Proceedings of National Level
Paper Contest for Ph.D. Students
on

Recent Trends in Computer & Information Science and Management & Engineering

Few Selected Papers

ISBN : 978-81-944670-1-4

21st November 2020

Editors
Dr. Krishna Prasad K.
Dr. P.S. Aithal
Dr. A. Jayanthiladevi

Srinivas Publication
(A Unit of Srinivas University)
Mangalore-575001, Karnataka, India
Email: srinivaspublication@srinivasgroup.com
Website: www.srinivaspublication.com

National Level Paper Contest for Ph.D. Students

on

Recent Trends in Computer & Information Science and Management & Engineering

ORGANISING TEAM

Chief Patron

Sri. CA. A. Raghavendra Rao
Chancellor, Srinivas University
Mangalore

Patrons

Dr. A. Srinivas Rao
Pro-Chancellor

Dr. P. S. Aithal
Vice-Chancellor

Prof. Shridhara Acharya
Event Director

Dr. Krishna Prasad K.
Event Convenor

Dr. A. Jayanthiladevi
Event Advisor

COLLEGE OF COMPUTER SCIENCE & INFORMATION SCIENCE
City Campus, Pandeshwar, Mangaluru– 575 001
Ph. No.: 0824 – 2441022, 2422851
Email: deanccis@www.srinivasuniversity.edu.in
Website: www.srinivasuniversity.edu.in

ACKNOWLEDGEMENT

After review, a few selected papers are published in this Proceeding Book. Special thanks to IEEE Bengaluru Section and IEEE Mangaluru Subsection for their Financial and Logistical assistance in organizing the event and publishing the Book.

ISBN: 978-81-944670-1-4

CONTENTS

A Conceptual Cloud-Based Framework for Fitness Tracking and Performance Prediction of Swimmers

Geetha Poornima K
Research Scholar College of Computer Science and Information Science, Srinivas University, Mangalore, India
Orcid ID: 0000-0001-9095-0349; E-mail: poornima.sanjay@gmail.com

***Abstract*—Swimming is considered a perfect exercise because the swimmer gets all the benefits of an aerobic workout without any negative impact on the muscles and joints. It can be practiced by people of all age groups. Like other competitive sports, swimming is very much influenced by the application of emerging technologies. To improve or sustain the performance a swimmer has to maintain physical fitness. Ensuring proper fitness parameters is a complicated job. Trainers and performance analysts aim to make use of emerging technologies such as Machine Learning (ML), Artificial Intelligence (AI), Internet of Things (IoT) and to evaluate different motor ability skills such as speed, power, endurance, coordination, flexibility, etc. to evaluate the performance. Performance prediction aims to analyze hidden relationships among different attributes to enhance the performance of swimmers. This paper sheds light on the different aspects of performance prediction and suggests a cloud-based conceptual framework that can be used to track the fitness of swimmers and predict their performance. A conceptual model is designed that uses a cloud computing platform and emerging technologies to track the fitness of the swimmers and assess their performance based on the fitness information.**

Keywords: Fitness, performance, wearables, prediction, parameters

I. INTRODUCTION

Data is in extensive demand these days. Every activity of an individual generates data that is of some significance for someone. The application of technology enables the efficient use of data. Technology has found its use in every sector. Sports are not an exception to that. Swimming is considered the complete physical activity that can be practiced by people of all ages including infants, elderly people, injured, overweighted individuals, and pregnant women. Fitness plays an important role in the performance of athletes. Different athletes show a different level of maturity and skillsets during the initial stages of training. Athletes are to be trained properly using game-specific variables to acquire fitness and improve performance. This can be achieved by using two models. They are generic training model and sports-specific training model. The generic model does not require any technical skills for the sports event. In the case of event-specific training, the swimmers are trained to adopt skills and tactics that are essential for their event. A swimmer must maintain fitness throughout the career. Training/coaching programs will help the athletes to acquire the skills that are essential to improve their performance. Fitness should be monitored strictly and carefully throughout the career of an athlete [1]. The goal of competitive swimming is to travel the race distance as swiftly as possible. The performance evaluation depends upon parameters such as velocity, stroke frequency, angle of hand movement, head positioning, etc. need to be analyzed. The influx of data has enormously affected the world of sports and games. The mighty wave of data has changed the nutrition and physique of athletes. With the help of adequate data, it is possible to analyze everything that the athletes do. Decision-making in the field of sports and games is also driven by data analytics. IoT devices, Global Positioning System (GPS) trackers, smart cameras, etc. monitor the on-field performance of the athletes, ML, AI, and predictive analytics (PA) are trying to find the next champion or superstar. The professional swimming field has embraced statistics and data in such a way that everything in sports is data-driven now. Swimmers and their coaches will have continuous pressure to improve performance. This has created a huge scope for PA in the field of sports to obtain accurate and real-time insights. To give a data-edge for sports these days teams appoint data analytics experts to improve the results [2]. Cloud computing provides a scalable storage platform for institutions and individuals. The service is provided transparently on a pay-as-you-use basis. When data is stored in the cloud, it can be accessed 'anytime' from 'anywhere' [3]. The performance prediction framework needs to store a large amount of data that can be done efficiently with the help of a cloud platform. Wearable devices or smart objects are considered as the building blocks of the intelligent sports framework. Daily workouts contribute more to the physical and emotional well-being of athletes. Wearable devices for sports offer several benefits to athletes and fitness freaks. These devices are economic and have changed the functionality of sports. Smart devices can be interconnected to automate the data collection process. When interconnecting multiple devices to automate the process of data collection. When data from these devices have used the protocols must ensure secure transmission of data. The smart devices can be used to monitor on a 24 X 7 basis to obtain deeper insights about the complete activities. The devices come with multiple sensors and applications to track blood pressure, heart rate calories burnt, distance covered, exercises carried out, sleep patterns, etc. Some applications measure emotions also [4]

II. OBJECTIVES

This paper focuses on the issues related to fitness tracking and performance prediction of swimmers. The main objectives include

- To understand the role of different fitness parameters in predicting the performance of swimmers.
- To elucidate the rote of different emerging technologies in fitness tracking and performance prediction.

III. METHODOLOGY USED

This study is carried out to analyze the relationship between different parameters in predicting the performance of swimmers. The work is conducted by using the information available online and from a few peer-reviewed

journals. The conceptual model designed in this paper is based on theoretical perception. It incorporates different technologies that are needed to track the fitness of swimmers and predict their performance. The concept is in its infantile stage and uses only abstract ideas. The feasibility of the development of this model requires further study in the related technologies.

IV. MOTIVATION

Participation in sports and games is compulsory during elementary schools; only selected students to engage in sports and games when it comes to the college level. Compared to the overall student strength of the college, only 5-10 percent of students participate in sports and less than 1 percent of athletes participate in swimming at the college level. Participation in swimming requires strong determination. In recent years, the success rate of college athletes is not up to the mark. The athletes lose consistency in performance due to unexplained circumstances. There is a need for an integrated fitness monitoring mechanism that allows physical directors and sports staff to keep track of fitness and predict its success. The framework should also provide athletes with customized training based on their fitness and skills.

V. RELATED STUDY

To maintain a steady performance or to improve the performance, an athlete has to maintain fitness. This requires regular exercise, following a strict diet, and regular practicing. Excess exercise may lead to overtraining syndrome which may lead to a decline in the performance, weakness during a performance, frequent sickness, etc. [5]. Coaches and athletes will also discover abrupt changes during learning Statistical methods can be programmed with simple algorithms and easily scaled to identify problem data that might help automate the coach's potential to obtain insight into the sudden decrease in the performance [6]. Weight training requires athletes to lift weights. When lifting weights, excess muscle power is needed. It is also considered as preventive solutions to back pain. It makes athletes lose their excess fat and strengthens their cardiovascular system. AI-based weight training techniques are used to provide rapid evaluation of parameters. Sensors are implanted on devices such as bench-press machines, weight training machines, leg press machines, etc. These sensors capture the necessary inputs when athletes use the respective machines. The inputs are then sent automatically to the evaluating system which is a computer-based framework. The parameters gathered are evaluated using artificial neural networks [7]. Maximum oxygen uptake is an important parameter to measure the endurance of athletes. It is also used to predict the risks of diseases in athletes. It serves as a metric to calculate disease risk, amount of cholesterol, excess body fat, and blood pressure. It provides a most accurate assessment of different parameters but requires sophisticated equipment, state-of-the-art laboratory settings, and trained staff which are expensive. Instead of establishing a laboratory ML and statistical techniques can be used to predict maximum oxygen uptake. Compared to the setting up of the laboratory, this technique is found to be less expensive and accurate. The data-driven technique can be applied to different athletes such as swimmers, hockey players, football players, and so on. Several parameters such as age, gender, height, weight, body mass index, exercise-specific information, etc. can be analyzed effectively to provide the desired outcome [8]. Lactate threshold (LT) is considered an important attribute in measuring the performance and fitness of athletes. During intense exercise involving exhaustive bouts swift energy is needed for athletes. LT is the intensity of exercise at which the accumulation of blood lactate starts to increase. The measurement of blood lactate demands expensive equipment. Blood samples are to be tested periodically to measure LT. This process is an expensive and non-practicable one. ML techniques evaluate the necessary parameters and check LT without the need for extraction of blood samples. This method is found to be an accurate and cost-effective one [9]. The performance of swimmers is strongly connected to their techniques. If the performance needs to be maximized their techniques are to be improved. The attributes such as body balance, body rotation, stroke efficiency, etc. need to be improved to increase the performance [10].

