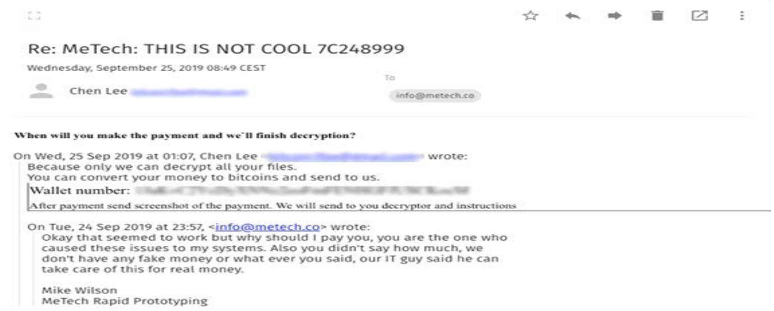
Tuesday, September 24, 2019 18:11 CEST

Chen Lee [redacted]

To info@metech.co

When will you send to us test file?

On Tue, 24 Sep 2019 at 17:40, Chen Lee [redacted] wrote:
> You can send to us 1 file for test. Not important please.
> Why would you do something like this to us? ---- Money.
>
> On Tue, 24 Sep 2019 at 17:33, <info@metech.co> wrote:
>> How do you not know you are the one who did this?!
>>
>> Again had to talk to my IT guy and he said that it was one computer with one file server that all the files were encrypted on but that I need to check all computers so I have had to shutdown our factory for a second day. Why would you do something like this to us?
>>
>> Mike Wilson
>> MeTech Rapid Prototyping
>>
>> On 2019-09-23 18:38, Chen Lee wrote:
>> > I`am an email operator. I`ll try to answer on all your questions and
>> > help to you to decrypt files. To start answer:
>> > 1. How many computers with local disks are encrypted?
>> > 2. How many external hard drives or NAS are encrypted?



Re: MeTech: THIS IS NOT COOL 7C248999

Wednesday, September 25, 2019 08:49 CEST

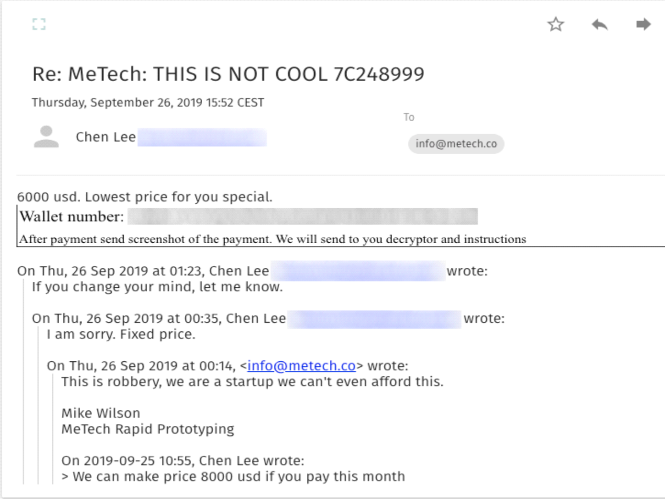
Chen Lee [redacted]

To info@metech.co

When will you make the payment and we`ll finish decryption?

On Wed, 25 Sep 2019 at 01:07, Chen Lee [redacted] wrote:
> Because only we can decrypt all your files.
> You can convert your money to bitcoins and send to us.
> Wallet number: [redacted]
> After payment send screenshot of the payment. We will send to you decryptor and instructions
>
> On Tue, 24 Sep 2019 at 23:57, <info@metech.co> wrote:
>> Okay that seemed to work but why should I pay you, you are the one who caused these issues to my systems. Also you didn't say how much, we don't have any fake money or what ever you said, our IT guy said he can take care of this for real money.
>>
>> Mike Wilson
>> MeTech Rapid Prototyping

We continued the exchange by attempting to haggle. Ultimately, we managed to reduce their price to US$6,000 worth of bitcoin from the original US$10,000.



Re: MeTech: THIS IS NOT COOL 7C248999

Thursday, September 26, 2019 15:52 CEST

Chen Lee [redacted]

To info@metech.co

6000 usd. Lowest price for you special.
Wallet number: [redacted]
After payment send screenshot of the payment. We will send to you decryptor and instructions

On Thu, 26 Sep 2019 at 01:23, Chen Lee [redacted] wrote:
> If you change your mind, let me know.
>
> On Thu, 26 Sep 2019 at 00:35, Chen Lee [redacted] wrote:
>> I am sorry. Fixed price.
>>
>> On Thu, 26 Sep 2019 at 00:14, <info@metech.co> wrote:
>>> This is robbery, we are a startup we can't even afford this.
>>>
>>> Mike Wilson
>>> MeTech Rapid Prototyping
>>>
>>> On 2019-09-25 10:55, Chen Lee wrote:
>>> > We can make price 8000 usd if you pay this month

When this attack had run its course, we simply reset the system after getting all the information we could.[7]

Preventing Ransomware Attacks:

1. Watch out for such known Files Extensions as list which are known ransomware extension shoot up. Extesion like \.enc|\.R5A|\.R4A|\.encrypt|\.locky|\.clf|\.lock|\.cerber|\.crypt|\.txt|\.coverton|\.enigma|\.czvxce|\.{CRYPTENDBLACKDC}|\.scl|\.crinf|\.crjoker|\.encrypted|\.code|\.CryptoTorLocker2015!|\.crypt|\.ctbl|\.html|\.locked| \.ha3|\.enigma|\.html|\.cry|\.crime|\.btc|\.kkk|\.fun|\.gws|\.keybtc@inbox_com| \.kimcilware.LeChiffre|\.crime|\.oor|\.magic|\.fucked|\.KEYZ|\.KEYH0LES|\.crypted|\.LOL!|\.OMG!|\.EXE|\.porno|\.RDM|\.RRK| \.RADAMANT|\.kraken|\.darkness|\.nochance|\.oshit|\.oplata@qq_com|\.relock@qq_com|\.crypto|\.helpdecrypt@ukr|\.net|\.pizda@qq_com| \.dyatel@qq_com_ryp|\.nalog@qq_com| \.chifrator@qq_com|\.gruzin@qq_com|\.troyancoder@qq_com|\.encrypted|\.cry| \.AES256|\.enc|\.hb15|\.vscrypt|\.infected|\.bloc|\.korrektor|\.remind|\.rokku|\.encryptedAES|\.encryptedRSA| \.encedRSA|\.justbtcwillhelpyou|\.btcbtcbtc|\.btc-help-you| \.only-we_canhelp_you|\.sanction|\.sport|\.surprise|\.vvv|\.ecc|\.exx|\.ezz|\.abc|\.aaa|\.zzz|\.xyz|\.biz|\.micro|\.xxx|\.ttt|\.mp3|\.Encrypted| \.better_call_saul|\.xtbl|\.enc|\.vault|\.xort|\.trun|\.CrySiS|\.EnCiPhErEd|\.73i87A|\.p5tkjw|\.PoAr2w|\.xrtn|\.vault|\.PORNO

2. Watch out for an increase in file renames File renames are not a common action when it comes to activity on network file shares. Over the course of a normal day, you may end up with just a handful of renames even if you have hundreds of users on your network. When Ransomware strikes, it will result in a massive increase in file renames as your data gets encrypted.

3. Create a sacrificial network share can act as an early warning system and also delay the Ransomware from getting to your critical data.

4. Update IDS System.

5. Use Client based Anti-Ransomware Agents.

III. CONCLUSION

The process

REFERENCES

[1] www.access.redhat.com/blogs/766093/posts/1976023

[2] blog.storagecraft.com/5-common-encryption-algorithms/

[3] www.darkreading.com/attacks-breaches/fbi-publishes-gandcrab-decryption-keys/d/d-id/1335258

[4] www.darkreading.com/attacks-breaches/ekans-ransomware-raises-industrial-control-worries/d/d-id/1336950

[5] www.darkreading.com/threat-intelligence/ryuk-ransomware-hit-multiple-oil-and-gas-facilities-ics-security-expert-says-/d/d-id/1336865

[6] https://www.fbi.gov/wanted/cyber/maksim-viktorovich-yakubets

[7] https://documents.trendmicro.com/assets/white_papers/wp-caught-in-the-act-running-a-realistic-factory-honeypot-to-capture-real-threats.pdf

[8] https://www.netfort.com/blog/methods-for-detecting-ransomware-activity/

ISSN 2394 - 7780

ANALYSIS OF MALWARE IN VIRUSTOTAL AND RISK ASSOCIATED WITH DISCLOSURE OF PERSONAL DATA

Tanushree Suresh Pai[1] and Neeta Khobragade[2]
[1]Department of Digital and Cyber Forensics, Institute of Forensic Science, University of Mumbai, Mumbai
[2]Head & Assistant Professor, Digital and Cyber Forensics, Institute of Forensic Science, Mumbai

ABSTRACT

Malware attacks have been ever increasing with increasing digitization. It has become important to develop the ability to forensically analyze the systems victimized by these attacks to prevent further damage due to data leaks. There is need to improve our detection capabilities with the rise of second generation malware which are using anti-forensic techniques to evade detection. With addition to antivirus software on individual system, even online malware analysis tools are being introduced to help users detect threats they are vulnerable to because of being connected to internet. The paper aims to study the analysis mechanism and results provided by the online tools, especially concentrating on a tool called VirusTotal.

Keywords: Malware, static analysis, VirusTotal, PE file, malware retrohunting, WINJA tool

INTRODUCTION

With increasing digitization, and rising over dependence on technology, users already have all or majority of their data in their mobile phones and computers. This has made it easier for criminals to misappropriate the technology and compromise the security features of the systems to gain unauthorized access to user data, and the most widespread method of doing so is by initiating a malware attack against the target system.

LITERATURE REVIEW

Malware

Malware or malicious software is any software that is created and transmitted with the intent of causing harm or detriment to the user of the system, the computer system or the network. A common feature of all the malwares is that they enter and spread throughout the system without user's informed consent or knowledge. The main goals that the attackers or the creators of such software have is to either disrupt the functioning of the system or to gain private or sensitive information with the objective of committing crimes such as financial frauds, espionage, data theft among other illegal activities. Malware is a general term used to refer to a variety of forms of hostile or intrusive software. It may also appear in the form of code, scripts or active contents. Malwares include Trojans, spyware, ransomware, viruses, worms, keyloggers, dialers, rootkits, malicious Browser Hijacking Objects (BHO'S) and other malicious programs. In order to protect our systems and information as well as to prevent further spread of the malware we need to perform malware analysis.

Malware Analysis

Malware analysis can be defined as the art of dissecting malware to understand how it works, how to identify it and how to defeat or eliminate it, along with studying the potential repercussions of the given malware. For purposes of security, two types of malware analysis are performed, the preliminary analysis being called as Static Malware Analysis followed by a more in-depth analysis known as Dynamic Malware Analysis. [1] The former is known as preliminary because it involves studying the code of the malicious software whereas the latter is accomplished by actually executing the code in a sandbox type of environment. It is important to perform both types of analysis when encountered with a suspicious file to gain complete knowledge of its functioning and its effects on the system. However some of the malwares currently in market have become more sophisticated by virtue of anti-forensic techniques such as detection of the tools used by the forensic analyst and prevention of analysis via anti-debugging, anti-disassembly, anti-emulation, anti-memory dumping, encryption, incorporation of fake signatures and code obfuscation due to which they can evade detection by static analysis. [2]

Static Analysis

There are various techniques of conducting Static Analysis such as file signature verification which involves comparing the cryptographic hash of the suspected file with identified bad signatures stored in a database, however one limitation to this is that almost every day new malwares are being created to exploit zero-day vulnerabilities, therefore it is difficult to have all the malicious signatures in the database. Also the file format of the suspicious file has to be examined for file metadata and functions. One of the common techniques used by any malware author to evade detection is to obfuscate the code wherein the author modifies the code so as to hide its true mode of execution. One way to do this is to pack the code or compress it, which results in great reduction of strings that form the program, which consequently reduces the size of the malware, making malware transfers easy, that further limits our analysis. To analyze such programs statically we need to employ certain tools that can unpack the program prior to its analysis. [3]

However it may be thought that all the above analysis method can be implemented only when our system has actually been attacked and most readers will argue that this will be of little advantage as the damage has already been done. For this purpose it is advised that all the computer systems whether used for private or commercial purpose have antivirus installed in them. In conjunction with this, some even suggest to use anyone of the multitude of online tools available for testing any suspicious link or a file before downloading it to the system. Online tools such as but not limited to Malwarebytes, Norton, Symantec, VirusTotal etc. VirusTotal is the most accepted one as it combines the detection capabilities of wide range of antivirus engines.

Tools for online malware analysis

Earlier anti-virus software where sufficient in order to protect systems from malware attacks that could infect systems by way of pen drive or Bluetooth. But with the advent of internet, it became important to protect systems from attacks that used the networks to spread throughout. For this purpose, online tools for malware analysis were developed in order to aid the detection of antivirus on our system as well as to help in identifying the malicious nature of links or files prior to downloading.

VirusTotal

VirusTotal was launched with an aim to provide free service by the Spanish security company Hispasec Sistemas in the year 2004. It was acquired by Google in June 2012 and currently owned by Chronicle Security Ireland Limited, an Irish limited company. Its main job is to analyze files and URL’s for different types of Malware infections such as virus, worms and Trojans etc. It is a platform formed by a collaboration between members of antivirus industry, researchers and end-users. It is formed on a simple agreement that the users get the analysis report and the antivirus engines get to add new malware in their database that they can use to improve their detection service.

VirusTotal utilizes over 70 antivirus scanners and URL/domain blacklisting services, along with many tools that extract signals from the content under analysis. It is free to use and can also be used for non-commercial purposes. The malware characteristics that are covered under analysis include known-bad signatures, metadata extraction, heuristics, and identification of malicious signals. It not only tells whether the submitted file is malicious, but also displays the label of infecting malware. [4]

It can be used to analyze files, file hashes, URL’s, IP addresses, domains which are given as inputs through a series of antivirus scanners and provides a wealth of information besides just the scanning results i.e. whether or not the sample is malicious. Here, it is necessary to take note of the point that VirusTotal does not prevent malwares from attacking our system neither does it delete any malwares from the system, it is just a means of detecting a threat.

a. VirusTotal Privacy Policy

Users may choose to register and create an account with VirusTotal, in such a case, VirusTotal may have access to user name, email address, unique username and password.

It shall also have access by virtue of Google analytics to device specific information such as hardware model operating system version, unique device identifier, mobile network information, Browser version information and IP address. It automatically collects information as to how the service was used, device state characteristics such as event crash information, standard http request headers, date and time. It stores cookies on user information and makes use of browser web storage and application data caches.

It also accepts that certain samples uploaded for analysis may contain private or sensitive information, probably due to the infecting malware, and that it still processes it and the raw data of the sample is even shared with the partners and customers however for legitimate purposes. [5]

b. Mediums of using VirusTotal.

Apart from uploading a specific file from your system folder or copying a link to the online VirusTotal portal and analyzing the same, we can also use it via the following means.

i. VirusTotal Browser Extensions (VT4Browsers): Downloading them allows you to automatically analyze any link or a file or even an email just by a right click on it. [6]

ii. VirusTotal Desktop Apps: VT provides a range of client side tools to simplify the interaction. The Windows Uploader could be used through command line however since 2017 onwards VirusTotal have dropped the support for it. Instead VirusTotal Uploader (3rd party open source Uploader), a Microsoft windows desktop application can be used. One benefit of using this is that, it can be used to detect active malwares on system and using it is as easy as a right –click on the file to be analyzed or simply dragging the file into the Uploader.[7]

iii. VirusTotal Android Application: However as these are not developed by VirusTotal, the user can use them on their own risk.

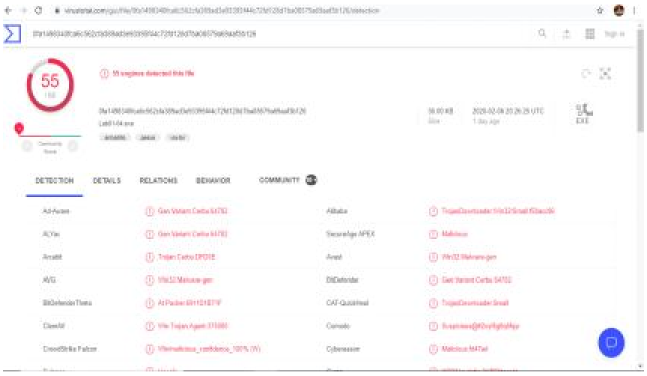
c. VT Internal Tools.

There is an internal VirusTotal tool called as Malware Retrohunting which allows to search text or binary patterns and hence assists in finding any malware. This is due to the fact that all the samples that are getting uploaded on the VirusTotal server are retained there.

There is one more tool called as WINJA tool which is designed basically to provide advanced functionality to users for establishing whether their system is a victim of malware attack or whether the files they download are malicious. The tool is a product of Phrozen Software. It utilizes the VirusTotal engine scanner to scan the samples that are uploaded to it, hence its detection is based on VirusTotal report. Further this tool also provides certain means to eliminate the detected malware. [8]

METHODOLOGY

For the purpose of my analysis, I downloaded an unknown malware say ABC.exe and uploaded the same on the online portal of VirusTotal and got the following results.



55 engines detected this file

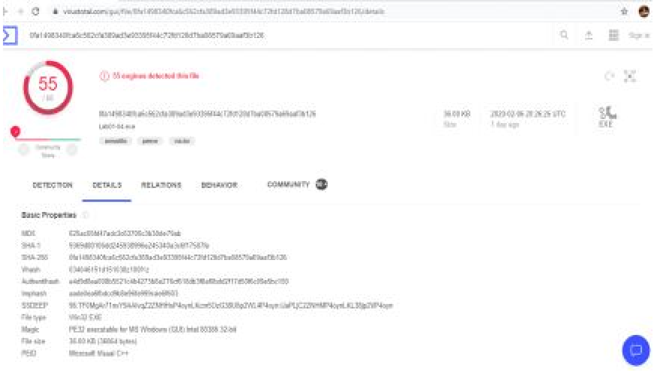
DETECTION DETAILS RELATIONS BEHAVIOR COMMUNITY

Ad-Aware	Gen:Variant.Cerbu.54782	Alibaba	TrojanDownloader:Win32/Small.[illegible]
ALYac	Gen:Variant.Cerbu.54782	SecureAge APEX	Malicious
Arcabit	Trojan.Cerbu.DFD1E	Avast	Win32:Malware-gen
AVG	Win32:Malware-gen	BitDefender	Gen:Variant.Cerbu.54782
BitDefenderTheta	AI:Packer.[illegible]	CAT-QuickHeal	TrojanDownloader.Small
ClamAV	Win.Trojan.Agent.[illegible]	Comodo	[illegible]
CrowdStrike Falcon	Win/malicious_confidence_100% (W)	Cybereason	Malicious.[illegible]

Fig.1: This shows detection of malware by different antivirus engines

There are following tabs that provide us various information about the file namely:

DETECTION: This tab shows how many and which of the antivirus engines have identified the file as malicious along with a short description regarding the malware. For this particular file, most antivirus scanners report it to be some kind of a Trojan. Here, it should be noted that the scanners which do not recognize it, does not mean that they are any lesser that the ones that recognize the malware.