VI. CONCEPTUAL FRAMEWORK

Coaching is a continuous process that aims to transform novice swimmers into elite ones. It is possible to incorporate technology at every step of the coaching process to improve the performance of swimmers. Fig-1 shows different technologies that can be incorporated at different steps of the coaching process.

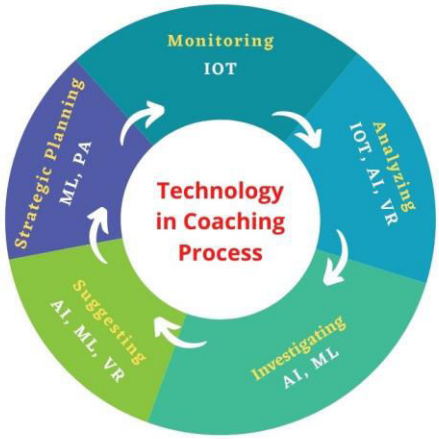


Fig. 1. Technology in the Coaching Process

1. **Monitoring:** During the period of coaching, the coaches are required to monitor the dryland exercises and other parameters such as stroke, power, speed, strength, agility, endurance, flexibility, etc. IoT enabled wearable devices can provide the much-needed quantitative information related to these parameters accurately.
2. **Analyzing**: Once the required parameters are collected, they are to be analyzed to find out the positive and negative attributes. Positive attributes are to be improved and the negative ones are to be minimized. Stroke type, number of breaths per length, time underwater, number of strokes, etc. need to be analyzed to assess the performance. The necessary parameters can be collected from wearables and AI algorithms can be used to provide valuable insights to coaches. VR can be used to analyze different moves of swimmers.

3. **Investigation**: In this step coaches critically evaluate the parameters such as the position of the head during swimming, angle of hand movement, and other aspects such as oxygen uptake, sustainability, Lactate tolerance (LT), the velocity of swimming, etc. that influence the performance. The statistical data collected during match or training, video recordings, etc. can be analyzed using AI algorithms and ML techniques to provide an accurate assessment of the parameters.
4. **Suggesting**: Once the essential parameters are analyzed ML, VR, PA, and AI can be used to provide suggestions for improvement of performance. Different types of predictions such as stress-level prediction, injury prediction, etc. can be done at this step.
5. **Strategic planning**: The ultimate aim of competitive coaching is to maximize the chance of winning the competition. It requires strategic planning. PA, VR, and ML can be used by coaches to provide a winning game-plan to the coaches [11].

Competitive swimming requires to find out the hidden relationship among different parameters. Fig-2. shows a conceptual model that can be used to predict the performance of swimmers based on their fitness.

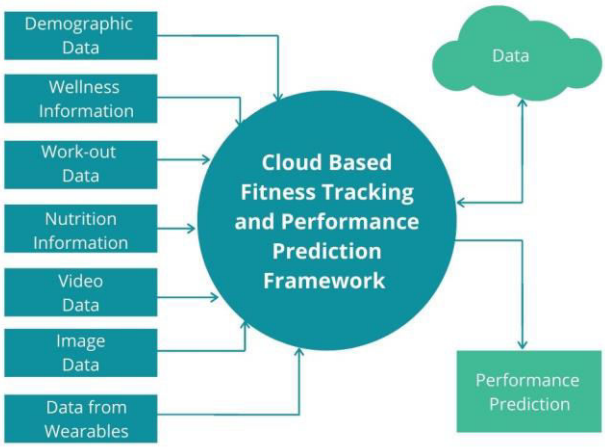


Fig. 2. Conceptual Model

The model displayed in Fig-2 requires a variety of inputs from distinct sources. Demographic data include height, weight, gender, year of birth, foot length, arm span, etc. It can be collected by providing a form to the swimmers during the admission process. Wellness information is related to the physical, mental, and emotional wellness, resting hours, etc. of swimmers. This needs to be collected every day before the training process. Swimmers are asked to enter this data honestly. Nutrition information is related to the diet the swimmer needs to follow. It includes the number of calories in the food to be taken and a complete plan containing the carbohydrates, proteins, sugar, fat, etc. that need to be followed during the training session. It also includes hydration details. Based on the workout target and food habits the nutrition information is swimmer-specific. Workout information contains the type of dryland workout to be followed to strengthen the muscles, increase endurance, etc. Fitness trials include both aerobic and anaerobic exercises. The entire fitness workout involves pull-ups, Vertical jumps, planks, V-ups, Clamp shells, superman hold, etc. These exercises are to be carried out to strengthen the upper body, core, and lower body. The qualitative information related to these exercises also varies from person to person. It is the job of the coach to design a proper training and workout plan to minimize exhaustion, overtraining, and injury risks. In addition to the fitness exercises, the swimmers are supposed to practice yoga, pranayama, and meditation. These practices are essential to create calmness, increase body awareness, release stress, relax the mind, and enhance concentration. The perfect fitness training is a combination of yoga, dryland workouts, games, stretching, animal flow, etc. For advanced analysis recorded videos of athletes are also used. They serve as a crucial input for the assessment of technics and tactics. For assessing the postures, hand-movement, angle of hand-rotation, etc. the pictures clicked during training sessions or during the competition will be useful. IoT devices such as goggles, accelerometers, smartwatches, smart bands, headbands, etc. can be easily worn during training sessions. They provide valuable information on quantitative data such as strokes, number of breathings, mobility, flexibility, calories burnt during workout and training, etc. The data from all these diverse resources will be in a different format. It will be stored effectively in the Cloud platform. Once all the necessary data is collected, the system can make use of emerging technologies such as AI, ML, VR, and PA to evaluate the parameters and provide actionable insights. The insights provided by the system can be used by the coaches to predict the performance of swimmers.

VII. QUANTITATIVE PARAMETERS USED

It is important to understand the relationship between different biomechanical variables and energetics related to swimmers. Researches show that the performance is strongly related to different energetic variables such as oxygen uptake, energy cost, velocity to achieve maximum uptake, LT, etc. These variables depend upon different biomechanical parameters such as the velocity of swimming, stroke frequency (SF), stroke length (SL), etc.

Each swimmer is made to swim a specified distance say 200 meters to measure different biomechanical parameters. The velocity is calculated as

$$v = \frac{d}{t} \quad (1)$$

according to (1) 'v' is the velocity of swimming; d is the distance covered by the swimmer and 't' is the time taken by the swimmer to cover the specified distance 'd'. To critically evaluate these parameters the distance 'd' say 200 meters is divided into 4 equal sessions say 50 meters each and the time taken to cover every 50 meters is recorded. This can provide valuable information related to the energetic parameter LT.

$$SL = \frac{v}{SF} \quad (2)$$

SL is the horizontal distance that the swimmer's body travels during the stroke cycle, SF is the number of strokes performed within a unit of time. (2) represents SL in terms of V and SF. An increase or decrease in 'v' is will directly affect SL.

$$SI = SL.\ v \quad (3)$$

The stroke index (SI) is another parameter that is used to assess the overall efficiency of swimming. It is calculated

 ISBN: 978-81-944670-1-4 | November, 2020

using the formula (3). To improve SI, v and SL are to be increased. It depends highly on the physical strength and endurance of the swimmer [12].

In addition to the parameters discussed here, there are some more variables related to limb kinematics, hip and center-mass kinematics, the friction of water and waves, neuromuscular variables, and genetic factors. When all the qualitative and quantitative parameters are analyzed accurate prediction of performance is possible [13].

VIII. TECHNOLOGICAL AID

A. Cloud Computing

Cloud is a centralized unit to be used to hold different kinds of data for analysis purposes. It becomes easy to analyze when data is stored in a centralized data warehouse. Cloud computing provides resources especially data storage and computing power to the user on an on-demand basis. It is the best example of resource sharing. It ensures agility, flexibility, and scalability for storing voluminous information needed to be stored for analysis. Cloud can be public, private, community, or hybrid. The public cloud can be accessed by anyone and it is available free of cost. A private cloud is owned by a single organization, a community cloud is meant to be used for a group of organizations, a hybrid cloud is a combination of the public and private cloud where non-critical data is stored on the public cloud and critical data is stored on the private cloud [3].

B. Predictive Analytics (PA)

PA is a branch of analysis that uses historical data and all the emerging technologies such as IoT, AI, ML, Deep Learning, etc. for decision making. It is used by industries to increase their business. It uses algorithms of data analytics and ML techniques to analyze data and determine the possibility of future events. It consists of a wide variety of techniques related to Data Mining, Predictive Modeling, Deep Learning, etc. Predictive models use historical data to identify possible risks and opportunities associated with a decision. The data required for PA come from a wide range of sources therefore it will not be a structured one. For efficient decision-making reasons, the data must be processed and analyzed. PA uses historical data to analyze and identify a solution to the problem then it uses real-time data to determine future events. Based on historic data available, a model that uses algorithms and techniques to analyze the situation is built. Then the model uses real-time data to predict future events exactly [14].

C. Artificial Intelligence (AI)

AI is a technology that is associated with the creation of computer systems that exhibit human-like behavior such as learning, reasoning, processing the natural language, etc. AI is extensively used in the analysis of sports-related data. It is used to detect anomalous data, to provide data visualization using neural networks and deep learning techniques. The natural language processing technique of AI is increasingly used by journalists to automate their sports coverage capabilities. AI-enabled chatbots are used by sports teams to manage fans' queries [15].

D. Machine Learning (ML)

ML uses statistical and probabilistic techniques that allow computers to learn from previous examples and to detect hard-to-discern patterns from huge noisy or complex data. Several ML algorithms are extensively used in predictive analytics. Depending on the nature of learning, they are classified into supervised, unsupervised, semi-supervised, and reinforcement learning techniques. Supervised learning is learning from examples. The supervised learning algorithm takes a known set of input data and known responses to the data and trains a model to generate realistic predictions as the response to new data. In the case of unsupervised learning, there is no training data. The insight is derived from the new test data. Semi-supervised learning uses both labeled and unlabeled training data. Reinforced learning is the one in which a machine is trained to take accurate decisions without human intervention [16].

E. Virtual Reality

It offers a dynamic assessment and will provide the most adaptive, real-world picturesque solution. It is used for simulating complex mathematical or computer modeling, organizing all data inputs and quickly and effectively producing outputs, and presenting data outputs in the form of graphs, reports, diagrams, and 2D or 3D visualization techniques. It can ensure that the outputs are more comprehensible, presentable, and connected and anyone without much previous knowledge and use [17].

F. IoT

IoT is an emerging technology with the help of which 'people' and 'things 'are connected 'anytime ', 'any place 'with 'anything' and 'anyone'. The new technology will make the day-to-day life of an individual easier, simpler, and better. One of the most popular and prevalent fields that use IoT in the sports sector. It aims to capture the quantitative data related to athletes. The different categories of smart things include on-body, off-body, and in-body devices. They enable automatic capture of data using sensors and the captured data can be stored on any storage medium without human interference [18].

G. Communication Technologies

Advancement in communication technology has increased the connectivity of different IoT devices. Different communication technologies include Low Power Wide Area Networks (LPWANs), 3G, 4G, or 5G technologies used in cellular networks, short-range wireless technologies such as ZigBee, Bluetooth, or Wi-Fi. Depending on the bandwidth required a suitable communication technology has to be selected to transfer data from wearables and other sources to the cloud [19].

IX. BENEFITS AND CHALLENGES

A. Benefits

Availability of up-to-date data: With the help of IoT devices, it is possible to gather real-time quantitative data related to the dryland workout sessions as well as the data related to training sessions easily. No need to key the data into the system manually as the up-to-date data of the necessary parameters is readily available.

Feedback capabilities: The advancement in wearable technology lead to the development of devices that are capable of providing accurate feedback related to

performance. This feedback is used by swimmers and coaches to enhance the performance.

Personalized training: Different swimmers show different levels of maturity, fitness, and skillsets. The use of technology can facilitate coaches to provide personalized training by evaluating their motor ability skills, techniques, and tactics.

Pervasive monitoring: With the help of wearable IoT devices the coaches can monitor the swimmers anytime anywhere. It is possible to make the devices send data related to different physical activities periodically without human intervention. This enables coaches to perform ubiquitous monitoring of swimmers.

Accurate prediction: With the help of predictive models and algorithms it possible to predict the performance of swimmers accurately. In most cases, data are automatically sent by wearable devices. Hence there is less room for manual errors.

Improved decision making: Predictive algorithms can generate accurate decisions to improve the performance of swimmers. The system helps the coaches a lot in effective decision-making.

Cost-effectiveness: Traditionally, performance prediction requires to employ many video cameras that produce high-quality videos. This requires high investment for devices and software. With the help of wearable IoT devices, it is possible to collect accurate data. The wearable devices are less expensive when compared to high-quality video cameras.

Increased Scalability: Any number of IoT devices can be interconnected to track different fitness-related parameters of the swimmers. They generate a huge amount of data with can be effectively stored on a cloud platform. Hence there is no limit to the amount of data as well as the number of devices that are used to monitor fitness.

Enhanced Transparency: The use of technology in swimming will facilitate the availability of accurate data related to the performance of swimmers thereby making the system more transparent. This makes the swimmer selection process for a competition easy and unbiased [3][18][20].

B. Challenges

Security and privacy issues: The analyzing systems require a huge amount of personal data of individual swimmers. The more the data, the accurate will be the prediction. When an individual swimmer's personal information is store.

Accuracy depends on the model: Accuracy of performance prediction depends upon the predictive model used and the techniques and algorithms that are used to implement the problem. The selection of a model suitable for the problem is always a challenge in the case of ML.

Scarcity of experts: Prediction of performance requires knowledge of Pa and ML algorithms and techniques. This is an area that always suffers from the scarcity of experts.

Negative impact on the performance: When the swimmers access data related to their performance there is a possibility of developing stress or depression. Coaches are not supposed to share performance-related parameters with athletes as they may directly affect their performance.

Selection of parameters: Different type of swimming requires different parameters to be analyzed. Parameters required for the assessment of freestyle, breaststroke, backstroke, and butterfly style are different. The selection of parameters requires a clear-cut idea of swimming. It is a real challenge for the data analyst [3][16][20].

X. LIMITATIONS AND FUTURE SCOPE

The conceptual model designed in this paper is based on the analysis of several researchers in the related field. The findings of different researchers are incorporated and the model is in its conceptual level and its implementation issues are to be investigated further. The model only focuses on the prediction of the performance of swimmers. The same model can be used to predict injuries, assess stress, the possibility of overtraining, outlier detection, and predict depression, etc. When the possible events can be predicted well in advance it is possible to minimize the risks associated with the same. Different ML and AI techniques that are used to build a predictive model are to be investigated further. The conceptual model focuses only on biomechanical and energetic parameters. To develop a full-fledged performance prediction solution for swimmers' limb kinematics parameters, generic factors, neuromuscular parameters, psychological parameters, etc. need to be analyzed.

XI. CONCLUSION

Fitness plays a crucial role in the success of swimmers. During the early stages of preparation, different swimmers show different degrees of maturity and abilities. Fitness and practice are directly associated with the success of a swimmer. Deviation in anyone will jeopardize the performance. While coaches and swimmers strive hard to find an efficient method to improve the performance, technology is silently doing it by providing valuable data required for analysis. Wearables can be used as reliable sources of information to improve performance. Prediction of performance requires the study of different variables and the hidden relationship among them. Adaptation of technologies will help the coaches to assess, compare, and manipulate dependent variables to enhance the performance. The use of technologies can also help the coaches to define goals, design personalized training sessions to enhance the performance of individual swimmers.

REFERENCES

[1] Harrison, C. B., Gill, N. D., Kinugasa, T., & Kilding, A. E . "Development of Aerobic Fitness in Young Team Sport Athletes" Sports Medicine, 45(7), 969–983., 2015.

[2] Dunn, A. M., Hofmann, O. S., Waters, B., & Witchel, E. "Cloaking malware with the trusted platform module" In Proceedings of the 20th USENIX Security Symposium (pp. 395–410), 2011.

[3] Cai, H., Xu, B., Jiang, L., & Vasilakos, A. V "IoT-Based Big Data Storage Systems in Cloud Computing: Perspectives and Challenges." IEEE Internet of Things Journal, 4(1), 75–87, 2017.

[4] Islam, S. M. R., Kwak, D., & Kabir, H.. The Internet of Things for Health Care: A Comprehensive Survey. 3, 2015.

[5] Mackinnon, L. T. "Overtraining effects on immunity and performance in athletes" . Immunology and Cell Biology, 78(5), 502–509, 2000.

[6] Silva, P., Lott, R., Wickrama, K. a S., Mota, J., & Welk, G. "Modern Techniques and Technologies Applied to Training and Performance

Monitoring" *International* Journal of Sport Nutrition and Exercise Metabolism, 32, 1–44, 2011.

[7] Novatchkov, H., & Baca, A. "Artificial intelligence in sports on the example of weight training" *Journal of Sports Science and Medicine*, *12*(1), 27–37, 2013.

[8] Press, D. "Machine learning and statistical methods for the prediction of maximal oxygen uptake": recent advances. 369–379.

[9] Etxegarai, U., Portillo, E., Irazusta, J., Arriandiaga, A., & Cabanes, I. " Estimation of lactate threshold with machine learning techniques in recreational runners." Applied Soft Computing Journal, 63, 181–196, 2018.