55 engines detected this file

DETECTION DETAILS RELATIONS BEHAVIOR COMMUNITY

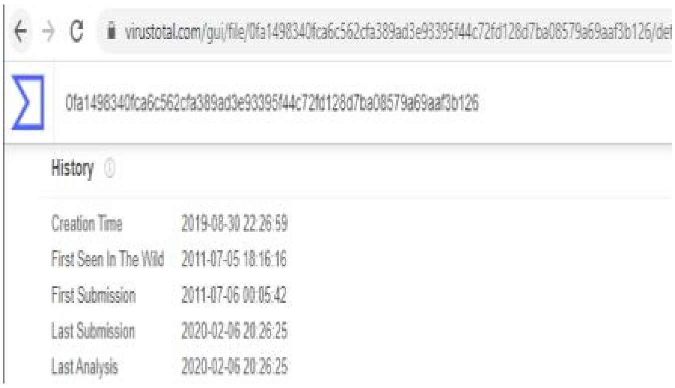
Basic Properties

MD5	[illegible]
SHA-1	[illegible]
SHA-256	[illegible]
Vhash	[illegible]
Authentihash	[illegible]
Imphash	[illegible]
SSDEEP	[illegible]
File type	Win32 EXE
Magic	PE32 executable for MS Windows (GUI) Intel 80386 32-bit
File size	36.00 KB (36864 bytes)
PEiD	Microsoft Visual C++

Fig2: Details tab which shows file hash values.

DETAILS: This tab provides additional information about the uploaded file such as hash value of the file using different algorithms such as MD5, SHA-1, SHA-256, Vhash, Authentihash, Imphash, SSDEEP, file type, file size etc. Here, it must be rightly mentioned that there is reason behind calculating so many different hash values for e.g.: The Imphash is calculated as per the libraries that are imported along with their linked functions by the uploaded sample in order to function effectively, so even if the malware author tries to change the signature of the file, this Imphash remains identical.

ISSN 2394 - 7780



virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/det

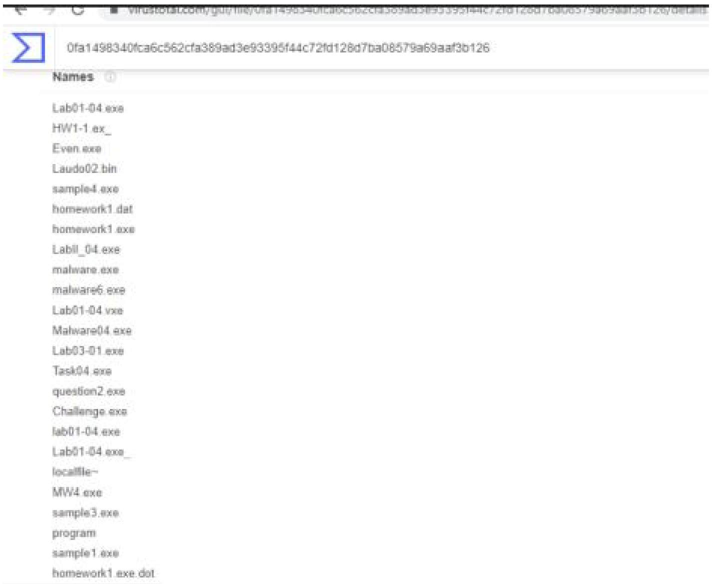
0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

History

Creation Time	2019-08-30 22:26:59
First Seen In The Wild	2011-07-05 18:16:16
First Submission	2011-07-06 00:05:42
Last Submission	2020-02-06 20:26:25
Last Analysis	2020-02-06 20:26:25

Fig 3: Details tab shows history of the file.

As per this data, the malicious file was first created on 30th August 2019, which is however not possible because the same file was first seen in wild on 5th July 2011 as per the analysis data. I have performed this analysis on 6th February 2020 as is also shown by the analysis data.



virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/details

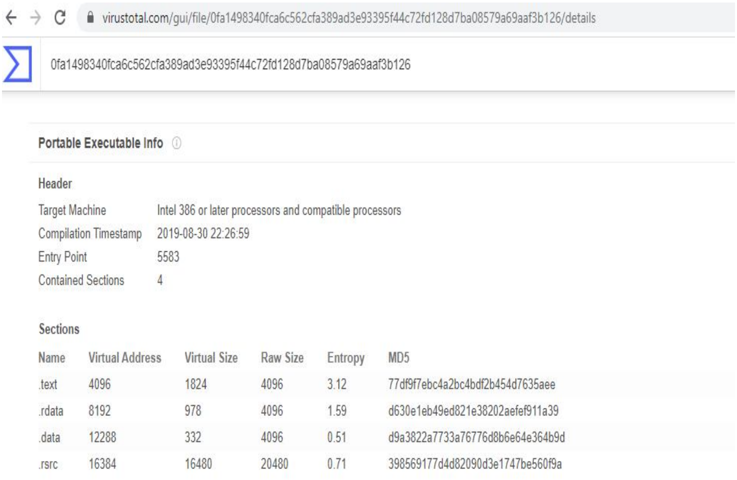
0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Names

Lab01-04.exe
HW1-1.ex_
Even.exe
Laudo02.bin
sample4.exe
homework1.dat
homework1.exe
Labll_04.exe
malware.exe
malware6.exe
Lab01-04.vxe
Malware04.exe
Lab03-01.exe
Task04.exe
question2.exe
Challenge.exe
lab01-04.exe
Lab01-04.exe_
localfile~
MW4.exe
sample3.exe
program
sample1.exe
homework1.exe.dot

Fig 4: Details tab also shows the names with which the file has been submitted to VirusTotal or found in wild.

This section shows the different names that have been used to identify the malware thus eliminating anonymity and establishing identity.



virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/details

0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Portable Executable Info

Header

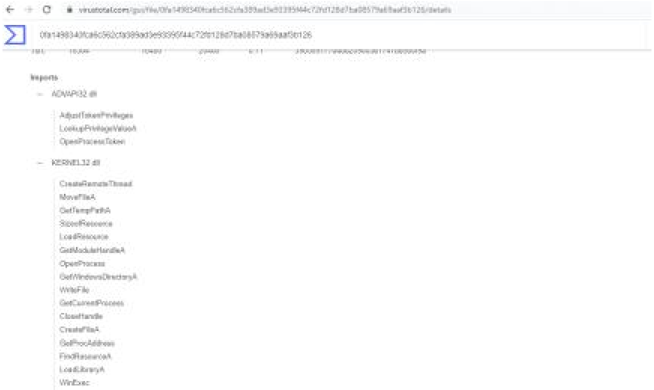
Target Machine	Intel 386 or later processors and compatible processors
Compilation Timestamp	2019-08-30 22:26:59
Entry Point	5583
Contained Sections	4

Sections

Name	Virtual Address	Virtual Size	Raw Size	Entropy	MD5
.text	4096	1824	4096	3.12	77df9f7ebc4a2bc4bdf2b454d7635aee
.rdata	8192	978	4096	1.59	d630e1eb49ed821e38202aefef911a39
.data	12288	332	4096	0.51	d9a3822a7733a76776d8b6e64e364b9d
.rsrc	16384	16480	20480	0.71	398569177d4d82090d3e1747be560f9a

Fig 4: Details tab also shows details associated with PE file.

Files with extensions such as .EXE and .DLL are commonly found to be of malicious nature. Files with the above extensions follow the PE (portable executable format). The PE file format is a data structure that contains the information necessary for the Windows OS loader to manage the wrapped executable code. Nearly every file with executable code that is loaded by Windows is in the PE file format. A PE file can be divided into two parts: the header and the section which are further subdivided into several parts. The header part contains details about metadata of the file such as location, size, time and date stamp. The size of the file in memory should be compared with the size of file on disk because if the virtual size is more for any of the sections than the size of raw data it indicates that the file may be malicious. The section part is divided into subsections namely: .text, .rdata, .data, .rsrc, .pdata, .reloc etc. These sections can be obfuscated to hinder analysis. [1]



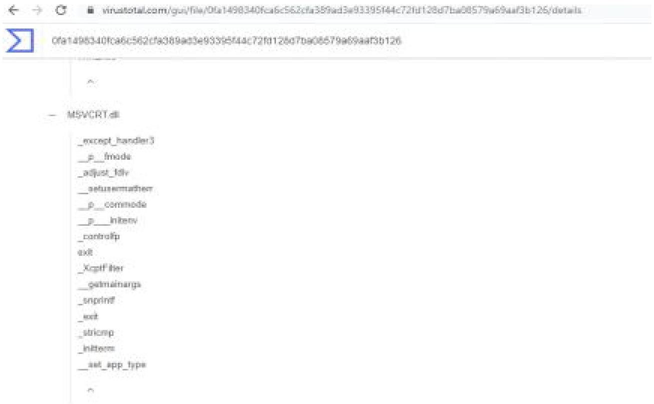
virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/details

0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Imports

- ADVAPI32.dll
  - AdjustTokenPrivileges
  - LookupPrivilegeValueA
  - OpenProcessToken
- KERNEL32.dll
  - CreateRemoteThread
  - MoveFileA
  - GetTempPathA
  - SizeofResource
  - LoadResource
  - GetModuleHandleA
  - OpenProcess
  - GetWindowsDirectoryA
  - WriteFile
  - GetCurrentProcess
  - CloseHandle
  - CreateFileA
  - GetProcAddress
  - FindResourceA
  - LoadLibraryA
  - WinExec

Fig 5: This shot shows the list of imported libraries and functions by the malware.

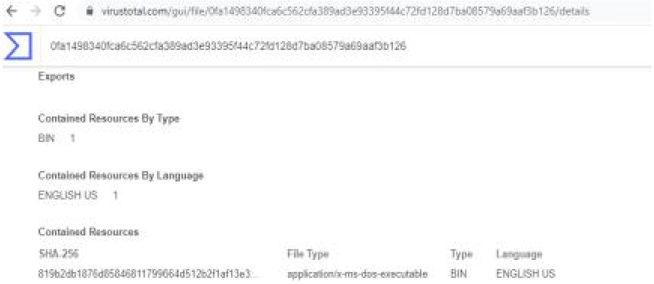


virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/details

0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

- MSVCRT.dll
  - _except_handler3
  - __p__fmode
  - _adjust_fdiv
  - __setusermatherr
  - __p__commode
  - __p___initenv
  - _controlfp
  - exit
  - _XcptFilter
  - __getmainargs
  - _snprintf
  - _exit
  - _stricmp
  - _initterm
  - __set_app_type

Fig 6: continued list of imported libraries and functions.



virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/details

0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Exports

Contained Resources By Type

BIN 1

Contained Resources By Language

ENGLISH US 1

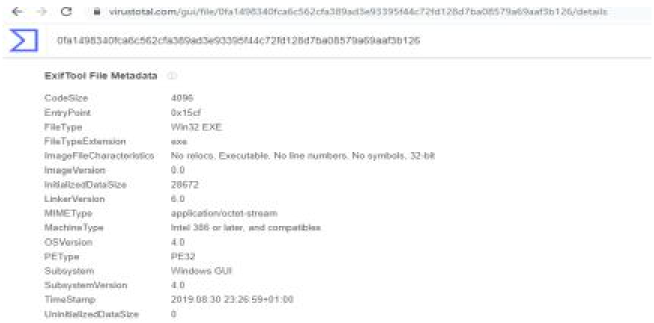
Contained Resources

SHA-256	File Type	Type	Language
819b2db1876d85846811799664d512b2f1af13e3…	application/x-ms-dos-executable	BIN	ENGLISH US

Fig 7: This shot shows the list of exported libraries and functions by the malware.

The knowledge of the imported and exported libraries and functions help us to understand about the actions of the executable.

The detailed list shows that the malware is not packed or obfuscated.



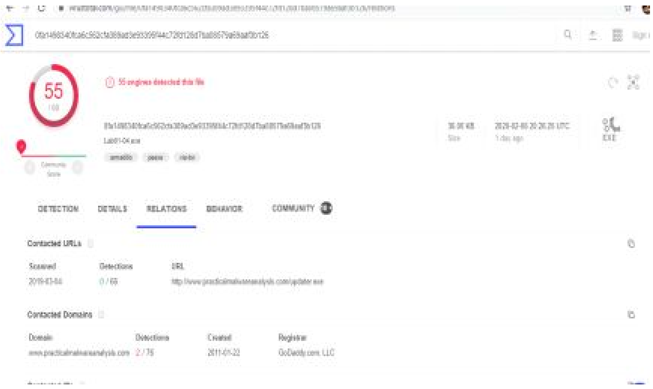
virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/details

0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

ExifTool File Metadata

CodeSize	4096
EntryPoint	0x15cf
FileType	Win32 EXE
FileTypeExtension	exe
ImageFileCharacteristics	No relocs, Executable, No line numbers, No symbols, 32-bit
ImageVersion	0.0
InitializedDataSize	28672
LinkerVersion	6.0
MIMEType	application/octet-stream
MachineType	Intel 386 or later, and compatibles
OSVersion	4.0
PEType	PE32
Subsystem	Windows GUI
SubsystemVersion	4.0
TimeStamp	2019:08:30 23:26:59+01:00
UninitializedDataSize	0

Fig 8: Metadata associated with malware obtained using Phil Harvey Exiftool.



55 engines detected this file

Lab01-04.exe

DETECTION DETAILS RELATIONS BEHAVIOR COMMUNITY

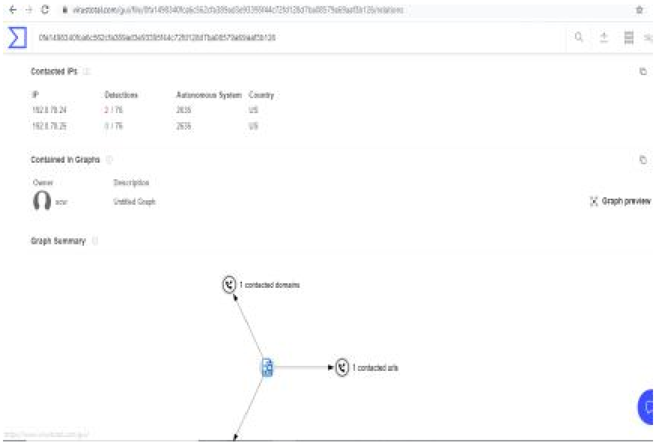
Contacted URLs

Scanned	Detections	URL
2019-03-04	0 / 66	http://www.practicalmalwareanalysis.com/updater.exe

Contacted Domains

Domain	Detections	Created	Registrar
www.practicalmalwareanalysis.com	2 / 76	2011-01-22	GoDaddy.com, LLC

Fig 9: RELATIONS tab: The above shot shows which IP addresses, url's and domains have been checked to identify and detect the malware.



Contacted IPs

IP	Detections	Autonomous System	Country
192.0.78.24	2 / 76	2635	US
192.0.78.25	0 / 76	2635	US

Contained In Graphs

Owner	Description
acw	Untitled Graph

Graph preview

Graph Summary

Fig 10: The above shot shows a graph.

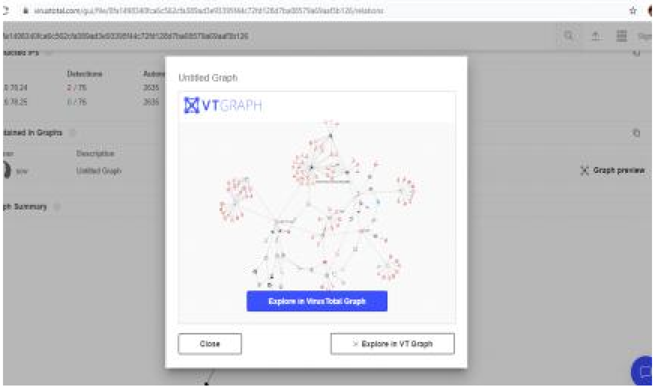
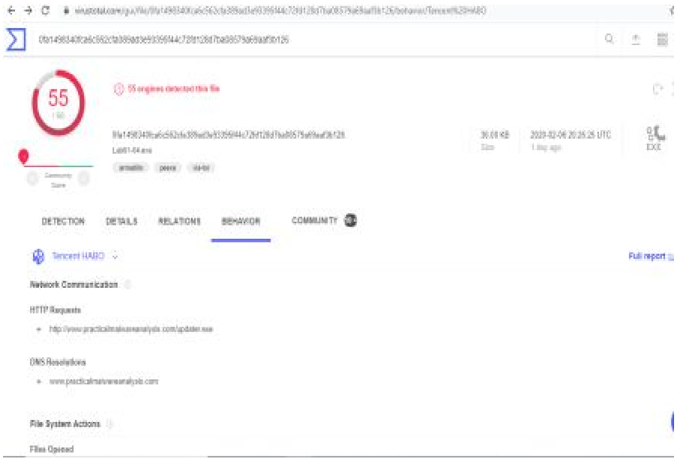


Fig 11: VT graph can be explored only after registering with VirusTotal.