[10] Bächlin, M., & Tröster, G. "Swimming performance and technique evaluation with wearable acceleration sensors" Pervasive and Mobile Computing, 8(1), 68–81, 2012.

[11] Jürimäe, J., Haljaste, K., Cicchella, A., Lätt, E., Purge, P., Leppik, A., & Jürimäe, T. "Analysis of swimming performance from physical, physiological, and biomechanical parameters in young swimmers" *Pediatric Exercise Science*, *19*(1), 70–81., 2007.

[12] M., T., A., D., J., M., & J., A. " Biomechanics of Competitive Swimming Strokes". Biomechanics in Applications, i., 2011.

[13] Barbosa, T. M., Costa, M., Marinho, D. A., Coelho, J., Moreira, M., & Silva, A. J. "Modeling the links between young swimmers' performance: Energetic and biomechanic profiles" . Pediatric Exercise Science, 22(3), 379–391, 2010

[14] Ahmed, E., Yaqoob, I., Hashem, I. A. T., Khan, I., Ahmed, A. I. A., Imran, M., & Vasilakos, A. V. The role of big data analytics in Internet of Things. *Computer Networks*, *129*, 459–471, 2017.

[15] Davenport, T., & Kalakota, R. DIGITAL TECHNOLOGY The potential for artificial intelligence in healthcare. Future Healthcare Journal, 6(2), 94–102, 2019.

[16] Zhu, P., & Sun, F. Sports Athletes' Performance Prediction Model Based on Machine Learning Algorithm. In Advances in Intelligent Systems and Computing (Vol. 1017). Springer International Publishing, 2020.

[17] Tanaka, Y., & Hirakawa, M . Efficient strength training in a virtual world. *2016 IEEE International Conference on Consumer Electronics-Taiwan, ICCE-TW 2016*, 6–7. 2016.

[18] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. "Internet of Things (IoT): A vision, architectural elements, and future directions." Future Generation Computer Systems, 29(7), 1645–1660, 2013

[19] Aqeel-ur-Rehman, Mehmood, K., & Baksh, A. "Communication Technology That Suits IoT - A Critical Review". *Communications in Computer and Information Science*, *366 CCIS*, 14–25 (2013)

[20] Allen, S. V, Vandenbogaerde, T. J., Pyne, D. B., & Hopkins, W. G.. "Predicting a Nation ' s Olympic-Qualifying Swimmers." 431–435. 2015.

The brief analysis on big data analytic technologies in agriculture sector

Vikranth K

Research Scholar, College of Computer Science and Information science, Srinivas University, Mangaluru, Karnataka, India
Assistant Professor, Department of Computer Science, Vivekananda College, Puttur, India
Email: vikranth.kadya@gmail.com
Orcid ID: 0000-0001-9549-1743

ABSTRACT

Agriculture and its practices are traditional and not reliable to the current competitive scenario to get full-fledged satisfaction and outcome of the business. In a country like India, people get migrated from agriculture to some other business because of a lack of technological innovation, a way of doing agriculture, a very small profit margin and no future information. In recent days had some research work to improve agriculture practice by using current and innovative technology in computer science and information science. The technological improvement in agriculture practice is referred to as smart agriculture or digital agriculture or precision agriculture. The main goal of all modern agricultural practice is to the effective utilization of existing technologies to perform drastically changes in way of doing agriculture thereby making agriculture as one of the comfortable and to make farmers to get more satisfaction in their profession. The different technologies such as IoT, Wireless Sensor Network, Data analytics has more contribution and requires a lot of research to further improvement. This article explains the effect of big data and data analytics in agriculture and explains different data analysis techniques used in different areas of agriculture. It also proposed the typical model of smart agriculture system that incorporates precision agriculture using data analytics.

Keywords: Data analytics, Agriculture, Big data, prediction, Decision support, Analytical techniques.

I.INTRODUCTION:

Agriculture is the backbone of the Indian economic system. The world population grows continuously in exponent form, and resources for crop production and output of agricultural activity is diminished. Therefore there have to be some innovative and reliable approaches that boost up agricultural activities. Agriculture is one of the oldest human activities that need to transfer into smart activity by using big data and related technology. The advancement in data analytical technology in data- driven agriculture solves today's many global problems. Data analytic techniques can be used in the accumulated bulk data to obtain information that is used by farmers to take appropriate decisions throughout the farming activities like planning, plantation, harvesting and marketing. The term data analytics refers to the process of examining large data sets to gain knowledge for proper decision making. It enables us to take raw data and uncover patterns to mine treasured insights from it [1]. The innovation of new technologies like the Internet of Things (IoT), Wireless sensor network (WNS) and data analytics shapes agricultural activities smart and smooth. In the future also agriculture will be expected use sophisticated technologies like robotic technology, sensors, aerial images and GPS technology. These advanced technologies allow farmers to peruse more profitable, efficient, and safe and nature-friendly agriculture. These technologies will bring several advantages to farmer societies such as empowers small farmers, manage crop diseases and pests, deals with environmental changes, make crop profit predictions etc. Predictive analytics gives actionable insights on available data and information that improves the on-time situational decision-making process during agricultural activity. The precision agriculture concept is relies on predictive analytics is depends on analyzing real-time data that comes from various sources of the farm such as soil, temperature, air, water, insecticides, moisture, equipment etc. The smart farming concept is mainly focused on usage of data analytics and its outcome in various functions of agriculture. Therefore these data analytic tools and technologies enable the farmers to timely information about agriculture thereby can take better decisions and policymakers to make better decisions on buying and selling agricultural products [2].

The remaining portion of the paper is systematized as following manner. In section 2 the effect of big data in agricultural practice is discussed. In section 3 and 4 the different big data analytical techniques and comparisons of different big data analytics were elaborated and section 5 narrates the implementation part of big data analytics in agriculture.

II. EFFECTS OF BIG DATA IN AGRICULTURAL PRACTICE:

The big data and data analytic solution has leveraged the data science in agriculture to automate and visualize the problems to the farmers. The precision farming or digital farming or smart farming generates the tremendous amount of data which stored in backend system can be easily analyzed with result will be stored in the front end helps farmer to undertake the appropriate decisions. The configured customized dash board can track all dynamic situations in farm and provide notifications about changes takes place in the farm. The data generated in different sources and equipment's are stored and updated automatically in the system [3]. These stored data or real time data generated by sensors affects the agricultural practice in following way.

A. Big data for weather prediction:

Practically the agriculture is mainly depends on natural conditions such as climate, soil, pests and weather. With use of big data and other related technologies farmers can easily analyze the extreme weather conditions that impact the agriculture. The big data platform will provide information on real time condition helps farmer to respond quickly. Also data from sensors and drone camera helps farmers to know expected growth rates [4].

B. Supply Chain Tracking:

The every stages of supply chain tracking are covered by smart farming. Big data are very useful for all the stake holders of supply chain management. At the time of production it was used by automated system and to reveal the conditions of production equipment. It helps farmers and suppliers to manage marketing activities which increases delivery reliability [5].

C. Risk Assessment:

Big data for risk management helps farmers to have prediction and manage risk connected with growing crops. The big data powered black chain platform is popular risk management for agriculture [6].

III. AGRICULTURE AND BIG DATA ANALYTICAL TECHNIQUES :

Smart farming and precision agriculture are 2 main areas found during the study of big data analytics in agriculture. The precision agriculture is the ideas of farming that include a concept called measurement, Observation and response [7]. Smart farming explains the connection of functions, variables and concepts. This section mainly focuses on well-known technologies related to big data in agriculture and its practical demonstration [8].

A. Predictive analytics:

It is an approach related to big data analytics can forecast future outcomes by analyzing current and historical data present in the database. Majorly it resolves with important techniques like regression analysis, decision trees and neural network. Further, it is a technique based on statistical data and process used in agriculture to predict crop, yield and consumer behavior [9]. The decision tree technique is influenced by a classification algorithm that forecasts the risk and rewards of various actions and further determines possible outcomes in the form of clean visualization methods of flow chart, histogram, pie chart, graph and other pictorial formats. The statistical modeling can be explained and used to an intensive deep learning model [10]. The multiple regression models are extensively used model in predictive analysis. This model is for predicting future trends and forecasting changes of one variable with respect to another variable.

B. Recommendations system:

It gives output in the form of advice based on the behavior and functionality of data and patterns. In agriculture recommendation system uses big data techniques such as collaborative filtering, content-based filtering and hybrid to analyze the data and recommends farmers based on climate, territory conditions or purchase of agricultural commodities [11]. The content- based technique depends on comparisons and similarities to make a recommendation system. The clustering and neural network methods are used to calculate utility prediction with the data. In a collaborative filtering technique which collects the data

from user in the form of feedback and converts them as rating in any given area [12]. It exploits some common behavior from rating among various users. The common methods used here are K-nearest neighboring (Clustering) using matrix factorization and neural nets. The Hybrid technique is a combination of content based and collaborative techniques. It solves the crunches of both models and provides better output for the recommendation system [13].

C. Data Mining:

In agriculture, data mining has major impact and contains several techniques. The pattern mining technique is used to find a pattern from a large data set. The association technique is the oldest one is also known as the relation technique [14]. Here the pattern is discovered based on relation between the items of the same transaction or activity. The classification is a typical mining technique based on a machine learning model. It uses mathematical context class techniques such as statistics, linear regression, decision tree etc. the clustering is one more data mining technique based on machine learning. Here clusters or groups of homogeneous objects are formed in the form of clusters. These objects are then put under a group of classes. The prediction data mining technique ascertains the association among dependent and independent variables in a given area. It is mainly used in text mining and sentiment analysis. The Sequential pattern mining is an approach in data mining that recognizes related kind of patterns, developments or trades in any given data set [15].

D. Spike and Slab regression analytic technique:

The two main coefficients used in probability distributions of linear regression models are spike and slab. Also it is one of Bayesian technique used in linear regression model that gains popularity it agriculture. [17].

E. Time series analytic technique:

The formerly used values are used to forecast the output. It uses the model of time series has a series of data points arranged in time [14]. Here time is considered an independent variable that has the ambition to forecast the future. It is used to forecast crop developments, environmental changes, price movement, etc. [18].

IV. ANALYTICAL TECHNIQUES IN AGRICULTURAL AREA :

The different analytic technique was implemented based on purpose and area of agriculture is shown in below table 1.

Agricultural Area	**Sources**	**Analytical Techniques**
Weather data	Sensors, Historical information, Weather station	Statistical analysis, Machine Learning (K-Means, random forest), Map Reduce
beast	Earth sensor, Optical sensor, feed, milk	Neural network, scalable vector machine, decision tree
Crops	Historical data, Satellite data, Ground sensor, Government data	K-means clustering, Fourier transform, scalable vector machine, wavelet filtering
Land	Chronological data, Geo-Special data, remote sensing data	Random forest, K-Means, Image processing
Weeds	Historical data, Drone images, Field sensor	Logistic regression, Neural Network, Image processing
Soil	Ground sensors, Historical data base, Optical data	Neural Networks, K-Means, Clustering algorithms.
Bio diversity	Historical data, Geo-special data, Government data	Statistical modeling
Food Security	Remote sensors, geo-special data, Historical data, Survey data	Simulation, neural network, statistical modeling, image

		processing
Farmer	Government data, social media, Historical data, optical sensors	Web services, web applications, mobile applications.
Distant Sensing	Drone data, satellite data, web based data.	Cloud computing using map reduce, decision support system, Geo special analysis.

Table.1: Big data analytic techniques in agriculture

V.IMPLEMENTING BIG ANALYTIC TECHNIQUES IN AGRICULTURE:

The proposed system collects the data from the environment, land, yield and other places which was transmitted over the network and stored in a data mart [19]. According to requirement data mart creates cluster of data. The clustered data will be analyzed with proper data analytic technique which will help to give prediction results to the farmer [20]. The decision support system will analyze the prediction result and helps to increase crop production as shown in figure 1.

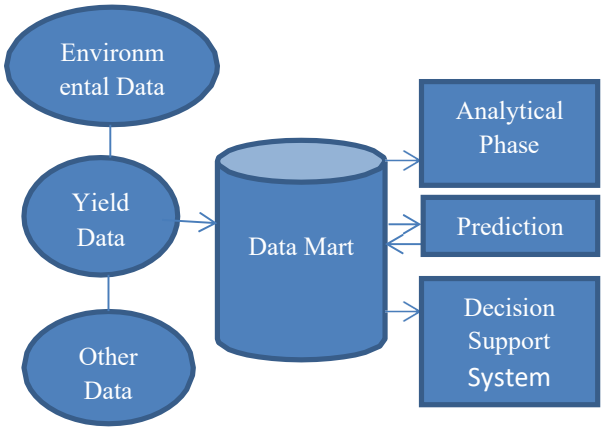


Fig.1 . Implementation of data analytics in agriculture.

VI.CONCLUSION:

Agriculture and its practices need to be shifted from its traditional method to the modern environment by using innovative technologies available today. These transformations should improve farmer's economy, lifestyle, way of thinking and entire business scenario related to agriculture. This paper narrates the big data and related technologies effects agriculture scenario and shows a brief comparison of different analytical technologies in the way it applied to different sectors of agriculture. Nevertheless, the increase in the use of big data and related technologies has huge potential in smart agriculture technology will boost to carry out more research and innovations. It will motive farmers to adapt in to modern technology to achieve a higher quality product, generate more revenue, and plan for the future to achieve sustainability in agri-business.

REFERENCES

[1] Tseng, F. H., Cho, H. H., & Wu, H. T. (2019). Applying big data for intelligent agriculture-based crop selection analysis. IEEE Access, 7, 116965-116974.

[2] Cardenas-Rodriguez, M., & Ferruzola-Gómez, E. (2019, November). A Brief Review of Big Data in the Agriculture Domain. In Technologies and Innovation: 5th International Conference, CITI 2019, Guayaquil, Ecuador, December 2–5, 2019, Proceedings (Vol. 1124, p. 66). Springer Nature.

[3] Kumar, M., & Nagar, M. (2017, July). Big data analytics in agriculture and distribution channel. In 2017 International Conference on Computing Methodologies and Communication (ICCMC) (pp. 384-387). IEEE.

[4] Li, C., & Niu, B. (2020). Design of smart agriculture based on big data and Internet of things. International Journal of Distributed Sensor Networks, 16(5), 1550147720917065.

[5] Pham, X., & Stack, M. (2018). How data analytics is transforming agriculture. *Business Horizons*, *61*(1), 125-133.

[6] Rajeswari, S., Suthendran, K., & Rajakumar, K. (2017, June). A smart agricultural model by integrating IoT, mobile and cloud-based big data analytics. In *2017 international conference on intelligent computing and control (I2C2)* (pp. 1-5). IEEE.

[7] Srichandan, P., Mishra, A. K., & Singh, H. Data Science and Analytic Technology in Agriculture. International Journal of Computer Applications, 975, 8887.

[8] Nuvvula, J., Adiraju, S., Mubin, S., Shahana, B., & Valisetty, V. (2017). Environmental Smart Agriculture Monitoring System Using Internet of Things. *International Journal of Pure and Applied Mathematics*, *115*(6), 313-320.

[9] Srija, V., & Krishna, P. B. M. (2015). IMPLEMENTATION OF AGRICULTURAL AUTOMATION SYSTEM USING WEB & GSM TECHNOLOGIES. International Journal of Engineering & Science Research, 5(9), 1201.

[10] Jadhav, G., Jadhav, K., & Nadlamani, K. (2016). Environment monitoring system using Raspberry-Pi. Int. Res. J. Eng. Technol.(IRJET), 3(04), 1168-1172.

[11] Raj, M. P., Swaminarayan, P. R., Saini, J. R., & Parmar, D. K. (2015). Applications of pattern recognition algorithms in agriculture: a review. *International Journal of Advanced Networking and Applications*, *6*(5), 2495.

[12] Patel, R., & Patel, M. (2013). Application of cloud computing in agricultural development of rural India. International Journal of Computer Science and Information Technologies, 4(6), 922-926.

[13] Mohapatra, S., Srichandan, P., Mohanty, S., Singh, H., & Patra, P. K. (2018, December). Smart Agriculture: An Approach for Agriculture Management using

Recent ICT. In 2018 International Conference on Information Technology (ICIT) (pp. 187-192). IEEE.

[14] Wickramasinghe, C. P., Lakshitha, P. L. N., Hemapriya, H. P. H. S., Jayakody, A., & Ranasinghe, P. G. N. S. (2019, December). Smart Crop and Fertilizer Prediction System. In 2019 International Conference on Advancements in Computing (ICAC) (pp. 487-492). IEEE.

Agrawal, S. (2018). Online Tool for Weed Detection and Prediction in Crops. Journal of Computational and Theoretical Nanoscience, 15(6-7), 2448-2453.

Sambrekar, K., Rajpurohit, V. S., & Deyannavar, S. B. (2018) . Design of a Cloud-Based Framework (Saa S) for Providing Updates on Agricultural Products. In International Proceedings on Advances in Soft Computing, Intelligent Systems and Applications (pp. 115 - 122). Springer, Singapore.

Paikekari, A., Ghule, V., Meshram, R., & Raskar, V. B. (2016). Weed detection using image processing. International Research Journal of Engineering and Technology (IRJET), 3(3), 1220-1222.

Goraya, M. S., & Kaur, H. (2015). Cloud computing in agriculture. HCTL Open International Journal of Technology Innovations and Research (IJTIR), 16, 2321-1814.

Nath, S., Debnath, D., Sarkar, P., & Biswas, A. (2019). Design of Intelligent System in Agriculture using Data Mining. International Journal of Computational Intelligence & IoT, 2(3).