55 engines detected this file

Lab01-04.exe

DETECTION DETAILS RELATIONS BEHAVIOR COMMUNITY

Tencent HABO

Full report

Network Communication

HTTP Requests

- http://www.practicalmalwareanalysis.com/updater.exe

DNS Resolutions

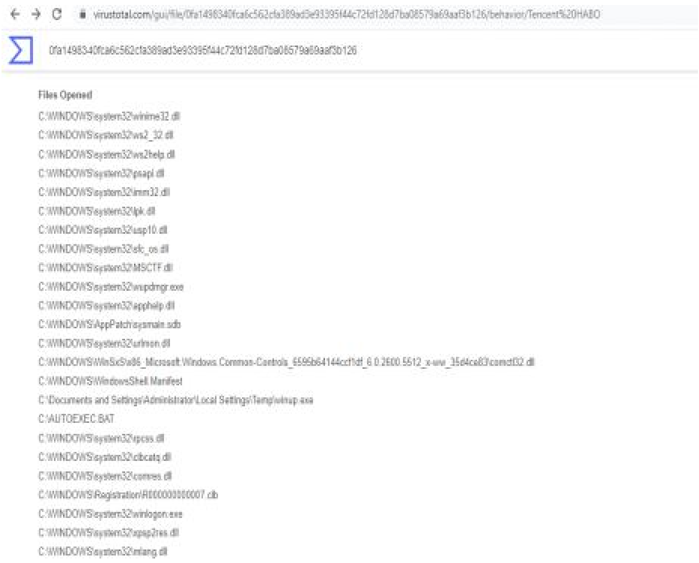
- www.practicalmalwareanalysis.com

File System Actions

Files Opened

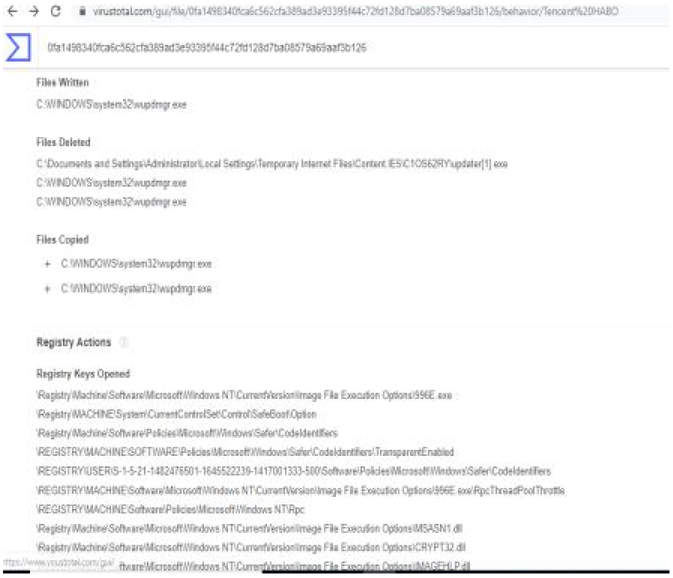
Fig 12: BEHAVIOR: Details shown under the behavior section.

The network communication section shows the network communications made by the file when being executed.



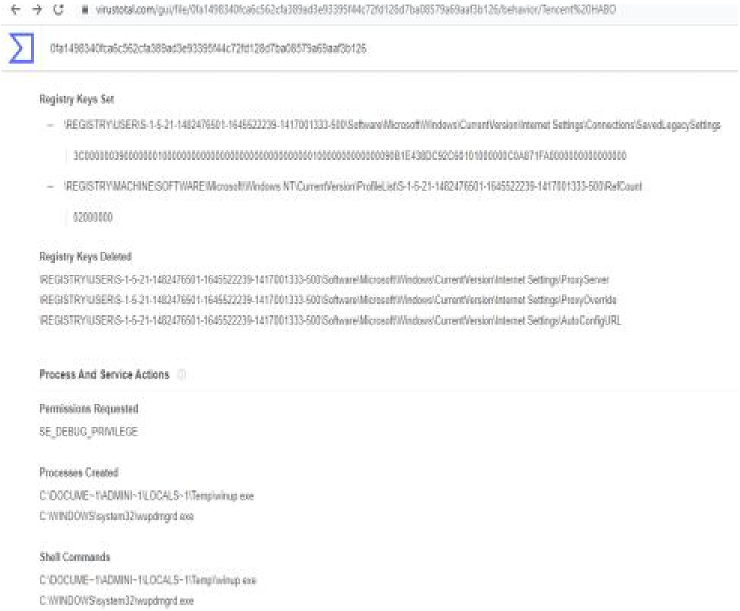
```
virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/behavior/Tencent%20HABO
0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Files Opened
C:\WINDOWS\system32\winime32.dll
C:\WINDOWS\system32\ws2_32.dll
C:\WINDOWS\system32\ws2help.dll
C:\WINDOWS\system32\psapi.dll
C:\WINDOWS\system32\imm32.dll
C:\WINDOWS\system32\lpk.dll
C:\WINDOWS\system32\usp10.dll
C:\WINDOWS\system32\sfc_os.dll
C:\WINDOWS\system32\MSCTF.dll
C:\WINDOWS\system32\wupdmgr.exe
C:\WINDOWS\system32\apphelp.dll
C:\WINDOWS\AppPatch\sysmain.sdb
C:\WINDOWS\system32\urlmon.dll
C:\WINDOWS\WinSxS\x86_Microsoft.Windows.Common-Controls_6595b64144ccf1df_6.0.2600.5512_x-ww_35d4ce83\comctl32.dll
C:\WINDOWS\WindowsShell.Manifest
C:\Documents and Settings\Administrator\Local Settings\Temp\winup.exe
C:\AUTOEXEC.BAT
C:\WINDOWS\system32\rpcss.dll
C:\WINDOWS\system32\clbcatq.dll
C:\WINDOWS\system32\comres.dll
C:\WINDOWS\Registration\R000000000007.clb
C:\WINDOWS\system32\winlogon.exe
C:\WINDOWS\system32\xpsp2res.dll
C:\WINDOWS\system32\mlang.dll
```

Fig 13: File System Actions: shows the files that are opened. These actions refer to those that are executed on the file system of the sandbox environment.



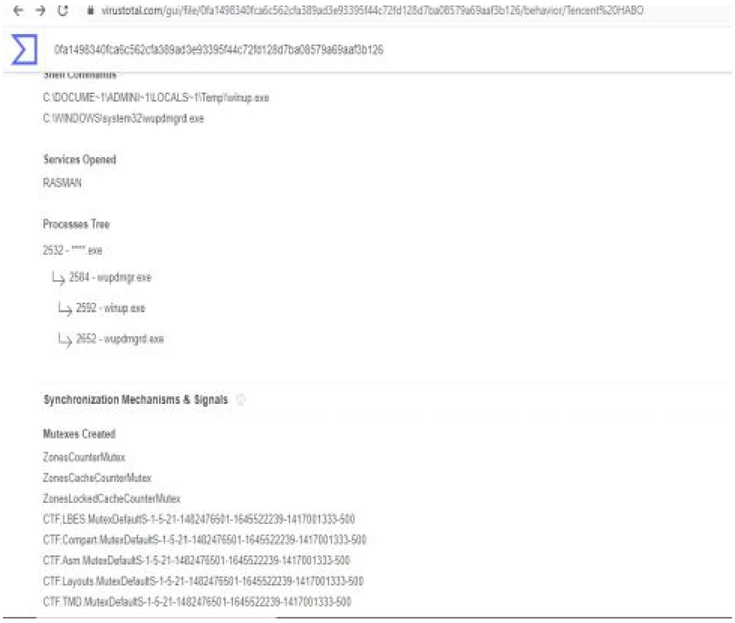
```
virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/behavior/Tencent%20HABO
0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Files Written
C:\WINDOWS\system32\wupdmgr.exe

Files Deleted
C:\Documents and Settings\Administrator\Local Settings\Temporary Internet Files\Content.IE5\C1OS62RY\updater[1].exe
C:\WINDOWS\system32\wupdmgr.exe
C:\WINDOWS\system32\wupdmgr.exe

Files Copied
  + C:\WINDOWS\system32\wupdmgr.exe
  + C:\WINDOWS\system32\wupdmgr.exe

Registry Actions

Registry Keys Opened
\Registry\Machine\Software\Microsoft\Windows NT\CurrentVersion\Image File Execution Options\996E.exe
\Registry\MACHINE\System\CurrentControlSet\Control\SafeBoot\Option
\Registry\Machine\Software\Policies\Microsoft\Windows\Safer\CodeIdentifiers
\REGISTRY\MACHINE\SOFTWARE\Policies\Microsoft\Windows\Safer\CodeIdentifiers\TransparentEnabled
\REGISTRY\USER\S-1-5-21-1482476501-1645522239-1417001333-500\Software\Policies\Microsoft\Windows\Safer\CodeIdentifiers
\REGISTRY\MACHINE\Software\Microsoft\Windows NT\CurrentVersion\Image File Execution Options\996E.exe\RpcThreadPoolThrottle
\REGISTRY\MACHINE\Software\Policies\Microsoft\Windows NT\Rpc
\Registry\Machine\Software\Microsoft\Windows NT\CurrentVersion\Image File Execution Options\MSASN1.dll
\Registry\Machine\Software\Microsoft\Windows NT\CurrentVersion\Image File Execution Options\CRYPT32.dll
https://www.virustotal.com/gui/ ftware\Microsoft\Windows NT\CurrentVersion\Image File Execution Options\IMAGEHLP.dll
```

Fig 14: Similar to the last shot, this one shows the other actions taken by the malware on file system such as deleting, writing and copying files.



```
virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/behavior/Tencent%20HABO
0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Registry Keys Set
  - \REGISTRY\USER\S-1-5-21-1482476501-1645522239-1417001333-500\Software\Microsoft\Windows\CurrentVersion\Internet Settings\Connections\SavedLegacySettings
      3C0000003900000001000000000000000000000000000000000100000000000000090B1E438DC52C6010100000C0A871FA0000000000000000
  - \REGISTRY\MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\ProfileList\S-1-5-21-1482476501-1645522239-1417001333-500\RefCount
      02000000

Registry Keys Deleted
\REGISTRY\USER\S-1-5-21-1482476501-1645522239-1417001333-500\Software\Microsoft\Windows\CurrentVersion\Internet Settings\ProxyServer
\REGISTRY\USER\S-1-5-21-1482476501-1645522239-1417001333-500\Software\Microsoft\Windows\CurrentVersion\Internet Settings\ProxyOverride
\REGISTRY\USER\S-1-5-21-1482476501-1645522239-1417001333-500\Software\Microsoft\Windows\CurrentVersion\Internet Settings\AutoConfigURL

Process And Service Actions

Permissions Requested
SE_DEBUG_PRIVILEGE

Processes Created
C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\winup.exe
C:\WINDOWS\system32\wupdmgrd.exe

Shell Commands
C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\winup.exe
C:\WINDOWS\system32\wupdmgrd.exe
```

Fig 15: The shot shows the actions taken by the file being analyzed on the registry of sandbox environment.



virustotal.com/gui/file/0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126/behavior/Tencent%20HABO

0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Shell Commands

C:\DOCUME~1\ADMINI~1\LOCALS~1\Temp\winup.exe

C:\WINDOWS\system32\wupdmgrd.exe

Services Opened

RASMAN

Processes Tree

2532 - ****.exe

↳ 2584 - wupdmgr.exe

↳ 2592 - winup.exe

↳ 2652 - wupdmgrd.exe

Synchronization Mechanisms & Signals

Mutexes Created

ZonesCounterMutex

ZonesCacheCounterMutex

ZonesLockedCacheCounterMutex

CTF.LBES.MutexDefaultS-1-5-21-1482476501-1645522239-1417001333-500

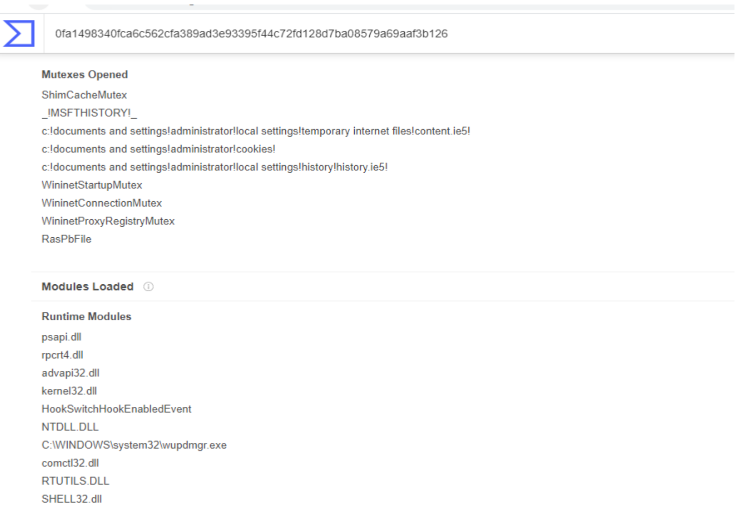
CTF.Compart.MutexDefaultS-1-5-21-1482476501-1645522239-1417001333-500

CTF.Asm.MutexDefaultS-1-5-21-1482476501-1645522239-1417001333-500

CTF.Layouts.MutexDefaultS-1-5-21-1482476501-1645522239-1417001333-500

CTF.TMD.MutexDefaultS-1-5-21-1482476501-1645522239-1417001333-500

Fig 16: Mutexes (mutual exclusion)



0fa1498340fca6c562cfa389ad3e93395f44c72fd128d7ba08579a69aaf3b126

Mutexes Opened

ShimCacheMutex

!MSFTHISTORY!

c:!documents and settings!administrator!local settings!temporary internet files!content.ie5!

c:!documents and settings!administrator!cookies!

c:!documents and settings!administrator!local settings!history!history.ie5!

WininetStartupMutex

WininetConnectionMutex

WininetProxyRegistryMutex

RasPbFile

Modules Loaded

Runtime Modules

psapi.dll

rpcrt4.dll

advapi32.dll

kernel32.dll

HookSwitchHookEnabledEvent

NTDLL.DLL

C:\WINDOWS\system32\wupdmgr.exe

comctl32.dll

RTUTILS.DLL

SHELL32.dll

Fig 17: modules loaded show the executables files required for an application to run.

CONCLUSION

After studying the analysis performed by a online malware detection engine known as VirusTotal it is matter to think about that when users upload samples on to the engine and if the malware is such that its aim is to extract personal information of the user then consequently even the sample uploaded may have personal information which is being further shared among a larger VT community. At this point, the user has no other option but to upload the file to prevent further harm risking the disclosure of his private information.

Also I encountered in my study that certain tools like WINJA automatically analyze all exe files and ongoing active processes on the system and although this is done with a complete legitimate objective of protecting the system, it may also risk the disclosure of sensitive data.

ACKNOWLEDGEMENT

We wish to acknowledge the valuable inputs provided by Mrs. Neeta Khobragade, HOD, dept of Digital and Cyber Forensic Science, Mumbai. We express immense gratitude to the insight and encouragement that she provided us.

REFERENCES

[1] Micheal Sikorski and Andrew Honig, Practical Malware Analysis.

[2] Craig Valli, The Malware Body of Knowledge 2008.

[3] Nirav Bhojani, Malware Analysis October 2014.

[4] https://support.virustotal.com/hc/en-us/articles/115002126889-How-it-works

[5] https://support.virustotal.com/hc/en-us/articles/115002168385-Privacy-Policy

[6] https://support.virustotal.com/hc/en-us/articles/115002700745-Browser-Extensions

[7] https://support.virustotal.com/hc/en-us/articles/115002179065-Desktop-Apps

[8] https://www.snapfiles.com/get/winja.html

ISSN 2394 - 7780

CLOUD COMPUTING ARCHITECTURE

Sachin Maurya
Student, Master of Science (Information Technology)
Patkar Varde College of Commerce, (Mumbai University), Goregaon West, Mumbai, Maharashtra

ABSTRACT

Cloud Computing now a days evolving rapidly and covering most of the IT region. It is the latest worldview of computer utilization. Cloud consumer can utilize resources, applications and infrastructure provided by cloud provide on a pay-as-you-use basic. Such service could be in the form of applications. Also massive storage infrastructure is available for storage of data provided by the consumers. This paper studies and analyses the current development in the cloud computing architecture.

Keywords: Cloud computing, architecture, Saas, Paas, Iaas, DaaS, Oaas, Sec-aaS, XaaS.

INTRODUCTION

We have consistently been putting away the information that we need onto our PC's storage devices and getting to it at whatever point required. however, presently innovation has dominated and the need to store everything on your physical hard disk is no longer there. Here, Cloud Computing comes into the scenario. Cloud computing is the strategy for computing wherein the information and projects are put away over the cyberspace and not on your physical storage device. The Internet is referred to as the Cloud in ‘Cloud Computing’. Like yours, others' and various associations' information are likewise kept over the Internet. With regard to an association's need, the prerequisites are significantly more. Servers, applications, Storage, etc. Making a physical foundation to store and introduce all these eventual tedious and costly. Rather, on the off chance that you get a share of what is already installed over the Internet for use, it is savvy and addresses your issues too. So Cloud computing is mostly founded on sharing computing resources.

CHARACTERISTICS OF CLOUD COMPUTING

1) On-demand self-service: User can easily audit cloud computing services without human synergy for the most part done intensive an online web-based self service portal.

2) Measured Services: Resource use are checked, estimated and reported straightforwardly depending on usage.

3) Broad Network Access: Cloud computing resources are available on the network, web platforms, for example, mobiles and workstations.

4) Rapid Elasticity: Assets are observed and released on demand and mechanized depends on the guideline, This guarantees your application will have absolutely the amount of resources whenever it will be require.

5) Resources Pooling: Suppliers serve various customers, clients or "occupants" with temporary and versatile administrations. These administrations can be managed to suit every customer's needs with no charges being supposed to the customer or end-user.

CLOUD COMPUTING ARCHITECTURE

DESCRIBING THE NIST MODEL

According to NIST, five critical on-screen characters have been recognized in cloud computing. The NIST shows the actors which are reviewed beneath.

Cloud Consumer: An individual or association that runs and establishes a business relationship uses cloud administrators from cloud providers.

Cloud Provider: An individual, Agency or entity occupied with providing cloud computing administration to intrigued person or company.

Cloud Auditor: An association is accountable for directing the autonomous assessment of distributed computing and deciding the viability and security of the framework.

Cloud Broker: An outsider organization or person which fills in as a middle person between consumer and suppliers. He/she is authorized for arranging policy and states of agreement for the acquisition of cloud administrations.

Cloud Carrier: A middle person, organization or individual that gives network and transport of cloud services from cloud provider to cloud clients.

CLOUD COMPUTING SERVICES

Cloud architecture can be isolated into two sections, back, and front end. Front end can be viewed to the customer through the Internet, allowing customer cooperation with the system. The back end incorporates the various cloud administration models.

Software-as-a-Service(SaaS): To utilize the supplier's applications running on a cloud framework the storage is provided to the consumer. The applications are accessible from various customer devices through either an interface, for instance, an internet browser (eg – email) or a program interface.

EXAMPLE OF SAAS

Email Applications such as gmail, Office application such as (spreadsheet, word editor, presentation application).

SaaS Applications – BigCommerce, DocuSign, Google Apps, MailChimp, Saleforce,Hubspot, Dropbox, Hubspot.

Platform-as-a-Service (PaaS): The capacity made available for the client is, to send customer-developed softwares using programming languages, libraries, administrations and tools onto the cloud foundation upheld by supplier. The customer doesn't oversee or control the hidden cloud framework including system, servers, operating system or capacity yet has authority over deployed applications and conceivable configured setting for applications hosting environment.

EXAMPLE OF PAAS

Business Intelligence, Database, Development and Testing, Integration, Application Deployment.

PaaS Applications - AWS Elastic Beanstalk, Force.com, Heroku, Magento Commerce Cloud, Windows Azure, OpenShift, Apache Stratos.