Park, H., Lee, E. J., Park, D. H., Eun, J. S., & Kim, S. H. (2016, October). PaaS offering for the big data analysis of each individual APC. In 2016 International Conference on Information and Communication Technology Convergence (ICTC) (pp. 30-32). IEEE.

A review of virtual work environment technologies in the wake of covid-19

Yogish Pai U
College of Computer Science and Information Science
Srinivas University
Mangalore, India
yogish77pai@gmail.com

Abstract— Covid-19 situation forced majority of the IT industries to shift their workforce to operate from virtual work environment by enabling work from home or work from anywhere concept. Teleworking is made easier by a combination of technology advances and the adaptability of the industries. Cloud computing platform, Desktop as a Service (Daas), Software as a Service (SaaS) and VPN as a service (VPNaaS) etc. helped the industry transition to a virtual work environment. Infrastructure as a Service (IaaS) and Platform as a Service (PaaS) also played a key role in enabling virtual work environment. In this paper, we will discuss these technologies and tools that have been placed in place to facilitate changes in the workplace during a pandemic. We are also going to discuss the security challenges and available solutions during the teleworking model. Some recommendations are also provided to overcome certain practical challenges.

Keywords—DaaS, SaaS, Teleworking, VPNaaS, Cloud

I. INTRODUCTION

Working model of majority of the IT firms changed drastically during Covid-19 Pandemic. Lockdown situation forced companies to switch to work from home model as it was the only survival option available at that point. Work from home or teleworking was not a new concept to IT industry as most of the companies had certain number of employees working under this model.[1] The challenge faced during pandemic was with the volume of employees to be shifted to teleworking model. Most of the companies were not prepared to enable teleworking for entire workforce. The companies operating 24x7 had shortage of computers as most of the computers were shared by more than one employee.

Another biggest challenge IT firms had to address was identification and implementation of right technology which supports smooth functioning of remote working. Not all the IT companies adopted cloud technology for their day to day operations. Many companies were still dependent on applications and tools hosted in on-premise servers. Some of the companies were using legacy applications which were not supported by latest cloud technology. There was a need for all the employees to get connected to the company network for several reasons, such as process related tools were accessible only over trusted office network, necessary documents and CRM hosted inside the office premises, patch management and group policy sync were possible only via office network. Thanks to modern day firewalls and Unified Threat Management devices which ships with built-in VPN modules, helped organizations to enable VPN access to office network without much configurations hassles. Microsoft and Amazon have played a major role in accelerating the digital transformation of IT companies by offering wide variety of cloud-based tools such as Microsoft Windows Virtual Desktop Solution, Amazon AWS Remote Workspaces etc.

II. VIRTUAL WORK ENVIRONMENT TECHNOLOGIES AND TOOLS

Fulfilment of virtual work environment purely depends on the nature of work or project requirement. Certain projects which are related to healthcare, Finance, personal data processing etc., may demand more security. In order to meet regulatory, statutory and business need company will have to carefully design its virtual work environment. In this paper we are going to review some of the technologies and tools that helped IT companies to build their virtual work environment effectively.

A. Virtual Private Network (VPN)

VPN helps organizations to create a secure channel between employees, working from home and office location, so that data located inside the office network can be accessed securely. Based on the security requirement company can choose SSL VPN or IPSec VPN. Convenience and Security plays a major role while deciding the VPN technology. VPN module is a built-in feature in most of the Unified Threat Management devices and Next Generation Firewalls available in the market. Some of the alternate options are to deploy dedicated hardware VPN Gateways, setup Software VPN Gateways within an office network or somewhere in the cloud.

VPN as a Service is another great savior for IT companies during pandemic situation. Large workforce can be routed via VPN concentrator hosted on cloud. VPNaaS is capable of providing robust and secure connectivity to its subscribers.[2]

B. Desktop as a Service (DaaS)

Desktop as a Service (DaaS): Virtual Desktop Infrastructure (VDI) Technology is not new to IT industry. DaaS is nothing but VDI that is deployed in cloud and made available to users on subscription basis. DaaS ensures stable and secured work environment. DaaS Service provider is responsible for all kind of administrative tasks such as patch management, system upgrade, data backup, security etc. Microsoft Windows Virtual Desktop (WVD), Amazon WorkSpaces, Citrix Managed Desktops (CMD), VMware Horizon etc., are some of the leading DaaS solutions available in the market.[3]

 ISBN: 978-81-944670-1-4 | November, 2020

C. *Software as a Service (Saas)*

Software as a service (SaaS) makes teleworking easier. SaaS tools are in larger demand during Pandemic situations due to their lower cost, high availability and scalability. SaaS tools are applications hosted on cloud by third party and made available to the companies or individuals on monthly or yearly subscriptions. The virtual work environment is incomplete without essential supporting tools that assist employees in a day to day work. Some of the major SaaS tools are explained here,[4]

Salesforce

Salesforce is a leading application for customer relationship management. It is a cloud-based technology that allows businesses to capture, store and manage customer data in a centralized portal. A CRM framework allows businesses to remain connected to consumers, streamline practices and increase profitability.

Zendesk

Zendesk is a ticketing system for SaaS customer care and support that helps businesses to manage incoming customer requests through multiple channels of communication. With an integrated ticketing system, it simplifies the customer experience and offers you the time to concentrate on stellar customer support. Zendesk is a common platform for all in one customer support.[5]

Microsoft Office 365

The new workplace collaboration product from Microsoft, Office 365, is a cloud-based suite of tools that allow organizations to manage content and collaborate in real-time with employees. All applications and services are linked to each other as well as to the larger web with Office 365, enabling collaboration, saving users time and helping them to work closer together through resources such as group chat, virtual meetings, sharing and co-authoring of data and , and emails.[6]

Slack 365

Slack is an instant messaging application that aims to facilitate business communication for organizations. It helps organizations to coordinate team discussions and open channels for specific topics.

Trello interactive project management

Trello makes it easy to connect with colleagues, collaborate on records, track workflows and reorganize them. What other workers and stakeholders are working on at a glance can be seen by both employers and team members. It is easy to add or move new tasks and append properties, checklists and labels where appropriate. Trello is customizable, as with most SaaS products, and fits with standard document formats.

Calendly meeting and conference scheduler

Calendly is developed to reduce the work involved in planning meetings. User will have to just input their availability details and rest of the arrangements part will be taken care of. Calendly is intelligent enough to understand the time zone difference and act accordingly. Calendly is capable of collaborating with other schedulers like MS Outlook or Google calendar.

Basecamp project management tools

Basecamp is an online project management platform that is cloud-based. Organizations can create topics for discussion, add material, including attachments, and submit content to individuals in the project. The calendar for Basecamp is very detailed. Projects, activities and deadlines can be seen from single view. Tracking each job, project team management for communication, message board management and file storage space allocation are some of the main features of basecamp.[7]

Zoom video conferencing

Organizations may use Zoom as a webinar tool or have an online meeting. A free zoom account allows twenty five people to participate in a single event. Depending on the form of subscription, Organizations can increase the number of participants. There is no upper limit on the number of conferences one can host, once organization have an active Zoom account. The platform comes with a screen sharing feature, making it easier to discuss and visually view various processes or help to solve any issues. Just like any other SaaS tools, it is possible to incorporate and expand the core functionality of the app with other tools. File sharing and cross-platform messaging include these. Another advantage of Zoom is that it allows workers to feel more associated with peers.[8]

Encrypt.me security software

Organizations shall be legally responsible for protecting any personal or consumer details that may be revealed when communicating remotely. There is typically a basic level of security built into SaaS packages, but if the company prefers additional protection, this is something you need to speak to your SaaS provider. Although each employee computer – computers, desktops, tablets and smartphones – should be protected individually.[8]

Adobe Document Sign

Authenticating of HR and legal documents such as letters, agreements, contracts with a signature have become troublesome due to the fact that most documents are now digital. In this scenario Adobe document sign tool can be a perfect solution. Other useful SaaS collaboration applications include Hubstaff, Bit.ai, Canva, Time Doctor and Process Street. Many of these systems are customizable so that resources unique to one's company can be added.[8]

III. VIRTUAL WORK ENVIRONMENTS

The virtual work environment is designed based on the organization's requirements. In the design of virtual work environments, many factors play a key role, such as cost, process confidentiality, legal and regulatory requirements, etc. One or a combination of different types of technology may be used in the virtual work environment. As the company's sustainability depends on the reliability and maturity of the service provider, the organization should carefully choose cloud service providers.