Infrastructure-as-a-Service (IaaS): The ability gave to the user is arrangement managing, storage, network and other basic computing assets where the buyer can send and subjective application which can incorporate OS and apps. The buyer doesn't control the hidden cloud framework yet has control over OS, storage and deployed apps.

EXAMPLE OF IAAS

Content-Delivery-Networks(CDNs), Backup and Recovery, Compute, Storage..

IaaS Applications: Amazon Web Service EC2, Digital Ocean, Google Compute Engine (GCE), Rackspace.

Database-as-a-Service(DaaS): DaaS is another administration that is turning popular now a days in cloud computing world. The thought behind DaaS is to maintain the simplicity and avoid the cost of running your own database. DaaS allows clients to pay for what they are truly using rather than the site. license for the entire database. Despite standard storage interfaces, for instance, RDBMS and archive systems, some DaaS commitments give table-style deliberations that are planned to scale out to store and restore an immense measure of information inside a stuffed time span, frequently excessively enormous, too much expensively or absurdly deferred for most business RDBMS to adjust. Instances of this kinds of DaaS incorporates Amazon S3, Google BigTable, and Apache HBase, and so on.

EXAMPLE OF DAAS

Ease of use, Power, Integration, Management.

DaaS Applications: Urban Mapping, Xignite, D&B Hoovers.

CLOUD COMPUTING DEPLOYMENT MODEL

A cloud deployment model is portrayed by where the structure for the deployment dwells and who has authority over that framework. Each model of cloud deployment satisfy different organizational needs. There are 3 different cloud deployment models.

Private Cloud: The deployment model backings all client who want to take advantage of computing asset, for instance, hardware or software on subscription.

Private Cloud: It is normally an infrastructure for a solitary association. This type of foundation can be inspected by an organization itself to help other user requirements. Private clouds are more pricy than public clouds due to the capital utilization included procuring and maintaining it.

Hybrid Cloud: An association take benefits of interconnected public and private cloud frameworks in this model. Many association uses this model if they need to expand up their IT foundation rapidly.

Community Cloud: This deployment model assists different associations sharing resources that are a part of a community. Examples of community cloud are -

Various police department in a country,

Various colleges under a single university, etc.

SECURITY OF CLOUD

There are various cloud Security features available such as data encryption, VM isolation, secure VM migration, platform, segregation between the users and protection, and many more.

SECURITY ENSURES

- Entry validation and approval
- Establish continuous accessibility
- Assuring customer confidentiality
- Subscriber identity management an sharp cloud foundation that is expected for immediate access of security features facilated virtual server environment.

CLOUD SECURITY ISSUES AND EXISTING SOLUTION

This segment talks about the particular security issues and existing answers for secure cloud environment.

Cloud Security Alliance (CSA) has analyzed and defined seven security threats to cloud computing

Insiders Abuse and Nefarious Use of Cloud Computing: Abuse and nefarious utilization of cloud computing is one of the significant risk distinguished by the CSA. Example - usage of botnets to spread spam and malware. Attackers can access to an open or public cloud, for example, and find an approach to transfer malware to many PCs and use the intensity of the cloud infrastructure to attack different machines.

SUGGESTED SOLUTIONS BY CSA

- Stricter starting registration and authentication methods.
- Upgraded credit score card extortion tracking and coordination

Insecure Application Programming Interfaces: As software interfaces or APIs are what clients use to have connection with cloud services, those must have exceptionally secure verification, access control, encryption and action observing instruments - explicitly when outsiders (third parties) start to develop on them.

SUGGESTED SOLUTIONS BY CSA

- Examine the safety version of cloud provider interfaces.
- Assure the best verification and access controls are completed working together with encrypted transmission.
- Perceive the requirement chain related with the API.

Malicious Insiders: The malicious insider danger is one that is significant as many suppliers don't uncover how they employ individuals, how they give them access to resources or how they provision them.

RECOMMENDED SOLUTIONS BY CSA

- Indicate human asset necessities as a piece of lawful agreements.
- Require straightforwardness into whole data security and the management practices, just as consistence announcing.
- Distinguish security break notice systems.

Shared Technology Vulnerabilities: IaaS providers usually share infrastructure. Sadly, the component on which this infrastructure is principally based was not intended for that. To ensure that buyers don't thread on each different's "area", checking and potential break-up is required.

RECOMMENDED SOLUTIONS BY CSA

- Observe surroundings for unauthorized adjustments/activity.
- Implement service level contracts for fixing and vulnerability treatment.
- Conduct weakness filtering and design reviews.

Data Loss / Leakage: Without a backup or by unauthorized access, data is always in danger of being lost or stolen. This is one of the top threats for organizations as the may lose their reputation.

SUGGESTED SOLUTIONS BY CSA

- Execute powerful API access control.
- Encrypt and secure the integrity of information in a transaction.
- Establish tough key production, organization, capacity, and demolition exercise.
- Legally indicate supplier reinforcement and maintenance methods.

Account, Service & Traffic Hijacking: Account, service and traffic hijacking is another difficulty that cloud clients ought to know about. These dangers assortment from man-in-the-middle assaults, to phishing and spam crusades, to denial-of-service assaults.

SUGGESTED SOLUTIONS BY CSA

- Keep away from sharing the account information among customers and organizations.
- Leverage robust two-way authentication procedures where achievable.
- Comprehend cloud supplier security policy and service level agreements (SLA).

Data Unknown Risk Profile: Security needs to constantly inside the top piece of the worry list. Code updates, vulnerability profiles, security practices, intrusion tries - everything that must persistently be remembered.

SUGGESTED SOLUTIONS BY CSA

- Exposure of useful logs and data.
- Halfway/Complete exposure of infrastructure details.
- Observing and warning on vital records.

RELATED WORK

In, private cloud-computing and assignment of control is expected. The fact was to take a view at problems in private clouds. The territories of pricing, guidelines and information assurance as far as private cloud was talked about. This guarantees an ideal advantage of such an interest over the long haul. Security in distributed computing: openings and difficulties were exhibited. The methodology inspected a few security issues. In, an audit on distributed computing: plan difficulties in design and security are proposed. Different ideas identifying with cloud computing structures were inspected in certain details. Security matter were likewise talked about with certain solutions proposed. In, distributed computing - ideas, architecture, and difficulties were introduced. The qualities, advantages and issues in distributed computing were examined.

CONCLUSION

After the review completed, it has been conceivable to find that the cloud computing offers endeavors, SMEs and people are able to utilize the versatile, on-demand resources from anyplace and whenever. Clients additionally don't need to stress over infrastructure cost engaged with in-house application development. The cloud design offers 3 kinds of services and 4 deployment methods. Ongoing study shows most noteworthy spending plan assigned to SaaS by ventures. what's more, the utilization of public cloud is required to keep developing more than private cloud.

ACKNOWLEDGMENT

I endorse that, this Review paper have been completed by me with the support of staffs of patkar varde college.

REFERENCES

[9] Isaac Odun-Ayo, Member, IEEE, Sanjay Misra, Member, IEEE, and Frank Agono, “Cloud Computing Architecture: A Critical Analysis” ,

https://www.researchgate.net/publication/327125094_Cloud_Computing_Architecture_A_Critical_Analysis

[10] Manju Sharma, Sadia Husain, Halah Zain, “Cloud Computing Architecture& Services” , College of Computer Science and Information Systems, Jazan University, Jazan- K.S.A, IOSR Journal of Computer Engineering (IOSR-JCE).

[11] "Cloud Computing – Architecture, Platform and Security Issues: A Survey", Department of Computer Science and Engineering, Faculty of Science and Engineering, International Islamic University Chittagong, Chittagong - 4318, Bangladesh.

[12] “Cloud Computing Architecture”, http://docplayer.net/49214358-Cloud-computing-architecture.html.

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

A COMPARATIVE STUDY OF CURRENT TRENDING ONLINE FOOD DELIVERY APPLICATION (i.e. ZOMATO VS. SWIGGY)

Aamina Qureshi and Prof. Rajesh Maurya
Usha Pravin Gandhi College Arts, Science and Commerce

ABSTRACT

Swiggy and Zomato have now become synonymous with food-delivery and yet it seems as if both startups have very different ideas on how they plan to conquer the world.

Food delivery through web applications and phone applications have become a duopoly. Zomato and Swiggy collectively satisfy over three-quarters of all orders. From the two, Zomato is more in demand. Almost 12% of the indian population has Zomato installed on their phones whereas Swiggy is installed by only 10% of the population according to data shared with Quartz *by the market research firm Unomer.*

Both Zomato and Swiggy have invested big amount of money to reach as many customers as possible. The premise is simple—Scale. Aggregators have to make huge investments in technology and manpower to fulfill orders that usually cost less. The more the customers, the better you chances of turning a profit. And both companies have spent crazy money trying to pull people with tempting discounts for far too long now.

Keywords: Complimentary services, Duopoly.

INTRODUCTION

A decade ago i.e. in 2008 an Indian food delivery start up was established by Deepinder Goyal as Foodiebay . It was retitled as Zomato in 2010. In 2011, Zomato extended across various cities of the country like Delhi NCR, Bangalore , Chennai, Pune and Kolkata. Afterwards, in 2012, the company expanded internationally in several countries like the UAE, Qatar, the United Kingdom, Sri Lanka and . Currently it has 5000+ employees working day and night to deliver food at doorstep. Currently there are 8 crore i.e. 80 million active Zomato users.

Swiggy is one of the India's largest and valued online food ordering and delivery platform. Started in 2014 by Nandan Reddy Sriharsha Majety, Rahul Jaimini. Swiggy was established out of Bengaluru, India and, has its headquarters in the same city. In 2019, it was operating in100 Indian cities.

In 2019, Swiggy started immediate pick up and drop service Swiggy Go. The facility was used to pick up and drop off variety of things, comprising laundry and important document or parcel deliveries to customers.

LITERATURE REVIEW

According to Serhat Murat Alagoz & Haluk Hekimoglu (2012), e-commerce is rapidly growing worldwide, the food industry is also showing a steady growth. In this research paper they have used the Technology Acceptance Model (TAM) as a ground to study the acceptance of online food ordering system. Their data analysis revealed that the attitude towards online food ordering vary according to the ease and usefulness of online food ordering process and also vary according to their innovativeness against information technology, their trust in eretailers and various external influences.

According to H.S. Sethu & Bhavya Saini (2016), their aim was to investigate the student‟s perception, behavior and satisfaction of online food ordering and delivery services. Their study reveals that online food purchasing services help the students in managing their time better. It is also found that ease of availability of their desired food at any time and at the same time easy access to internet are the prime reasons for using the services.

According to Sheryl E. Kimes (2011), his study found that perceived control and perceived convenience associated with the online food ordering services were important for both users and non-users. Non-users need more personal interaction and also had higher technology anxiety to use the services. According to Leong Wai Hong (2016), the technological advancement in many industries have changed the business model to grow. Efficient systems can help improve the productivity and profitability of a restaurant. The use of online food delivery system is believed that it can lead the restaurant‟s business grow from time to time and will help the restaurants to facilitate major business online.

According to Varsha Chavan, et al, (2015), the use of smart device based interface for customers to view, order and navigate has helped the restaurants in managing orders from customers immediately. The capabilities of wireless communication and smart phone technology in fulfilling and improving business management and service delivery. Their analysis states that this system is convenient, effective and easy to use, which is expected to improve the overall restaurant business in coming times.

RESEARCH METHODOLOGY

In this section, the researcher discussed the methods to prove whether our NULL hypothesis or Alternate Hypothesis is accepted or rejected. For this the researcher collected quantitative data through surveys i.e. Google form.

H0: There is no significant difference between consumers of Zomato vs. Swiggy.

H1: There is a significant difference between Zomato vs Swiggy i.e. Zomato is better than Swiggy.

To prove the Hypothesis the researcher kept a sampling frame of 80 people and the exact sampling the researcher collected was from 70 people. For this, the researcher used non-probability sampling method given below:-

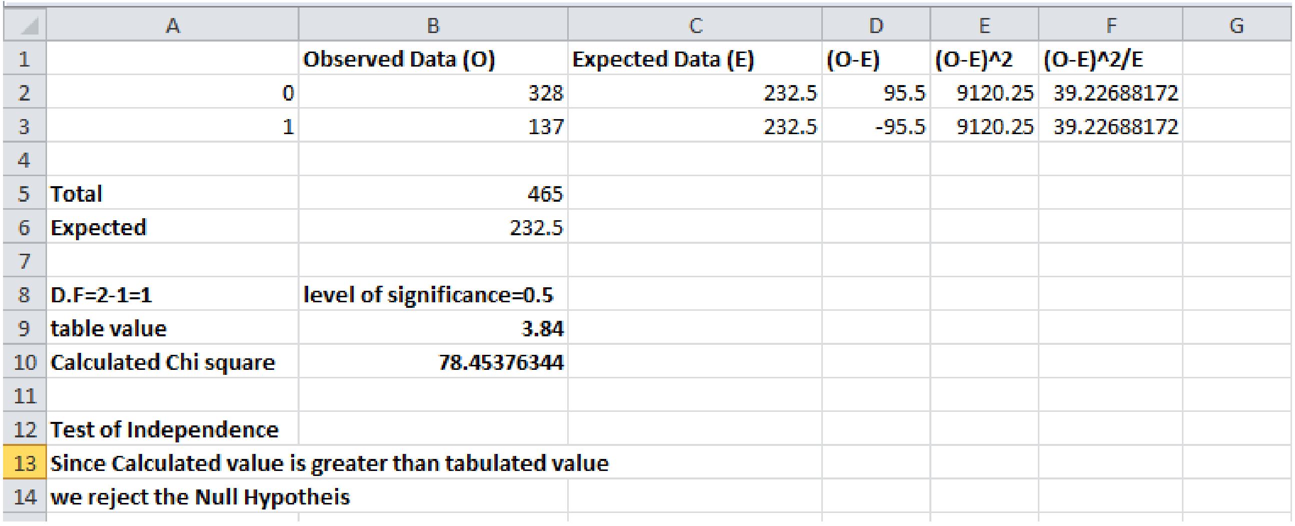
4. Convenient sampling
5. Judgmental sampling
6. Snowball sampling.

LIST OF QUESTIONNAIRE USED IN MY SURVEY

18. Out of the two which app do you prefer?
19. Which features you prefer the most in your favourite food delivery app.
20. Navigation through which app is easier than the other?
21. According to you which app is more costly?
22. Which out of the two offer flexible payment method?
23. Which app delivers food in a timely manner?
24. Which Food delivery app provides Food worth the price?
25. Do you know about the following ? (Zomato gold, Swiggy super, etc..)
26. Rate Zomato?
27. Rate Swiggy?

RESULT

On the basis of data collected above and to prove our hypothesis the researcher used Chi-Square test for hypothesis testing. The reason behind using Chi-Square test was that for it the researcher required more than 50 samples and also it's a non parametric test. The researcher used goodness of fit method to prove the hypothesis.



	A	B	C	D	E	F	G
1		Observed Data (O)	Expected Data (E)	(O-E)	(O-E)^2	(O-E)^2/E	
2	0	328	232.5	95.5	9120.25	39.22688172	
3	1	137	232.5	-95.5	9120.25	39.22688172	
4							
5	Total	465					
6	Expected	232.5					
7							
8	D.F=2-1=1	level of significance=0.5					
9	table value	3.84					
10	Calculated Chi square	78.45376344					
11							
12	Test of Independence						
13	Since Calculated value is greater than tabulated value						
14	we reject the Null Hypotheis						

From above data, the calculated Chi-Square value is 78.45376344.

Tabulated value of χ^2 test for 1 Degree of Freedom at 5% level of significance is 3.84: So the calculated value of χ^2 test is 78.45376344 is highly significant and Null Hypothesis is rejected at 5% level of significant.

ISSN 2394 - 7780

Hence the researcher concludes that H1(Alternate Hypothesis is accepted and its means there is significance difference Zomato and Swiggy i.e. Zomato is better than Swiggy).

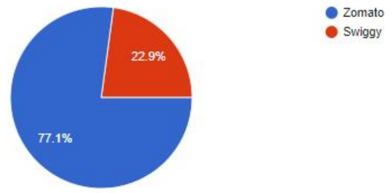
DATA COLLECTION

This data was collected through Google form and responses were collected from 70 people. There were total 10 questionnaire.

Detailed Description of questionnaire and responses collected are given below:

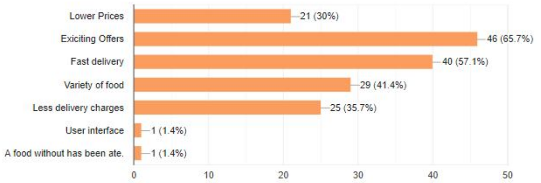
Out of the two which app do you prefer?

70 responses



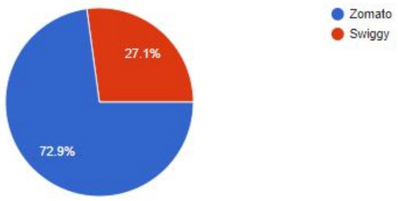
Which features you prefer the most in your favorite food delivery app?

70 responses



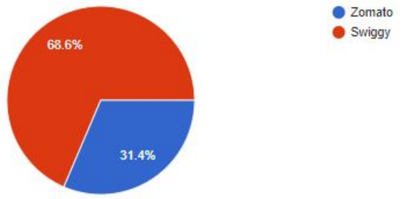
Navigation through which app is easier than the other?

70 responses



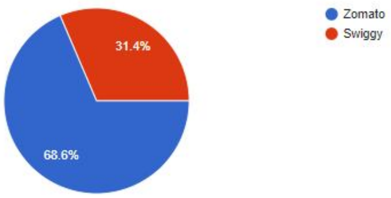
According to you which app is more costlier?

70 responses



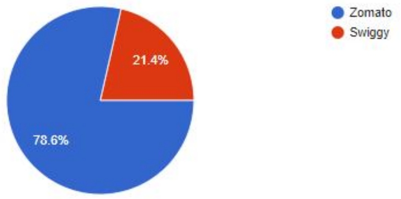
Which out of the two offer flexible payment method?

70 responses



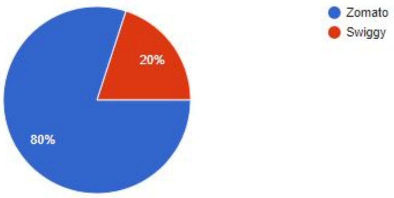
Which app delivers food in a timely manner?

70 responses



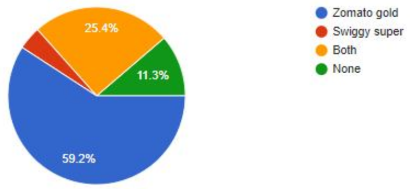
Which Food delivery app provides Food worth the price?

70 responses



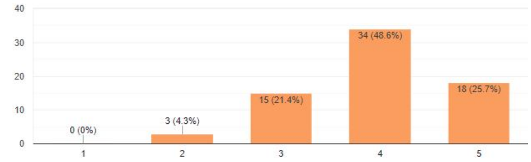
Do you know about the following ?

70 responses



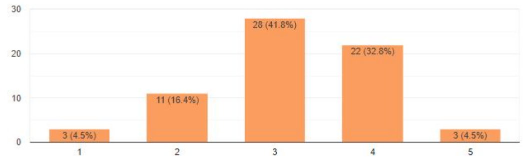
Rate Zomato

70 responses



Rate Swiggy

67 responses



CONCLUSION

This paper presented the significance difference Zomato and Swiggy that both are giving the best features to tackle their competitors and attract more customers but still difference in offers, delivery time and ease use makes Zomato stronger than Swiggy. But Swiggy‘s complementary services is adding fuel in increasing its market share and value resulting in increasing number of customers giving tougher competition to Zomato. As competition is growing Zomato need to be updated and provide more complementary service to keep his position strong in market.

REFERENCE

[1] International Journal of Multidisciplinary Research and Development Online ISSN: 2349-4182, Print ISSN: 2349-5979 Impact Factor: RJIF 5.72 www.allsubjectjournal.com

[2] https://en.wikipedia.org/wiki/Swiggy

[3] https://en.wikipedia.org/wiki/Zomato

[4] economics times

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

FAKE NEWS DETECTION, TECHNOLOGY: BLOCKCHAIN

Priya Shah and Neelam Naik
Usha Pravin Gandhi College of Arts, Science, Commerce

ABSTRACT

Blockchain technology has unlocked the doorway of generating decentralized applications, where security may be a big concern. Here, any transaction ever held is recorded permanently. Over the years, some non-reputable sources have been publishing fake news. Due to the lack of any regulatory system, this news cannot be verified. Hence, these unreliable sources can publish whatever they want and even in some cases, it makes chaos in society. In recent times due to the ease in internet availability and social media, inappropriate news can spread more quickly than ever before. In some cases, fake news is more attractive than the real one. Thus, most of the people become misguided. Using the benefits of Blockchain peer-to-peer network concepts, we'll discuss a way to detect fake news in social media."

Keywords: Block chain, Breadth First search, Ethereum, Decentralized.

INTRODUCTION

In this generation, social media is source of all types of worldwide news. But when an individual use this to spread phony news, it fails. Since clickbait articles take a short time to circulate exponentially in social media. In days the news is becoming viral worldwide. Such speculators benefit from the tendency of people to share the appeal news without knowing its authenticity or consequences. [1]

What is fake news?

Fake news is defined as information which has no fact behind it but presented factually accurate and consumed by millions through radio, websites or social media.

What is Blockchain?

“"A blockchain is a digital ledger that is decentralized, distributed, and often public, and is used to record transactions across many computers so that any record involved can not be changed retroactively without changing all subsequent blocks.

Researcher can leverage the benefits of blockchain protection, immutability and transparency in social media to make trust among shared news. Researcher will provide hypothesis, which might be implemented during a decentralized social media and provide authenticity of some news among users. [1]

Figure 1 shows number of active users through which medium, users usually gets informed about fake news. January 2019 records.”

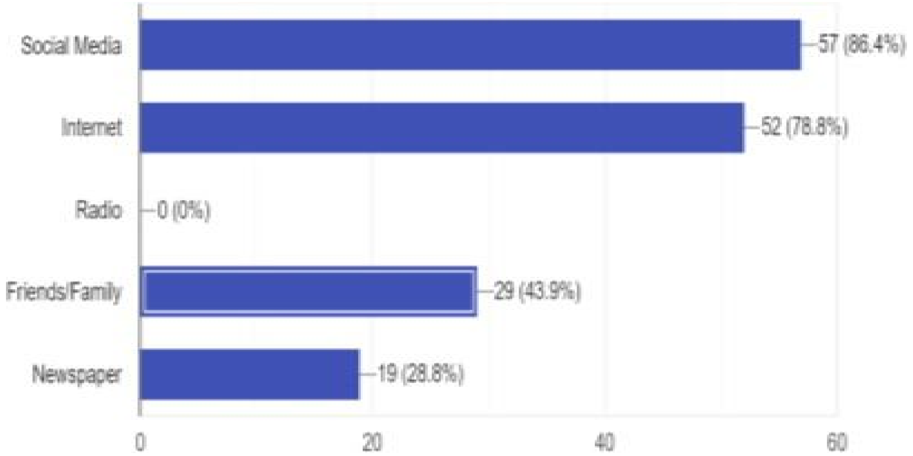


Figure 1: No. of active users

So, as Figure 1 says that Social media is at the epicentre. The fake news published in socialmedia seems similar to be actual. The purpose of posting incorrect newscast stories is to mark public's desirability. Thus, they'll earn extra currency with extra page opinions. Nearly all country uses socialmedia as a basis of news. This paper researcher will deliberate a way to detect these false news stories. Researcher will use the concept of decentralization, Ethereum, smartcontracts and use Breadth first search algorithm for calculating users proximity.

LITERATURE REVIEW

1. In a research learning, Gilda discovered the application of NLP techniques to identify fake news.
2. Research by Prannay S Reddy as Naïve Bayes Classifier to identify fake news
3. YoungkyungSeo, Chang-Sung Jeong presented a fake news detection model using media reliability.

RESEARCH METHOD

"Research study Blockchain technology and Ethereum chain (it is a platform for sharing information across the globe that cannot be manipulated or changed) are used as before stated. These two concepts have been discussed broadly in the next section."

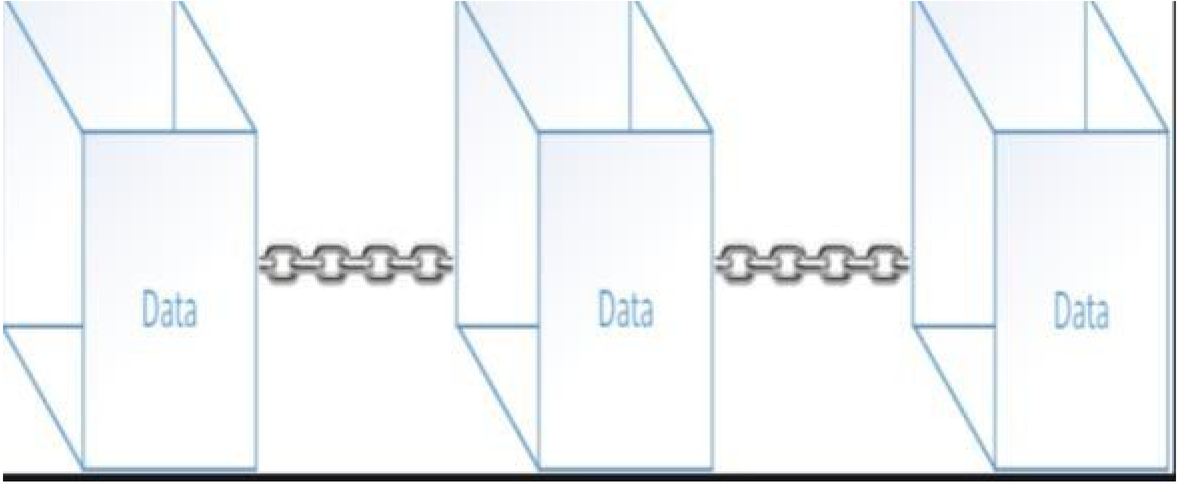


Figure 2: -Blockchain

A Blockchain

In the simplest terms, Blockchain can be described as a data structure that holds transactional records and while ensuring security, transparency, and decentralization. You can also think of it as a chain or records stored in the forms of blocks which are controlled by no single authority.

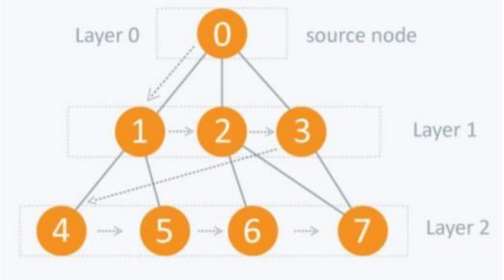


Figure3: Breadth First Search

B Breadth First Search

BFS is a basic algorithm for graph searching, which stands for Breadth First Search. It is used to explore organized tree graphs in a way to fully explore a certain level before hitting the next (deeper) level. BFS start traversing from a source node, then traverse the neighboring nodes repeatedly until it traverses the entire graph. We use this BFS definition to find the closeness and assign weight to our users.

C Ethereum

The Ethereum is a smart-contract based, decentralized network. Smart contracts in Blockchain are very successful in facilitating, checking a contract agreement. The users involved in the contract must agree to the terms and conditions for starting a contract. Once the specified conditions are met, the program will administer the pledge automatically. The transaction is done via the Ethereum chain undoubtedly trustworthy.

BACKGROUND

"In this present context, it is hard to determine if a news is deceptive (fake). Whereas in Blockchain we can offer a system where this validation can be done secretly. The idea is that we will integrate social media in a Blockchain in such a way that, random users will act as news validators. Because of anonymity (invisibility), they can validate news without any external pressure. Therefore, they cannot be biased by any other person (no third party involvement). After publishing news, the news will deploy as a transaction in a chain. After a certain level of virality the validator users will get a request to verify the news. As a validator, they will assign a correctness value for the news. The mean of those values will be the authenticity of that news. Because of the decentralized and anonymous system, their verification will be more transparent and trustworthy. After the verification is done, the news will have an authenticity rating on the top. This rating are going to be added wherever the news is shared."

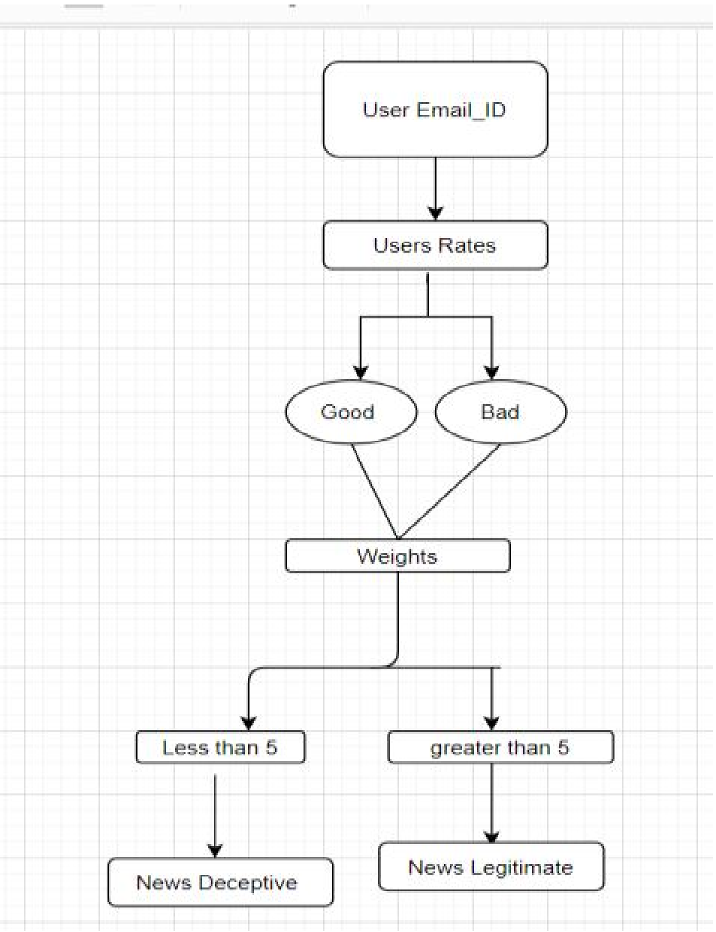


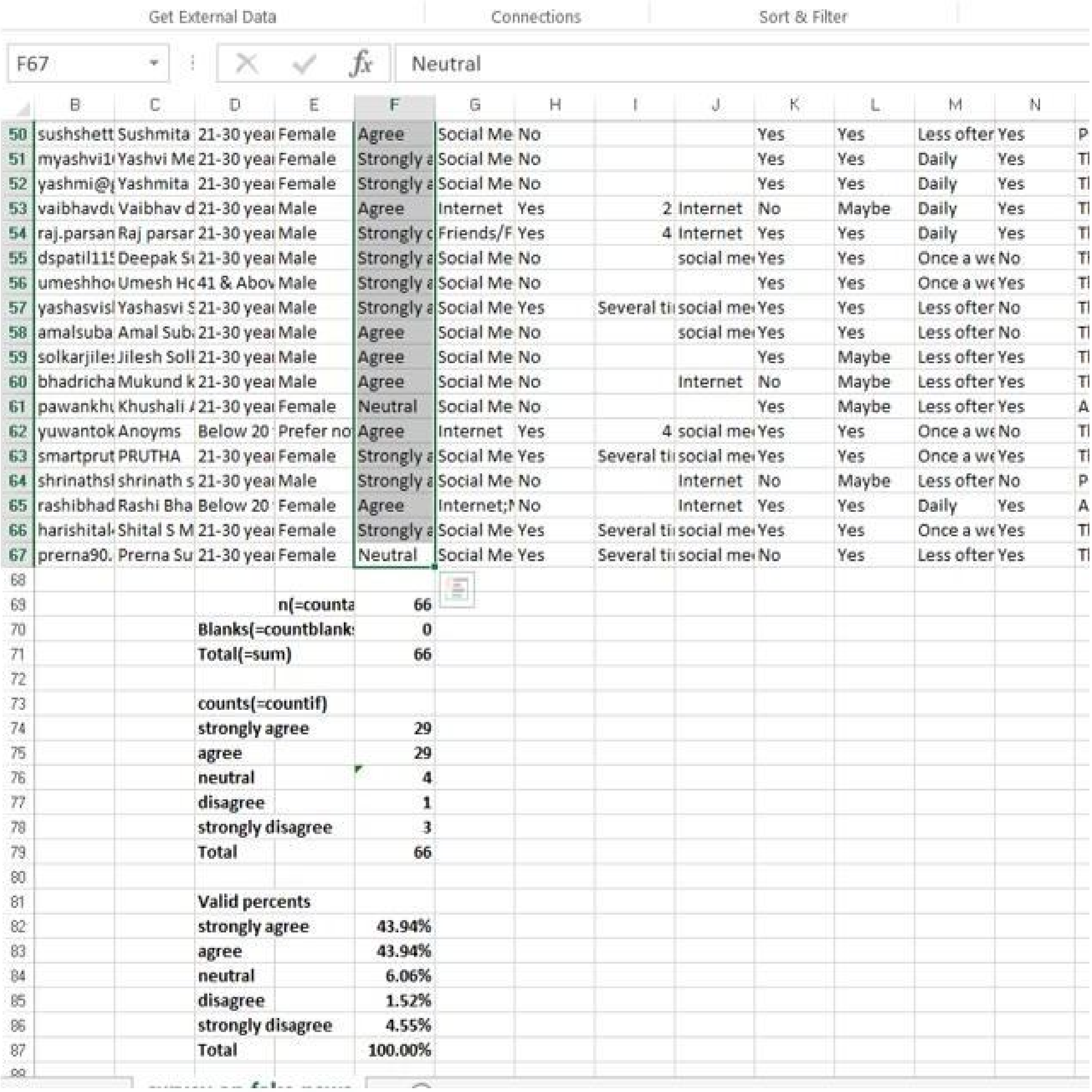
Figure 4: Flowchart

IMPLEMENTATIONS

"As we specified earlier, socialmedia are profitable to remain combined into Blockchain. From there, we will acquire the consumer Email-id. Each and every time news has been produced, it will be newscast through the chain by the operation. Now, researcher resolve analysis only those specific news' which have exceeded a precise limit of virility, example- 5000+ shares. The newscast will cover chain. Initially this news will have no grade. As time goes, validators will deliver their analyses, then the newscast will collect up with a grade to the consumers. This grade signifies legitimacy of specific newscast. Here, we propose a weight-based proof. A weight is linked to each different consumer. This weight defines the chance of being selected as a validator. Consider two people, single by weight of 4.8 and another through 4.0. The user with 4.8 weight will have an improved possibility of being nominated as an exact newscast validator than the 4.0 weighted consumer. Another weight is related to our authentication phases. The validators will remain nominated based on contiguity, i.e., the occurrence area happening which the newscast consumes been made. The validators which are from the similar area have the highest weight for grade calculation on a scale of one to five. The sense of preparing weight, the subsequent importance will go to the validators who exist the nearby neighbor to that precise area and the well-known newscast entries. It works in way that is monitored by Breadth First tree." Once the critic provides their appraisal in terms of a statistical number, their search will be multiplied with their allocated weight. All the multiplication outcomes will be added. And then the sum will be standardized to a scale of (1 to 5) to generate the last ratings. If somebody from this phase makes a blunder while spotting the false news, their weight will remain compact. This decreasing will remain based on their grade. The additional their rating deviates from the average rating of the news, the more their biases will reduce. But if their grade is nearby to the actual rating of the news then their weight will not be condensed." The additional they will be away from rating, the more their impact will be condensed."