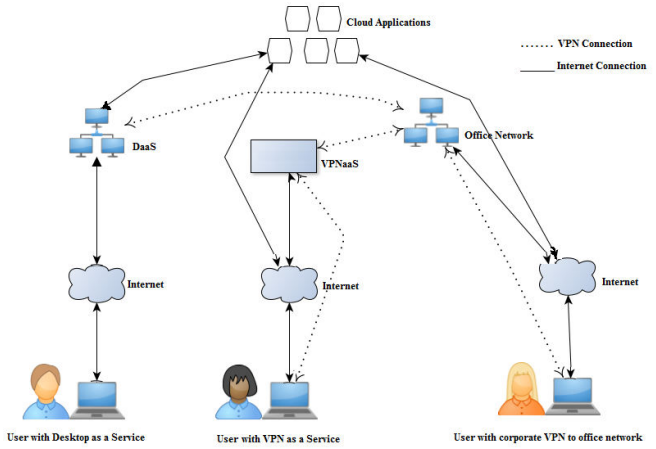


Fig. 1. Virtual Work Environment

Fig. 1. above gives an overall idea of a typical Virtual Work Environment. Three different scenarios such as Users with Desktop as a Service, Users with VPN as a Service, Users with corporate VPN to office network have been considered here. Many more components such as firewall, switches, routers and end point security systems could be a part of typical virtual work environment and these components have not been addressed in this paper.

A. *Users with Desktop as a Service*

In this scenario desktop as service plays a major role as most of the applications, client VPN, security settings are configured in virtual desktop. End user computer acts as a dumb terminal.

- Company sign ups with DaaS service providers such as Microsoft, Amazon, VMWare or Citrix for virtual Desktops on cloud service.
- IT department configures virtual machine as per company requirement such further hardening or installing third party software, installing VPN client to get connected to company network or clients work environment.
- Virtual machine login credentials will be shared with employee. Employee can login to the virtual machine and start working as if he is working from office system.
- User can access cloud based applications from same virtual machine and all business related data will be safe in the cloud with all the security in place.

B. *Users with VPN as a Service*

It has been observed that some of the IT companies found it challenging to arrange VPN access to their entire workforce due to shortage of VPN licenses or limited resources. VPN as a service helped such companies to get connected to corporate network or client's network overcoming resource challenges.

- Company and VPNaaS provider gets into service agreement.
- Service provider enables cloud based VPN service to which multiple users can establish the VPN connection.
- Service provider establishes a site to site VPN connection between company network or client network to the cloud.
- VPN login credentials for cloud based VPN will be allocated to each user.
- Users can now access resources from office network or from client network via cloud based VPN.
- If there is no Virtual Desktop Infrastructure available inside the office network, the employee's computer would be used to do office work.
- With or without a VPN, cloud applications can be accessed according to the access rules configured at the cloud service provider end.

Above model helps organizations to overcome resource bottleneck, licensing challenges etc. Cloud service provider will be responsible for the VPN uptime, scalability and availability of the service.

C. *Users with corporate VPN to office network*

A simple and most common way of enabling virtual work environment is represented by this model. Here users directly gets connected to office network in order to access resources available via office using SSL VPN or IPSec VPN.

- Company deploys VPN gateway by itself
- Company creates login credential for each remote working employee and shares it with them individually.
- Employee either installs VPN desktop client or uses his browser to establish a VPN tunnel to the office network based on the type of VPN protocol enabled.
- Once VPN is established, user can access necessary data and applications via office network. Employee's computer will be used to carry out office work if Virtual Desktop Infrastructure is not available within the office network.
- Cloud Applications will be accessed with or without a VPN as per the access rules configured at cloud service provider end.

In above scenario company will have added responsibility of keeping VPN gateway up and running all the time. High availability, redundancy, service uptime etc. are additional workload compared to first two models. If company has sufficient VPN licenses then there will be some savings towards cloud services. This model is widely deployed among SMB sector since most of the next generation firewalls and unified threat management devices has built-in VPN module.

Company can opt for any one of the above model or go for combination of any of the above discussed models based on business requirement and network architecture.

IV. Data and End Point Security

Work from anywhere on virtual environment increases security threat to the business and data as each end point is exposed to the public internet. Insecure WiFI connections, outdated security patches, lack of awareness can lead to security incidents. Organization will have to give proper

attention to each components involved with virtual work environment. Starting from user WiFi connection to cloud service providers security competency needs to be evaluated and corrected if any breaches found. Zscalar is one of the cloud based security tool that provides Internet security, data loss prevention, SSL decryption, bandwidth control etc. [9]

V. Conclussion

Timely availability of the cloud based tools and advancement in virtual platform technology helped IT Industries to build quick and safe virtual working environment during pandemic. Even though plenty of SaaS and IaaS applications are available in the market, industry has to evaluate the service providers carefully or else virtual work environment might become a big disaster as all the teleworking computers are directly connected to the public internet. Organizations will have to give more importance to security and continuous improvement to overcome any unseen challenges.

References

[1] Sreeramana Aithal, & Suresh Kumar P. M. (2016). Working from Home - A Transition in the concept of Workplace. International Journal of Current Research and Modern

[2] VPN-as-a-Service. (n.d.). Retrieved November 12, 2020, from https://aisn.net/vpn-as-a-service/

[3] Microsoft Windows Virtual Desktop vs Amazon AWS Remote Workspaces. (n.d.). Retrieved November 12, 2020, from https://www.remoteworkstations.com/microsoft-windows-virtual-desktop-vs-amazon-workspaces

[4] Team, C. (2020, June 09). Best tech tools for the virtual workplace: Workable. Retrieved November 14, 2020, from https://resources.workable.com/tutorial/tech-tools-for-virtual-workplace-digital -transformation

[5] Sureka, A. (n.d.). 28 Important Zendesk features for Better Customer Support. Retrieved November 12, 2020, from https://www.clariontech.com/platform-blog/28-important-zendesk-features-for-better-customer-support

[6] Wright, N. (2019, August 09). Everything you ever wanted to know about Office 365. Retrieved November 12, 2020, from https://www.nigelfrank.com/blog/everything-you-ever-wanted-to-know-about-office-365/

[7] Project Management – Basecamp: Features, Advantages & Disadvantages. (2018, November 12). Retrieved November 14, 2020, from https://www.edupristine.com/blog/up-and-running-with - basecamp

[8] The top 7 SaaS tools for remote working employees. (n.d.). Retrieved November 14, 2020, from https://www.soldo.com/gb/resources/the-top-7-saas-tools-for-remote-working-employees/

[9] What is Cloud Security?: Understand the 5 Core Principles. (n.d.). Retrieved November 14, 2020, from https://www.zscaler.com/solutions/cloud-security

Using Machine Learning Algorithms Detecting Distributed Denial of Service Attacks

Sangeetha Prabhu
College of Computer Science and Information Science
Srinivas University
Mangalore, India
sangeethaprabhu.ccis@srinivasuniversity.edu.in

Abstract – **The excessive dependence on the Internet has opened possibilities for elevated cyber threats and perpetration of a broad range of cyber crimes, resulting in massive economic losses and user data privacy violations. One of the recent but harmful releases to the library of malicious software is the DDoS. The DDoS attacks are a major threat to information security. These threats are sometimes obtained from virtual servers in the cloud, rather than from attacker's own device, to accomplish confidentiality and greater network bandwidth. The DDoS attacks have grown more complex and continue to grow in number day after day, thereby making it harder to track and fight such attacks. Thus, there is a necessity of intelligent intrusion detection system (IDS) to track and identify any anomalous activity of the network traffic. This paper concentrates on network and application layer attack and strategy of DDoS. This research work includes Intrusion Detection Methodologies, Machine Learning techniques to identify the DDoS attacks.**

Keywords—Cyber Security, DDoS, DDoS Detection, Traffic Analysis, malicious activity, Intrusion Detection

I. INTRODUCTION

The Denial of Service (DoS) is a network security issue whenever a directed service assault resulted in a valid user's lack of availability or system deterioration. .The service can be a single computer, a set of computers or even a system. The DoS attack is effective if an intruder can try to bring the functionality of the specific part into a state of ignorance for authorized customers[1]. This assault could be taken out on the OSI and TCP/IP prototypes in a range of methods and on various layers. The basic implementation strategy of each type of DoS attack is influenced by a variety of factors, including the software package that used produce network attacks, targeting protocol, target nature, etc..

The victim's primary objective is to position the target resource in a state of confusion to the authorized customers. While numerous security measures may be introduced to essential infrastructure to avoid such assaults, the weaknesses that remain in the frameworks are computer world facts that are abused by attackers. Some forms of DoS assaults listed in [2], powered by different techniques, aimed at specific sources & using their subjection are highlighted in Table 1.

Table: Various kinds of Denial of Service attacks [1]

DoS attack Types	Common targets	Subjection
Data flooding	Network bandwidth or capacity of networks or servers	Limited network bandwidth/Limited server capacity of processing requests
Attack on network devices	Network devices / hardware like switches routers, and firewalls	Vulnerableness/bugs in device software
Protocol attack	Protocol services	Restrictions and drawbacks of the protocol, like ARP IP spoofing, poisoning, and TCP SYN floods
Application attack	Application services	Application limitations and weaknesses
Operating system attack	Operating System services	Vulnerableness/bugs in Operating System software

The thermal DDoS attack, which would be accountable for further than 65 percent of such assaults, is among the most harmful attack users on a Internet [3]. In a thermal DDoS attack, inside an effort to overwhelm the victim's computational capabilities or the close networking devices, many hackers organize the transmission of a greater incidence of random information. On one hand, since the primary Service routers usually use the FIFO and DROP-TAIL having to queue techniques, the strong efficiency scores for this sort of attack arise because they do not distinguish between kinds of data, causing similar failure rates for assaults and logical approach[4]. While legitimate traffic appears to withdraw in order to avoid more congestion, this commitment is not made to assault traffic and allows the links to be overwhelmed. Legitimate traffic is as a result, even obstructed[3].