ISSN 2394 - 7780

RESULT ANALYSIS



Neutral

	B	C	D	E	F	G	H	I	J	K	L	M	N	
50	sushshett	Sushmita	21-30 yea	Female	Agree	Social Me	No			Yes	Yes	Less ofter	Yes	P
51	myashvi1	Yashvi Me	21-30 yea	Female	Strongly a	Social Me	No			Yes	Yes	Daily	Yes	Tl
52	yashmi@	Yashmita	21-30 yea	Female	Strongly a	Social Me	No			Yes	Yes	Daily	Yes	Tl
53	vaibhavdu	Vaibhav d	21-30 yea	Male	Agree	Internet	Yes	2	Internet	No	Maybe	Daily	Yes	Tl
54	raj.parsan	Raj parsar	21-30 yea	Male	Strongly c	Friends/F	Yes	4	Internet	Yes	Yes	Daily	Yes	Tl
55	dspatil11!	Deepak S	21-30 yea	Male	Strongly a	Social Me	No		social me	Yes	Yes	Once a we	No	Tl
56	umeshho	Umesh Hc	41 & Abov	Male	Strongly a	Social Me	No			Yes	Yes	Once a we	Yes	Tl
57	yashasvisl	Yashasvi S	21-30 yea	Male	Strongly a	Social Me	Yes	Several ti	social me	Yes	Yes	Less ofter	No	Tl
58	amalsuba	Amal Sub	21-30 yea	Male	Agree	Social Me	No		social me	Yes	Yes	Less ofter	No	Tl
59	solkarjile:	Jilesh Soll	21-30 yea	Male	Agree	Social Me	No			Yes	Maybe	Less ofter	Yes	Tl
60	bhadricha	Mukund k	21-30 yea	Male	Agree	Social Me	No		Internet	No	Maybe	Less ofter	Yes	Tl
61	pawankhu	Khushali /	21-30 yea	Female	Neutral	Social Me	No			Yes	Maybe	Less ofter	Yes	A
62	yuwantok	Anoyms	Below 20	Prefer no	Agree	Internet	Yes	4	social me	Yes	Yes	Once a we	No	Tl
63	smartprut	PRUTHA	21-30 yea	Female	Strongly a	Social Me	Yes	Several ti	social me	Yes	Yes	Once a we	Yes	Tl
64	shrinathsl	shrinath s	21-30 yea	Male	Strongly a	Social Me	No		Internet	No	Maybe	Less ofter	No	P
65	rashibhad	Rashi Bha	Below 20	Female	Agree	Internet;I	No		Internet	Yes	Yes	Daily	Yes	A
66	harishital	Shital S M	21-30 yea	Female	Strongly a	Social Me	Yes	Several ti	social me	Yes	Yes	Once a we	Yes	Tl
67	prerna90.	Prerna Su	21-30 yea	Female	Neutral	Social Me	Yes	Several ti	social me	No	Yes	Less ofter	Yes	Tl
68														
69				n(=counta	66									
70			Blanks(=countblank:		0									
71			Total(=sum)		66									
72														
73			counts(=countif)											
74			strongly agree		29									
75			agree		29									
76			neutral		4									
77			disagree		1									
78			strongly disagree		3									
79			Total		66									
80														
81			Valid percents											
82			strongly agree		43.94%									
83			agree		43.94%									
84			neutral		6.06%									
85			disagree		1.52%									
86			strongly disagree		4.55%									
87			Total		100.00%									

From this we can see that users strongly agree that due to phony news the world gets affected primarily through social media.

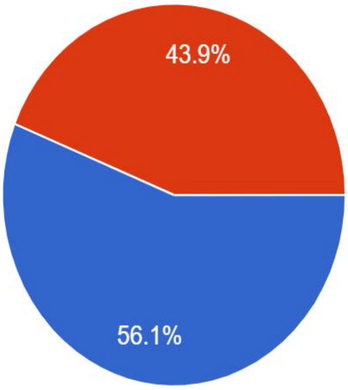


Figure 5: Victim of fake news.

According to the pie chart analysis 56.1% people are victim of deceptive news and 43% are not.

From the execution part, a final grade was generated by performing some simple calculations and using BFS, blockchain and smartcontract models. The derived score which is on a scale of 1 to 5, is allocated for broadcast. In order that one may confirm the fakeness of the news. If news becomes 1, within the grade scale of 1 to 5. The extra grade is accomplished, the more consistent the news is. If somewhat news gets 5 out of 5, then the news is powerfully trustworthy. "

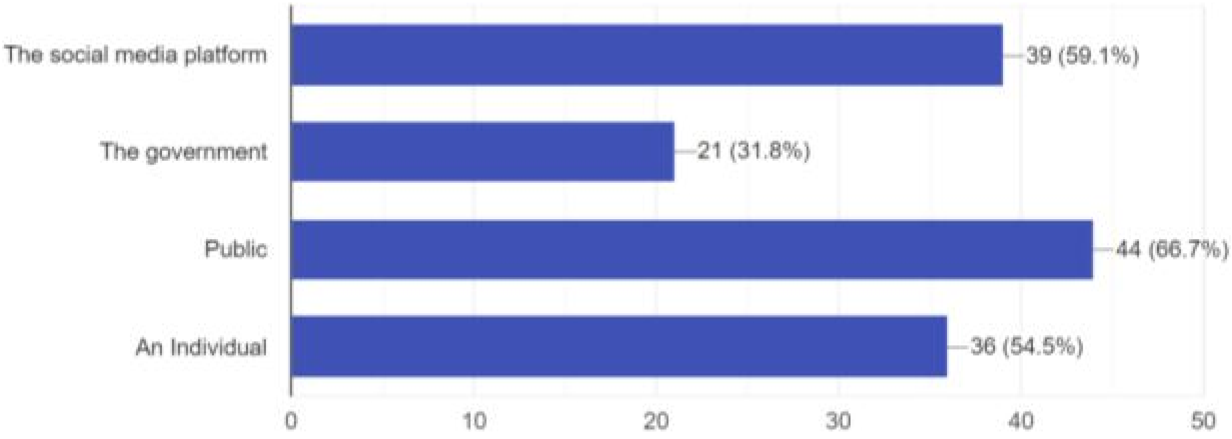


Figure 6: Identification of fake news

Figure 6 shows that fake news is spread frequently to social media platforms to the end users (39%) approx. and then commonly through public.

BOUNDARIES

"There are some significant challenges to the adoption of the blockchain for instance, political obstruction for publishing solid news against govt., therefore the regulatory of govt. Is the real challenge. Sometimes if the minors are under the political influence, they may remark fake news to be valid which has been published in their political party's support. So, using the Ethereum blockchain, it's difficult to detect the news supported politics and religion. For its veridical verification system, journals and news portals have to face job risk as it drives them to a competition of obtaining ratings. Although blockchain & Ethereum chain can save users from being misguided by reading fake news in social media, the consuming of computational power is additionally cannot be neglected."

INFERENCE

Despite having some drawbacks, the proposed method would certainly be helpful in detecting fake news in social media. This news is misguiding people just to know more views on the website to dishonestly earn extra money. Whereas my hypothesis makes sure these wrongdoings are committed.

REFERENCES

[1] Shovon-Paul,Amit kumar Das "Fake News Detection in Social Media using Blockchain", International Conference on Smart Computing & Communications (ICSCC), 2019

[2] Shlok Gilda, "Evaluating machine learning algorithms for fake news detection", Conference on Research and Development (IEEE) 2017

[3] Prannay S Reddy, Diana Elizabeth Roy, M. Keerthana , "Fake News Detection Using Naïve Bayes, SVM, Neural Networks",(Researchgate) 2019.

[4] Youngkyung Seo, Deokjin Seo, Chang-Sung Jeong , "Fake News Detection Model Using Media Reliability"(IEEE) 2018

A COMPARATIVE STUDY OF SINGLE INVESTMENT OPTION AND DIVERSIFICATION OF PORTFOLIO

Akash Kale and Prashant Chaudhary
Usha Pravin Gandhi College Of Arts Science And Commerce, Vile Parle - West, Mumbai

ABSTRACT

Once an individual comes to know he/she has enough money to survive for the next 6 months without earning a single penny he/she starts investing the rest of the money, rather some part of the saving into the investment options. India has various investment options to deal with. Some are easy and some are intricate to understand. People usually do think very much before investing their money into these options and end up in bad investment. They do not apply the rule of diversification of the investment portfolio of various options that react differently to the same economic event. The paper analyses the investment behaviour of people.

The author also researches the best investment options for age, income, and risk tolerance.

Keywords: Investment, diversification, investment Portfolio, Stock Portfolio Selection.

INTRODUCTION

Investment refers to a commitment of funds to at least one or more assets that will be held over some future period of time. There are many ways to save money like cutting down unnecessary shopping, spending on luxury lifestyle, etc. Everyone wants financial freedom and have everything they want in life and the ability to afford those things. So the investment is important because it helps in financial independence, growth of income let’s say passive income, fulfilling personal goals and reduces future risks.

Senator Elizabeth Warren popularized the so-called "50/20/30 budget rule" in her book named “ALL YOUR WORTH: THE ULTIMATE LIFETIME MONRY PLAN”. So according to 50-30-20 rule is to divide up after-tax income and allocate into needs, wants and savings respectively. That means on should spend his 50% of the money on needs for living which are must-have or must-do. And rest half to split into 20% for savings and debt repayments and 30% into everything else that you might want.

So as oxygen is necessary for life and survival the same thing investment and saving are necessary for future security and future goals. The future is always unsure and predictable and one should be always ready for the challenge. In India, there are multiple options available for investing and saving. It is necessary for investor to have adequate knowledge about the option he/she choose for investment.

Investment attracts all the people irrespective of age, occupation, gender, income, etc. Care should be taken investment not increase taxable income.

REVIEW OF LITERATURE

Several studies especially related Investments are studied by authors for this paper.

C.Saathiyamoorty, (2015) explained that the main aim of the investment is an expectation of good returns. While taking investment decision various factors must be considered like age, education, and income of the investor, etc

Mak and Lp (2017) explained financial investment behavior and sociology, psychological and demographic factors like income level, age, marital status, and investment experience influence investment behavior.

Virani (2013) determined that the income of the investor as a major factor drives the decision of the investor. Bank deposits were found imperative option as an investment to fulfill goals like kid’s education, a marriage, and security after retirement.

Bashir et al., (2013) explained relation between demographic constraints and the investment.

Patel & Patel (2012) studied people with salary. Youth with salary are aggressive investors and are interested investing with full dedication they studied the spending and saving pattern of the salaried person and investing behavior and found out that these people prefer a safe and secure investment with high risk but also yields good or high returns.

RESEARCH OBJECTIVES

The general purpose of the study is to determine the best investing option. The level of risk involved in the investment. How people invest their savings and number of investment option they choose to invest in.

The motives of this research paper are as follow:-

1. To identify if people believe in investing in only one investment option
2. Do they believe in diversification of portfolio and what all options they choose to invest?

RESEARCH DESIGN & METHODOLOGY

A. Sources of data

This study involves primary and secondary data. Primary data collected through online survey forms. Data gathered from different age group and income brackets were considered. Independent variables like age, income range, percentage of saving, expected returns, risk tolerance are few decision driving factors. These factors help to decide for investors People do invest in expectation of good returns, which makes necessary for the investor to choose the investment options carefully.

In this study, data will be collected using surveys form, from the people who invest and later on will be analyzed by statistical technique and tools like Descriptive Statistics, Percentage Method, and Ranking Method. The conclusion would be interpreted accordingly;

B. Tools used for Data Analysis

The data collected was analyzed through Percentages, frequencies and chi-square tests are applied for the analysis of data. Charts are also prepared.

C. Period of the Study

The study was conducted during October 2019 to December 2019.

D. Sampling Design

The study covered the investors from Mumbai, Pune. The population represents all the investors investing in different options with different age group, income slab, and percentage of investment, occupation, and education.

LIMITATIONS OF THE STUDY

People usually don't like to discuss their income and investment so primary data collection was a challenge. Only 100 people responded to the survey.

HYPOTHESIS TESTING

H1. There is no significant relationship between education and portfolio diversification.

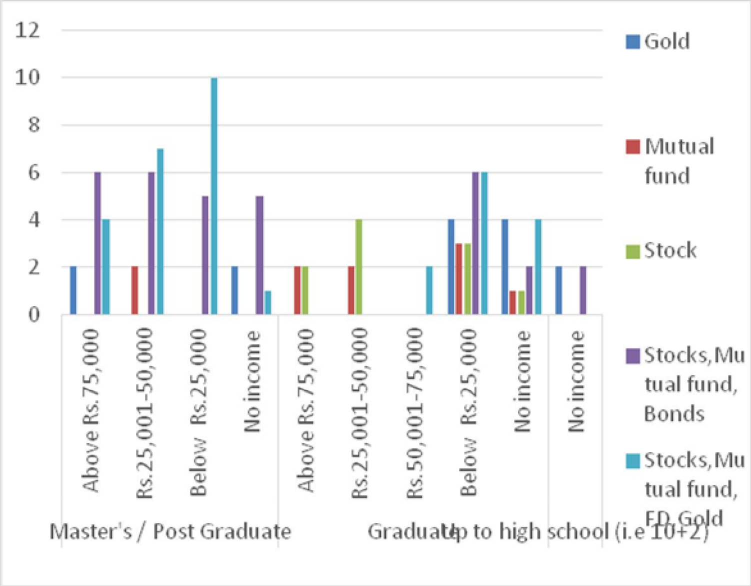
H2. There is a significant relationship between education, income, and portfolio diversification.

ANALYSIS & FINDINGS

Table 1 Investment distribution according to the age.

	Gold	MF	Stocks	Stocks, MF, FD, Bonds	Stocks, MF, FD, Gold	Total
PG	4	2	1	21	22	50
Graduate	8	12	14	4	8	46
10+2	2	0	0	2	0	4
Total	14	14	14	28	30	100

Source: Primary Data.



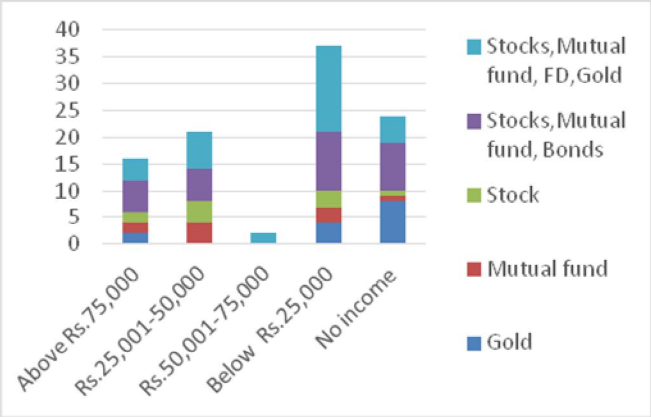
This Table indicates that people with higher education do prefer more than two investment options.

According to the above analysis 55% of people who are graduate and post-graduate/ Master's choose diversification of the portfolio over a single investment option

Table 2. Education and Income vs Investment options.

	Gold	MF	Stocks	Stocks, MF, Gold	Stocks, MF, FD, Gold	Total
75k above	2	2	2	6	4	16
50k-75k	0	0	0	0	2	2
25k-50k	0	4	4	6	7	21
Below 25K	4	3	3	11	16	37
No income	8	1	1	9	5	24
Total	14	10	10	32	34	100

Source: Primary Data.



The above table reveals very few people invest in a single investment option. As it is always recommended by experts to not invest all in a single investment option. The table also indicates as people do invest in more than two investment options. The table also indicated that there is a relationship between income and diversification of the portfolio. People having higher education and good income prefer investing in various investment section. For example people, 39% of the people of the total sample having income ranging from Rs.25000-Rs.75000 believe in diversification. This shows that higher education

Null Hypothesis 1

There is no significant relationship between education and portfolio diversification.

Chi-square value = 45.40023

Table Value = 15.51

Significance level = 5%

RESULT

The χ2 value is larger than the table value thus we reject the hypothesis. Therefore there is a significant relationship between education and diversification of the portfolio.

Null Hypothesis 2

There is a significant relationship between education, income, and portfolio diversification. So the researchers accept this hypothesis.

Chi Square =22.81014

Table Value = 26.30

Significance level = 5%

RESULT

The χ2 value is larger than the table value thus researchers accept the hypothesis. Therefore there is a significant relationship between income and diversification of the portfolio.

CONCLUSION

So from the above-given analysis, we can derive that there is a significant relationship between incomes, education versus investment portfolio. Usually, people with higher education and good income think pragmatically and make the decision suitable for the environment.

Also, a single option investment is way riskier to invest in. Because a change in trend, rules or any factor affecting the performance of the option can result in your portfolio to bleed red i.e. facing loss. Whereas investing your saving by dividing it into various options like Mutual Fund, Fixed Deposit will not affect much if there is any economic change and can bring you good returns as compared to a single option investment.

ISSN 2394 - 7780

VIRTUAL REALITY THERAPY FOR TREATMENT OF PHOBIA

Yashvi Mehta[1] and Smruti Nanavaty[2]
M.Sc. [I.T.] Student[1] and M.Sc. [I.T.] Coordinator[2], Usha Pravin Gandhi College

Virtual reality treatment (VRT), has been as of late utilized in the treatment of subjects determined to have fear, a confusion that is described by stamped tension upon introduction to certain situations, object, creature or movement. Research predominantly centers around the plausible utilization of VRT for treatment of fear when contrasted with other ordinary strategies like psychological conduct treatment, prescriptions, and so on. Viable computer generated experience frameworks are recommended to conjure nearness, which in term inspires an enthusiastic reaction, assisting with driving a fruitful treatment result.

Presentation treatment is one strategy for conduct treatment utilized for treatment of fears using technology called Virtual reality to create patient's dreaded condition such that the patient is able to communicate within the virtual world with no risk.

INTRODUCTION

A phobia is a type of anxiety disorder. It is a relentless, persistent, excessive, unrealistic fear of an object, person, animal, activity or situation which causes one to want to avoid it or endure it with copious amount of anxiety and/or distress. [10,11]

Some phobias are very limited and specific. They can be majorly be classified into three types - agoraphobia (fear of being outside), social phobia (fear of public speaking, having to meet unknown people, or other social situations), and specific phobias (fear of other items or situations).[1]

For example, a person might fear cats (ailurophobe) or dogs (cynophobia). In such cases, the individual lives generally liberated from uneasiness by evading the thing he/she fears. Some phobias though are much more difficult to avoid. For example, fear of heights can suddenly be triggered by driving over bridge or by looking outside the window of a place of business . Claustrophobia (fear of closed/confined spaces) can be triggered when riding in an lift or using a small washroom.

Phobias are diagnosed when anxiety in response to their fear is so intense that it causes significant distress or hampers normal functioning of the person. Phobias are highly treatable as compared to all other psychological disorders. Phobias can be triggered by daily life activities and/or life situations. They often tend to run in families.

Often, phobias often go unreported. Despite that, statistics for people who have phobia are estimated to be more than 6 million in United States alone.

If left untreated, phobia might affect the person’s life deeply. They may alter their lives drastically and attempt to hide or avoid it, leading to problems with health both physically and mentally, job location, social and recreational activities, daily tasks such as changing the route to the office while struggling to cope.

People often get panic attacks, cold sweat, black outs when triggered with their phobia. Some even pass out. Intensive psychic symptoms (loss of control, fear, anxiety) and vegetative symptoms (palpitations, fainting, sweating, breathing issues).

Phobias are evaluated by health care professional who analyze the symptoms, conduct medical interviews and also physical examinations.

The treatment of phobias include clinical treatments such as the use of cognitive behavioral therapy, medications, desensitization. These medications include beta-blockers, antidepressants ,tranquilizers,etc.

The most likely way for people to deal with phobia is simply avoiding the trigger/cause of their phobia. They try to make positive self-statements and talk to peers about it.

A. Cognitive Behavioral Treatment

Cognitive behavioral treatment (CBT) for phobias try to separate anxiety response and cause. CBT achieves this partly by identifying problematic or irrational thinking patterns and helping people learn more adaptive ways of thinking/imagining about triggers. Therapists provide them with behavioral methods to help them face their fear without anxiety.

B. Medication

Medications are not the permanent solution for phobia. They are often prescribed on a short-term basis to treat effects of phobia such as anxiety. Such medications are classified in three types namely:

1. antidepressants
2. tranquillizers
3. beta-blockers

C. Antidepressants

These are often prescribed to help reduce anxiety. Selective serotonin reuptake inhibitors (SSRIs) are most often prescribed to treat anxiety, social phobia or panic disorder

Common side effects of these treatments include:

1. Insomnia
2. Agitation
3. Nausea
4. Restlessness
5. Upset stomach
6. Headaches

Sometimes initially, these medications make your anxiety worse and can possibly cause sexual problems.

Drugs like Anafranil (Clomipramine) can cause

- Dry mouth
- blurred vision
- Drowsiness
- Tremors
- Constipation
- Irregular heartbeat (palpitations)
- Urinary hesitancy (difficulty in urinating)

They can be lowered slowly. It is dangerous to stop taking antidepressants suddenly.

D. Tranquillizers

Benzodiazepines are a group of medicines that are categorized as minor tranquillizers. They are used temporarily at the lowest possible dose for treating severe anxiety.

Tranquilizers too, should be stopped gradually to avoid withdrawal symptoms.

E. Beta-blockers

Beta-blockers lower your heart rate and blood pressure. Inderal (Propranolol) is a common beta-blocker to treat anxiety. Possible side effects include:

- Stomach problems
- Tiredness
- Insomnia
- Cold fingers

F. Virtual Reality

Virtual Reality (VR) is a 3D environment generated using computers. The VR works by sending information, such as images or sounds, to a user’s senses so that a brain of the user is truly in the generated environment. This understanding can be proven from interactivities or responses of the user [2].

We can create virtual environments (VEs) using the advanced computers we have. VEs can create a scenarios similar to the actual phobic stimulus. [3].

Virtual reality therapy (VRT), has recently been suggested for the treating specific phobias. Virtual Reality Therapy is based on systematic desensitization. VRT provides situations for the patients who have difficulty imagining situations well. It allows greater control over gradual exposure to the stimulus in vivo at much cheaper rates than having it live in real world. [4,5].

LITERATURE REVIEW

Clinical psychology defines phobia as an anxiety disorder, mainly characterized by intense irrational fears triggered from specific objects or situations. The intensity of fear is not directly proportional to the potential amount of danger of the trigger. All phobias have a common characteristic :-the concern that the feared object may also occur outside the present place and time which leads to anticipated fear [6].

Agoraphobias are characterized by fear of situations or places. Say anything from small confined places or large overcrowded places. In such cases, patients' concerns are often regarding immediate help not being available. In advanced cases, patients isolate themselves from outside world. They do not step out of their house.

Second category is social phobias. Traits like low self-esteem, fear of people, overcrowded places, fear of critique, or fear of different modes of transportation are often found in patients with social phobia. Such phobias can also be induced by public performances or at times, they can stem from simply communicating with an individual. All of the mentioned situations discourage patient to have any form of social contact. In some cases, the social phobia is not necessarily caused by fear of other people, instead, it is caused by the patient themselves. Patient's failures or the thought of what others might think of them is a cause for this. Small things like a presentation in front of a small group might mean panic attacks and high-level anxiety for patients suffering from social phobia.

Third category of phobias is isolated phobias also known as specific phobias. These phobias are caused by fear of certain situations and objects. An example of such phobias is fear of germs, spiders, snakes, bacteria, darkness, blood, etc.

Phobias even though are classified into categories, they cannot always be assigned one specific category. They can fall into two or more categories. Let's take an example: A patient feels extreme levels of anxiety while crossing through a busy junction. Patient's feet are trembling and thus, he/she has difficulty breathing and is unable to think rationally. Thus, the patient fled the place after panicking. This is a typical reaction to Agoraphobia, provided that the reaction is caused by a fear of large open spaces, while away from the safety of your home. The situation, could also be categorized as social phobia, if the reason for panic and anxiety was being in the presence of large number of people or being at the center of attention . The most common method for treating phobia is psychotherapy along with pharmacotherapy, which is a mixture of CBT, relaxation procedures and psychosocial methods.

New practical techniques to overcome pathological fears have also been developed. The main aim here is to provide the patient with correction experience where patient is forced to face the item or situation that causes their phobia which is created in an artificial environment.. Sterling technique is systematic de sensibilization. It is divided into three parts:

1. Evoke a comfortable state.
2. Create a set of situations such that it gradually evolves in content and intensity from menial harmless situation to very fearsome situation causing a phobic response.
3. Get patient accustomed to the created situations and having them endure it in a as pleasant state as possible. Getting the feared circumstance together with an enthusiastic state, which for the most part oozes dread, happens in genuine circumstances - "in vivo" or in circumstances most distinctively envisioned by patients - "in vitro" in research facility.

Excessive fear is suppressed by purposeful and repetitive imagination of ideas. Circumstances that produce dread alongside muscle and mental unwinding, despite what might be expected of dread in mental limit, are motoric and autonomic . [7 ,8]

In psychotherapy there are various other methods; implementation of them can be combined with newer types of treatment - therapy through virtual reality.

Due to the fact that some human senses, such as sight and hearing may be easily deceived, it is easy to simulate objects, situations and environments for patients that would cause phobia with the help of the tools of virtual reality. This method of therapy may include a number of features such as computer graphics in real time, position control devices, sensor inputs, etc., that allow the patient to supposedly "dive" into the environment carefully modified by computers.

In psychotherapy there are different techniques; usage of them can be combined with more current sorts of treatment - treatment through Virtual reality.

Human senses, for example, sight and hearing can be effortlessly deceived. It is anything but difficult to recreate scenarios, circumstances and conditions for patients that would cause fear with the assistance of Virtual reality. This technique for treatment may incorporate various highlights, for example, position control devices, computer graphics in real time, sensor inputs, etc. These highlights will allow the patient to "dive" into the environment and experience it like real time.

METHODOLOGY

a. Data Collection

The data collected is primary data. This data is collected using survey method using Google forms as the tool. The survey was conducted basically with idea to understand various factors related to phobia and to have an answer to questions like does phobia affect the normal life of people, the most common phobia seen in today's generation and also does the medical or therapy factors help a person to cure phobia. While doing this it was seen that most people mix phobia with many other things and don't really understand it well. In order to avoid this confusion, the people were given as many details as possible during filling the survey and before the question being asked in order to avoid confusion. The sample size is of a 100 responses. Here, random sampling technique has been used for picking up the samples. Random sampling techniques is seen to work better than many other techniques where authors of the research have a diversity in the population data and the sample needs to be selected without any bias. The survey had a variety of options and also it was given a choice for the user to answer something else apart from the options listed to be flexible of choice.

b. Gender grouping

To understand the effects of phobia researchers divided the analysis on two different groups, the male and female group. Data in both the groups were consisting of equal data points and the analysis made was to understand that how much effect it has on one basic differential personality group. The statistical results turned out to be accurate enough to conclude a point.

c. Hypothesis

Hypothesis are the educated assumptions made while researching on the treatments being opted for and the treatment outcome on the basis of the treatment applied. Here basically the researchers looked into two topics mainly being the choice of treatment and its result. Having two parameters in hand researchers decided to apply Ztest to their problem area. The null hypothesis is set that the taken treatment is successful. Accordingly, the alternative hypothesis suggests that the taken treatment is not successful. The null hypothesis if accepted then we reject the alternative hypothesis and if the alternative hypothesis is accepted then we reject the null hypothesis.

ANALYSIS

a. Gender being more affected by phobia.

The sample data was equally grouped and furthermore clustered into groups of male and female who are affected by phobia. A scale of distinct values was chosen in order to see the difference clearly. The following graph gives a visual observation. Authors of the research observed that more male are affected than females by phobias

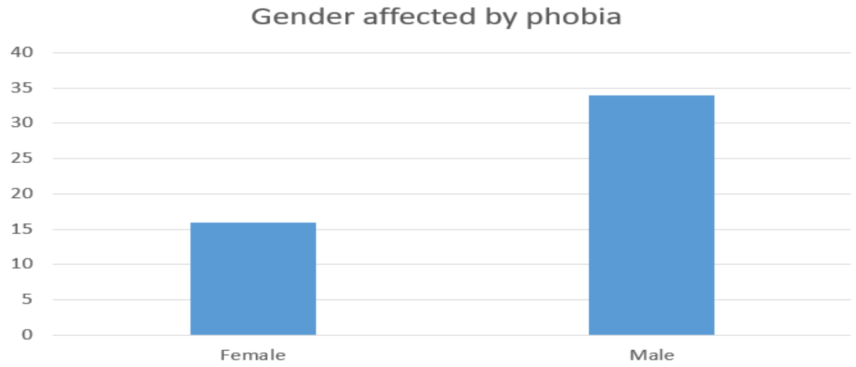
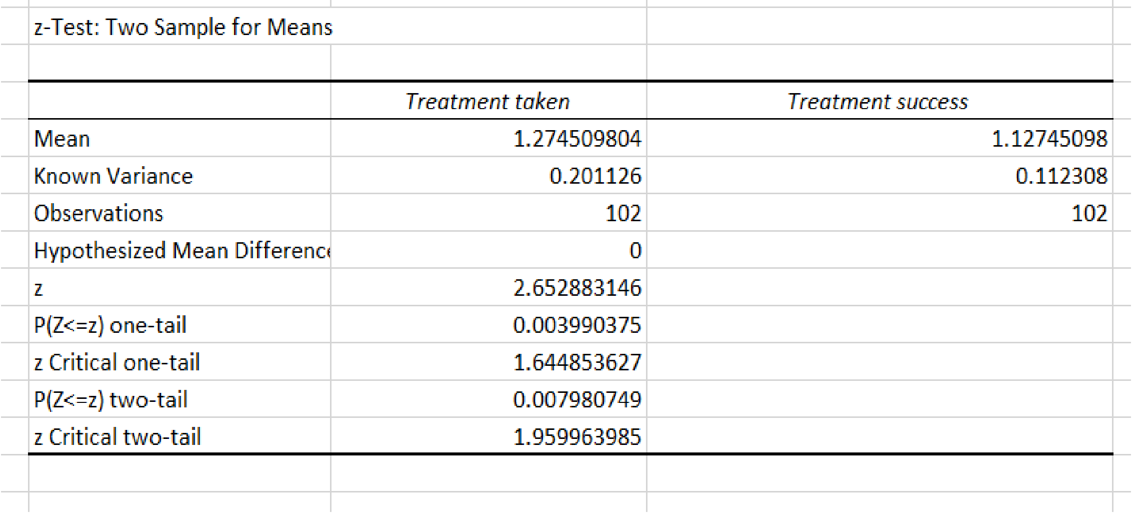


Figure 1: Gender affected by phobia

ISSN 2394 - 7780

b. Hypothesis.

The null hypothesis is set that the taken treatment is successful. Accordingly, the alternative hypothesis suggests that the taken treatment is not successful. The null hypothesis if accepted then we reject the alternative hypothesis and if the alternative hypothesis is accepted then we reject the null hypothesis. In our case we took two parameters that explained this situation. One being that how many people actually go for a clinical or psychological therapy treatment in order to treat their phobia. And the other parameter or factor being that laid importance was the outcome of these tests, were they really helpful or was the outcome as expected for the people being treated. If so then the tests were successful else not as said in the hypothesis. Performing Z Test for two sample means said that if the p value is greater than or equal to the alpha value which is 0.05 then the null hypothesis is accepted that the test is successful and the alternative hypothesis is rejected that the test is not successful. But if the p value is less than the alpha value which is 0.05 then we reject the null hypothesis that the tests are successful and accept the alternative hypothesis that the tests are not successful. In this case, the researchers observed that we rejected the null hypothesis and accepted the alternative hypothesis as the p value was less than the alpha value that is 0.05.



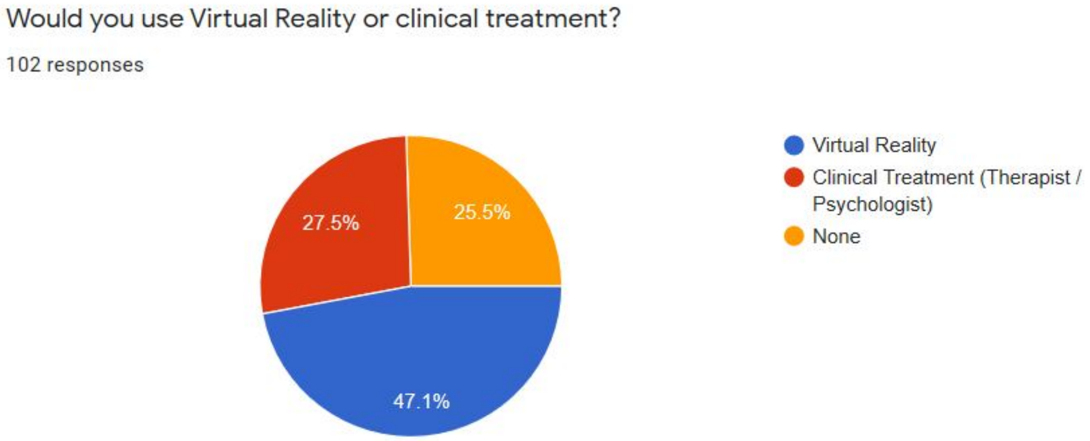
z-Test: Two Sample for Means

	Treatment taken	*Treatment success*
Mean	1.274509804	1.12745098
Known Variance	0.201126	0.112308
Observations	102	102
Hypothesized Mean Differenc	0	
z	2.652883146	
P(Z<=z) one-tail	0.003990375	
z Critical one-tail	1.644853627	
P(Z<=z) two-tail	0.007980749	
z Critical two-tail	1.959963985	

Figure 2: Hypothesis test.

C. Choice of VR.

Firstly, and most importantly, people do not prefer to go to therapists and especially talk about it. Going to therapists is still considered as bad since people start associating the work sickness when going for therapy. Therapy is for people who need help, not sick people. Unfortunately, the stigma associated with it hasn’t gone away. People would much rather spend their entire lives in fear, stress and avoiding their triggers at all costs. Another disadvantage to it is in treatments like CBT, people have to talk. It’s not easy to talk about things you fear much less talk about it continuously to stop getting afraid of them. Medications have associated side effects to them. Thus, Virtual Reality may also help in turning away from such a stigma as it provides the client with an immersive game-like experience which may be a little bit more welcomed by the larger masses and it is a refreshing choice to navigate through all of those disadvantages. VR gives much more control as compared to putting yourself in live situations which are difficult to control.



CONCLUSION

As per the first hand survey, based on our analysis, it is highly likely that people are more likely to use Virtual reality as an aid to phobia as compared to other traditional methods. Virtual reality offers several advantages over the actually putting the patient in real time uncontrollable situations. The treatment can be carried out in the therapist's office. This is cheaper for the patient and less restrictive for the therapist. They do not need to accompany the patient to expose them to real world stimulus. This allows better immersive experience as compared to plain therapy.

VRT can help the lives of many people. It holds advantages over other treatment like its cost effectiveness and control, ease of use.

REFERENCES

[1] International classification of diseases 2013. [On-line]. [19] M. North, S. North, and L. Coble, “Effectiveness of virtual environment desensitization in the treatment of agoraphobia”, [September 11, 2015].

[2] Virtual Reality Society, "Virtual Reality: what is it and why is it important to know about?" Virtual Reality, 2016.

[3] M. M. North, S. M. North, and J. R. Coble, Virtual Reality Therapy: IPI Press, 1996.

[4] B. O. Rothabaum, L. F. Hodges, R. Kooper, D. Opdyke, J. S. Williford and M. M. North, “The efficacy of virtual reality graded exposure in the treatment of acrophobia,” Amer. J. Psych., vol. 152, pp. 626–628, 1995.

[5] M. M. North and S. M. North, “Relative effectiveness of virtual environment desensitization and imaginal desensitization in the treatment of aerophobia,” Arachnet Electr. J. Virtual Culture, vol. 2, 1994.

[6] A. Heretik, “Úzkostné (neurotické) poruchy”, in: Heretik, A. sr., Heretik, A., jr. (eds) a kol. Klinická psychológia. Nové Zámky: Psychoprof, spol. s.r.o., 2007, pp. 217–241.

[7] S. Kratochvíl, Basics of Psychotherapy. (Základypsychoterapie.) Praha: Portál, 4. vyd., 2002.

[8] J. Langmeier, et al., Child psychotherapy. (Dtskápsychoterapie.) Praha: Portál, 2. vyd. 2000.

[9] Submitted to Management Resources College Student Paper

International Journal of Advance and Innovative Research
Volume 7, Issue 1 (XIV): January - March, 2020

ISSN 2394 - 7780

A SURVEY ON CHOICE OF ELECTRICAL SERVICE PROVIDER BASED ON CUSTOMER SATISFACTION (TATA VS ADANI)

Vaibhav Dubey and Prashant Choudhary
Usha Pravin Gandhi College Arts, Science and Commerce

ABSTRACT

The Electrical service providers adopting fundamental changes in this country. Because of which the competition had being generated among various company. Researcher had observed that the electrical service providers was not able to full fill all the needs. Earlier the Author [1] were described the various methodology to classify Electricity Customers based on clustering algorithm and self-Organizing maps. Author [2] was describing the clustering technique to classify Electricity customer for load forecasting. In this situation the researcher was Classifying the electrical service provider. Customer satisfaction was very much important because this was a key factor which describing the customer relationship with its current service provider.

In this paper the researcher will classifying best electrical service provider based on its services. The services were depending on several parameters based on its functionality. And this was classifying best electrical service providers. The researcher's assumption that there was no impact on consumer regarding the electrical services. After this researcher was determine the statistical analysis for generating the report. After the testing the

researcher had accepted or reject the assumption based on the analysis report.

Keyword: Customer satisfaction, Customer problems, Best electrical utility.

INTRODUCTION

Tata power sector had seen many challenges over the last 10 decades. While passage of the Electricity Act 2003, addressed some concerns of the industry, many issues are yet to be addressed. Today, the sector is reeling with several challenges. Poor financial outlook of Discoms where the losses have been increasing to levels far higher than previous years is a matter of great concern. Power distribution stillremainsasegmentthatneedssignificantreform-intervention.Tata Power has exhibited exponential growth despite the challenges faced by the power sector.

AdaniGroup is exploringacquisitionof Vidarbha Industries Power (VIPL), a subsidiary of Reliance Power which supplies electricity to Adani Electricity Mumbai ,VIPL operates two 300 megawatt (MW) units at Butibori in Maharashtra, but it has not been generating any power since mid-January after Coal India stopped supplying coal over an ongoing litigation which is in the Delhi High Court and issues relating to payment. VIPL had shut its first unit of 300 MW before that. The spat has forced Adani Electricity to buy power from the open market to meet its obligation to its 3 million customers in the city.

The researchers had classifying the best electrical service provider based on customer satisfaction. There was a competitive environment where different service providers providing its service as per the consumer needs. The electrical utility had been measured by its services, and the services was depending on several attributes such as power cut off, customer care services, extra cost, maintenance, etc.

In which consumer impact was dependable on its current electrical service provider. The rules and policies were also important of electrical service providers. The measurable attributes were depending on the consumer respond which was calculated. The quantitative measures were taken by the researcher for better accuracy of electrical service provider. There were so many electrical power agencies such as TATA power, ADANI electricity, NTPC, NHPC, etc. The researcher was gathering different responses of each consumer depends on its electrical service provider. The researcher was investigating the phenomena and generating insight on respondent data .The level of consumer satisfaction had little impact on their profit or market share as it would in the case of regular product or services in competitive markets .Therefore the need had emerged to conduct in regular intervals interaction in terms of survey , online forms etc. of consumer satisfaction in order to recover consumer needs as well as to handle and moderate the quality of service provided by the different electrical distributor companies . The researcher was using different mathematical model and this model was responsible to measure the accuracy. The mathematical model was applied over the variables, and generating report over the assumption by the analysis as per the researcher objectives.

The different methodologies were used for analytical procedure, in this scenario researcher was used different testing for generating on researcher's objective.

LITERATURE REVIEW

With reference to [1] in 2002 conclusion, application of clustering algorithm and self-organizing maps were used to classify electricity customers in this the author was classifying the consumer into different groups. Based on consumption pattern of the consumer into different classes. The technique was developed and each consumer had clustered into different groups based on its tariff. The researcher was used hierarchical clustering and Fuzzy k-means methodology for isolation of two different cluster.

[2] Gave an overview about application of clustering technique to electricity customer classification for load forecasting with the development of smart grid and opening-up Progress of the electricity market, customer relationship management plays a more and more Important role. Clustering to key customers is an effective way of customer analysis to power Enterprises, especially under the background of big data in power system database. This paper introduces two kinds of commonly used clustering algorithms. Then clustering is performed with two propose algorithms under given number of clusters. In this the researcher was used consumption patterns defined, clustering techniques are generally applied to perform consumption grouping. In this section two mostly used clustering technique is used for consumption pattern. The main purpose of clustering was to improve the similarity of key customer in the same cluster as much as possible and lower it among different cluster.

[3] In this paper the present practices of “Trouble call management” based on manual processing. One of the reasons for the problem reported by the customer not being properly attended was the lack of ready availability of customer data. The Andhra Pradesh state electricity board introduced Trouble call system. In this paper it was concluded that automating the complaint handling system will provide the benefits of increase in the efficiency and production, Improve customer relations and Better accountability.

[4] In this research paper the customer satisfaction in the Hungarian electricity distribution.

The satisfaction and importance should be measured separately in the household and business Segments. The customer satisfaction index should be a comprehensive measuring tool, which Express not only the satisfaction with the core offering of the electricity supplier.

The first point of interest in the survey was to find out how do consumers perceive the importance of electricity supply among other major public utilities and the satisfaction compared to the other utilities. The result of the survey indicates that the importance of electricity supply was on the top in both main segment of consumers.

The survey has also made it possible by revealing the structural difference in averages at both national and enterprise level.

RESEARCH METHODOLOGY

In this section, we will discuss the methods we use to prove whether our NULL hypothesis or Alternate Hypothesis is accepted or rejected. For this we collected quantitative data through surveys i.e. Google form.

H0: There is no impact on a consumer of electrical service providers.

H1: There is an impact on a consumer of electrical service providers.

Hence, we have to measure certain attributes of electrical service providers to estimate the performance. Because of which we can easily classify the best electrical service provider between different electrical power agencies.

To prove our Hypothesis, we kept a sampling frame of 57 people and the exact sampling we collected was from 51 people.

List of questionnaires used in my survey –

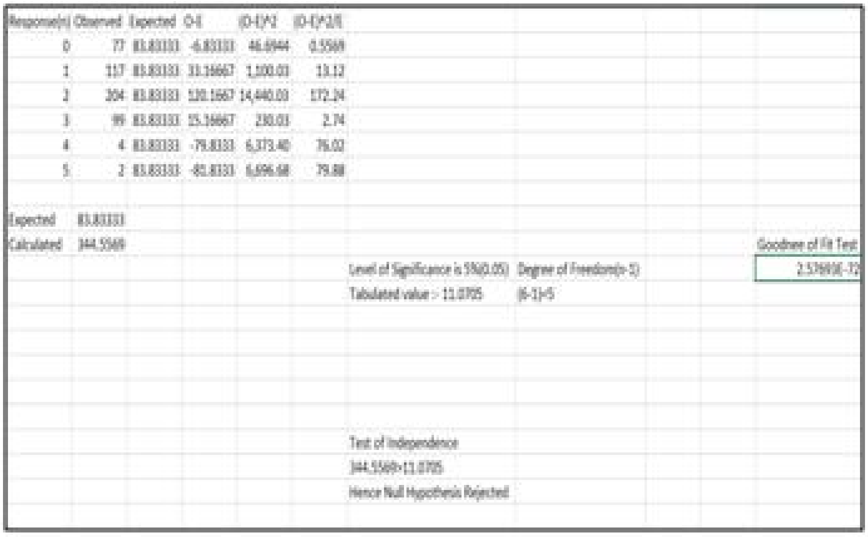
1. Current electrical service provider?
2. On what basis have you selected your current electrical service provider?
3. Are you satisfied with your electrical service provider?
4. If no, what is the problem?
5. Rate your electrical service provider.
6. From how many years are you using your current service provider?
7. Which method do you prefer for payment of your electricity bill?

ISSN 2394 - 7780

8. If Online, do you get any benefits?
9. Is the customer care service of your provider available 24*7?
10. How frequently does power cut-offs takes place per day?

RESULTS

On the basis of data acquisition from the responses the researcher did the quantitative analysis on a system variable. Because of which the researchers proved the hypothesis using Chi-Square Test for statistical analysis. From the online survey the researcher got 52 responses and, in the Chi-Square Testing the threshold sample was 50 responses. Hence, this is the reason why the researcher used Chi-Square Test. In the Chi-Square Testing the researchers was used test of independence to prove its assumption.



Response(n)	Observed	Expected	O-E	(O-E)^2	(O-E)^2/E
0	77	83.83333	-6.83333	46.6944	0.5569
1	117	83.83333	33.16667	1,100.03	13.12
2	204	83.83333	120.1667	14,440.03	172.24
3	99	83.83333	15.16667	230.03	2.74
4	4	83.83333	-79.8333	6,373.40	76.02
5	2	83.83333	-81.8333	6,696.68	79.88

Expected 83.83333
Calculated 344.5569

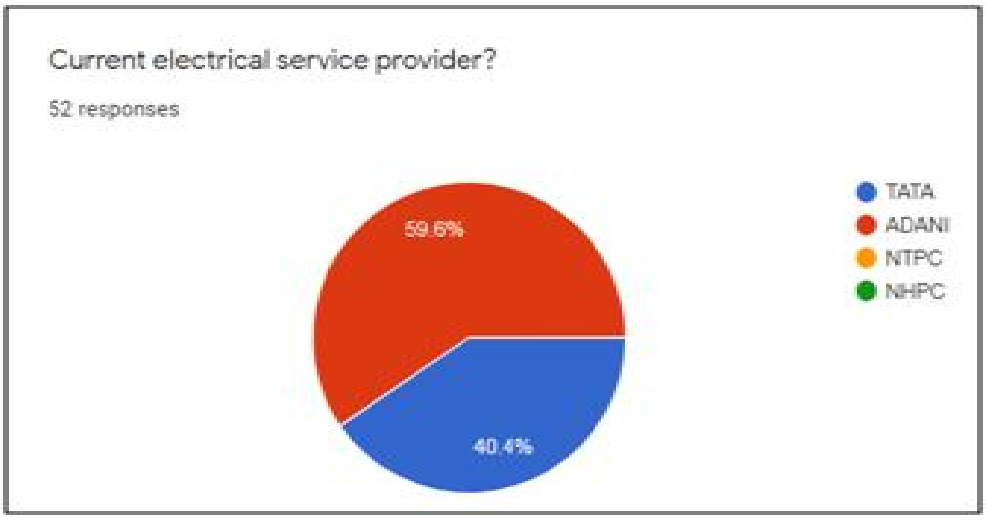
Goodness of Fit Test
2.57693E-72

Level of Significance is 5%(0.05) Degree of Freedom(n-1)
Tabulated value :- 11.0705 (6-1)=5

Test of Independence
344.5569>11.0705
Hence Null Hypothesis Rejected

From above data, Chi-Square value =344.5569.

Tabulated value of χ2 test for 5 Degree of Freedom at 5% level of significance is 11.0705. So, our calculated value of χ2 test is 344.5569, it is highly significant and Null Hypothesis is rejected at 5% level of significant.



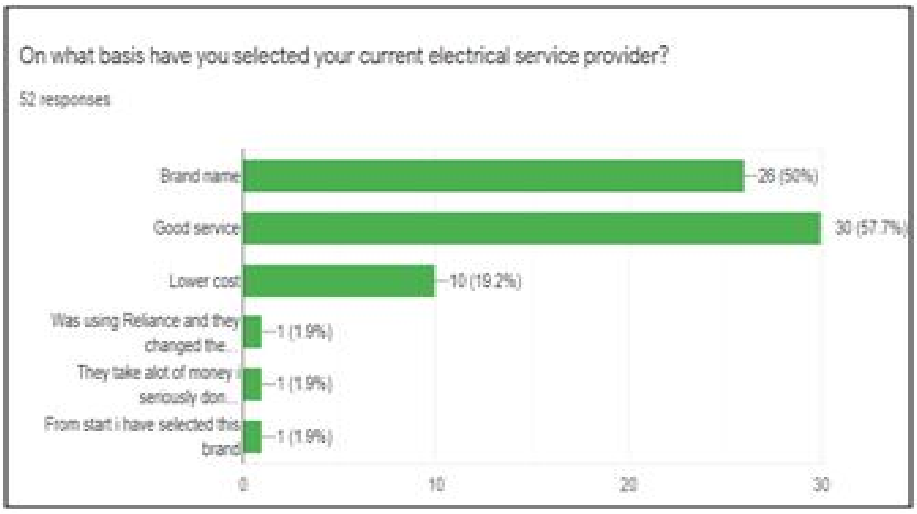
Data collection

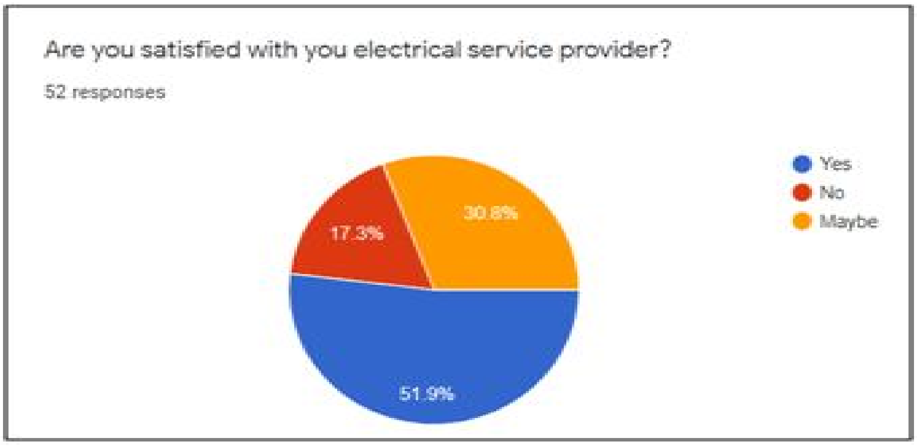
This data was collected through Google form and responses were collected from 70 people. There were total 17 questionnaire from which we selected these 10 most important questions needed to prove our objective.

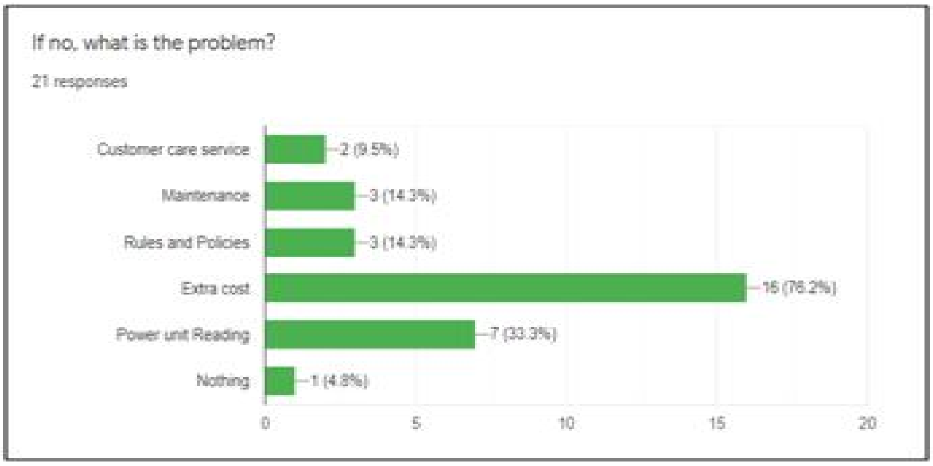
Detailed Description of questionnaire and responses collected are given below:

Conclusion

In this paper the researcher presented the significant difference between different electrical service







provider based on its measurable attributes .The TATA and ADANI both were providing its services to each consumer in a particular region but both the companies were not able to fulfill all the needs of consumers and because of which the consumer was not satisfied with its current electrical service provider . From various options of a problems the extra cost is an important factor for consumers.

This is the reason why the researcher was classifying the best service provider among TATA and ADANI. In current trend there is a competitive environment among various electrical power agencies. And the consumers were not able to do the classification between the electrical service providers, Therefore the researcher was used this methodology for choosing the best electrical utility. As per the researcher there are only two companies that are giving best services to consumers by the Google Form Survey. The researcher applies statistical analysis on a few groups of people therefore the researcher got this type of outcome, there are also several companies in India like NTPC, NHPC, etc. And finally, the researcher wants to conclude research with some Suggestions/Feedback given by respondents in image given below:

MANUSCRIPT SUBMISSION

GUIDELINES FOR CONTRIBUTORS

1. Manuscripts should be submitted preferably through email and the research article / paper should preferably not exceed 8 – 10 pages in all.

2. Book review must contain the name of the author and the book reviewed, the place of publication and publisher, date of publication, number of pages and price.

3. Manuscripts should be typed in 12 font-size, Times New Roman, single spaced with 1" margin on a standard A4 size paper. Manuscripts should be organized in the following order: title, name(s) of author(s) and his/her (their) complete affiliation(s) including zip code(s), Abstract (not exceeding 350 words), Introduction, Main body of paper, Conclusion and References.

4. The title of the paper should be in capital letters, bold, size 16" and centered at the top of the first page. The author(s) and affiliations(s) should be centered, bold, size 14" and single-spaced, beginning from the second line below the title.

First Author Name1, Second Author Name2, Third Author Name3

1Author Designation, Department, Organization, City, email id

2Author Designation, Department, Organization, City, email id

3Author Designation, Department, Organization, City, email id

5. The abstract should summarize the context, content and conclusions of the paper in less than 350 words in 12 points italic Times New Roman. The abstract should have about five key words in alphabetical order separated by comma of 12 points italic Times New Roman.

6. Figures and tables should be centered, separately numbered, self explained. Please note that table titles must be above the table and sources of data should be mentioned below the table. The authors should ensure that tables and figures are referred to from the main text.

EXAMPLES OF REFERENCES

All references must be arranged first alphabetically and then it may be further sorted chronologically also.

- **Single author journal article:**

Fox, S. (1984). Empowerment as a catalyst for change: an example for the food industry. *Supply Chain Management*, *2*(3), 29–33.

Bateson, C. D.,(2006), 'Doing Business after the Fall: The Virtue of Moral Hypocrisy', Journal of Business Ethics, 66: 321 – 335

- **Multiple author journal article:**

Khan, M. R., Islam, A. F. M. M., & Das, D. (1886). A Factor Analytic Study on the Validity of a Union Commitment Scale. *Journal of Applied Psychology, 12*(1), 129-136.

Liu, W.B, Wongcha A, & Peng, K.C. (2012), "Adopting Super-Efficiency And Tobit Model On Analyzing the Efficiency of Teacher's Colleges In Thailand", International Journal on New Trends In Education and Their Implications, Vol.3.3, 108 – 114.

- **Text Book:**

Simchi-Levi, D., Kaminsky, P., & Simchi-Levi, E. (2007). *Designing and Managing the Supply Chain: Concepts, Strategies and Case Studies* (3rd ed.). New York: McGraw-Hill.

S. Neelamegham," Marketing in India, Cases and Reading, Vikas Publishing House Pvt. Ltd, III Edition, 2000.

- **Edited book having one editor:**

Raine, A. (Ed.). (2006). *Crime and schizophrenia: Causes and cures.* New York: Nova Science.

- **Edited book having more than one editor:**

Greenspan, E. L., & Rosenberg, M. (Eds.). (2009). *Martin's annual criminal code:Student edition 2010.* Aurora, ON: Canada Law Book.

- **Chapter in edited book having one editor:**

Bessley, M., & Wilson, P. (1984). Public policy and small firms in Britain. In Levicki, C. (Ed.), *Small Business Theory and Policy* (pp. 111–126). London: Croom Helm.

- **Chapter in edited book having more than one editor:**

Young, M. E., & Wasserman, E. A. (2005). Theories of learning. In K. Lamberts, & R. L. Goldstone (Eds.), *Handbook of cognition* (pp. 161-182). Thousand Oaks, CA: Sage.

- **Electronic sources should include the URL of the website at which they may be found, as shown:**

Sillick, T. J., & Schutte, N. S. (2006). Emotional intelligence and self-esteem mediate between perceived early parental love and adult happiness. *E-Journal of Applied Psychology*, *2*(2), 38-48. Retrieved from http://ojs.lib.swin.edu.au/index.php/ejap

- **Unpublished dissertation/ paper:**

Uddin, K. (2000). A Study of Corporate Governance in a Developing Country: A Case of Bangladesh (Unpublished Dissertation). Lingnan University, Hong Kong.

- **Article in newspaper:**

Yunus, M. (2005, March 23). Micro Credit and Poverty Alleviation in Bangladesh. *The Bangladesh Observer*, p. 9.

- **Article in magazine:**

Holloway, M. (2005, August 6). When extinct isn't. *Scientific American, 293*, 22-23.

- **Website of any institution:**

Central Bank of India (2005). *Income Recognition Norms Definition of NPA.* Retrieved August 10, 2005, from http://www.centralbankofindia.co.in/ home/index1.htm, viewed on

7. The submission implies that the work has not been published earlier elsewhere and is not under consideration to be published anywhere else if selected for publication in the journal of Indian Academicians and Researchers Association.

8. Decision of the Editorial Board regarding selection/rejection of the articles will be final.