II. CONCEPT OF DENIAL OF SERVICE

A. *Related Work*

Successful identification of irregular network activity is the premise for maintaining the regular company performance and ensuring a reasonable standard of service. The detection techniques of DDoS attacks at domestically and overseas are actively investigating and

inventing new, statistical-based and ML are the popular identification technique methods, like entropy identification, SVM[5].

a. Methods Statistical-Based.

To detect intrusions in flow, Tao[6] utilizes entropy adjustment. When an event is identified by the protection system, this will obstruct or restrict irregular flow and identify the position of the intruder. To distinguish DDoS attacks against flash crowds, data distance is used. If the communication barrier is within a specified threshold in the suspect flow, this will be defined as a DDoS attack, otherwise it would be a temporary obstruction of the channel. Mousavi [7] suggests a framework for detecting technology depends on target entropy. In order to measure entropy, the study uses the relationship among source & target IP addresses to establish the standard by metrics to determine how attack takes place. Entropy is, proportionally speaking, never deemed a reasonable measure[6] since it has a reasonably high mixed and inconclusive value. A new approach is proposed by Dong et al.[8] to identify the DDoS attack and further discover the infected ports attached by cyber hackers. First, utilizing SPRT, which has restricted false positives and false negatives failure rates, they identify the stream hazards connected with a terminal, instead make decisions. To identify DDoS, Zhang et al.[9] used the ARIMA process. Studies by Kumar[10-11] help establish the differences between both the amplitude response of the Smurf attack flow and the initial legitimate traffic.

b. **Methods based on Machine Learning**

sKim et al.[12] evaluate that even an HTTP Distributed Denial of Service attack happens by measuring the amount of variance of respective target domains, the estimated number of important queries, the estimated byte of information transmitted, the estimated depth of the reference tree, as well as the comparison between usual and unusual traffic. Jacob et al.[13] represent a method that uses learning algorithms to construct a graph-based framework and determine whenever the gap between both the network environment and the unusual link triggers an assault. BOTFINDER[14] identifies contaminated domains in the system using bot's elevated internet traffic property only. It requires users to recognize the main characteristics of interaction control and order, based observer traffic created by bots in the managed environment. Using such functions, BOTFINDER builds prototypes which can be implemented to classify compromised host at channel exit points. The analysis found that the concentration of statistical identification has detection accuracy performance, but constructing the prototype takes a lot of time. The accuracy rate of machine learning algorithms is strong, however the detection efficiency is poor.

III. UNDERSTANDING DDoS ATTACK

An intruder tries blocking-authorized customers of a utility from accessing the service during the DDoS attack. Distributed Denial-of-Service is the attack executed several sites, so it's not just either one or two IP addresses have to be blocked. The attack could target a weakness in the network of a third party, for example DNS or NTP because you are potentially getting packets that could not be shuttered from mainstream websites such as corporations or institutions, although there are active initiatives to find and notify such sites of the issue and need them to fix their service. We highlight the specifics for clarification of certain kinds of attacks. If the intruder X has IP address 4.3.2.1 and the target Y has IP address 8.7.6.5, X will transmit the packet to abc.com with Y IP address 8.7.6.5 as the root and say "instruct me more about Z. So abc.com sends the intruder X a lot of information, which he did not ask for. If X does that to xyz.com, def.com, etc., all of which ask them to deliver 8.7.6.5 info, this is a DDOS attack[15]. As a consequence, the suspect's link buffer will be loaded with unresolved interactions which will never be accomplished, thus preventing it from responding to new inquiries that may be legitimate.

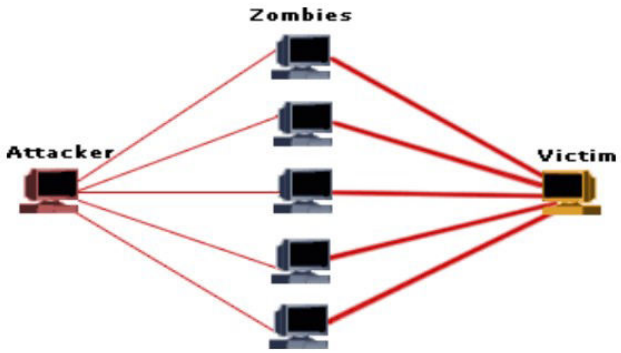


Fig. 1. DDoS attack[15]

On various layers of the OSI model, Intrusions can be inserted. We explain in this section the monitoring of the new network attacks and the application layer.

A. Network Layer Attacks

The network layer threats include a Smurf intrusion and a UDP flood threat. In a Smurf attack, the target is overwhelmed with echo-reply frames from the ICMP. This intrusion uses IP airwaves, in which packets are transferred from a system on a local network to an IP routing protocol, the packets are delivered to all devices on a certain network. The intruder transmits packets with fake websites source Address aimed at the recipient under such conditions. As these packets are transferred at the address of the transmissions, all nodes within the network receive them. As the root IP address is faked as the attacker's address, each device reacts to the target computer. In UDP Flood intruder passes a significant number of UDP packets to remote nodes of their targeted system, this results in network congestion and exhaustion of usable capacity to the victim device for valid requests[15]. A victim device will examine the queuing request at the port of destination when acquiring a UDP packet. If no request is pending on the terminal, it will create unavailable target ICMP packets for the faked source node.

B. Application Layer Attacks

The new forms of attacks in the application layer are UDP flood and HTTP flood. The intruder takes advantage of HTTP GET enquiries in the HTTP flood to target a web browser or software. It is also considerably difficult to catch

and stop these threats. A HTTP client is a kind of internet browser communicating to a server by forwarding either a GET or POST kind of a request from HTTP. The user sends an HTTP response to user in an inaccurate HTTP flood attack using the GET form but in a new context[15]. The client would not send the entire HTTP header, but it sends a portion of it. At various points, the client regularly sends corresponding header to hold the connection active. By submitting several imperfect queries, the services of the server are drained. Such queries use all the necessary resources on the network, thereby rejecting the queries of authorized customers. Furthermore, SQL Injection DDoS are the most recent DDoS application level threat, in which hackers begin from client end, such as browsers, by injecting a Ransomware and routing it to the network side [16].

IV. DDOS ATTACKS DETECTION METHODOLOGIES

With the rapid growing popularity of cloud services, it continues to be a major research topic to guarantee the protection and accessibility of services, information resources. DDoS attacks are not just a new challenge; they are a serious security problem and a wide-ranging research topic. This section addresses the various techniques of DDoS intention and launching, which can be used to execute or encourage DDoS attacks and examine techniques and security strategies for malware detection.

A. Intrusion Detection Methodologies

Some techniques for attack detection [17] are listed here. Three categories of such techniques are categorized: Signature-based Detection, Anomaly-based Detection & Stateful Protocol Analysis. The positives and negatives of the methodologies of identification are also listed.

a. **Signature-based Detection:** By monitoring activities and detecting patterns that suit the signatures of known threats, the Signature-based ID technique can identify intrusions. The significant actions needed to perform the assaults and arrangements in which they must be accomplished are defined by an attack signature. In addition, this only identifies threats whose signatures have recently been stored in the database. The signatures must be changed periodically for tracking purposes. Much as potential hazards are frequently issued, creating need for signature updates, new risks are regularly discovered against your servers[18]. Only the set behavioural pattern fits well against such a strategy. They struggle to deal with human-created threats or worms with behaviour patterns that change it.

b. **Anomaly-based Detection:** Because of its capacity to track novel threats, the AD drew many researchers. Detection is based on identifying the behaviour of the service. The behaviour of the system is consistent with the behaviour required. So it is acknowledged otherwise the activity in the AD is caused. The recognized network behaviour can be planned or taught by the requirements of the system administrator [17]. The main aspect of AD compared to signature-based machines is if it drops out of usual traffic conditions, a unique assault in which no signature will occur can be identified.

c. **Stateful Protocol Analysis:** The SPA demonstrates that the protocol specifies may be known and traced by IDS. AD considers pre-loaded system or network-specific profiles while SPA relies on developer built standardized profiles to standard methods that define how individual protocols should and should not be used. In an attempt to comprehend and mitigate the risks, SPA provides essential capability. The SPA depends on well-behaved and well-defined protocol frameworks as an intrusion identification tool[17]. SPA is less reliable in situations at which a protocol is confidential, improperly specified or the design of a vendor differs from the template.

B. Machine Learning Techniques

We briefly define the different ML algorithms in this part and the problem domains in which they are commonly used. As in literature, several decision tree and rule induction strategies have been proposed. A probabilistic classifier is the Naive Bayes [18] algorithm, which believes that impact of uncertain parameters on a specific position is conditionally independent of other parameters. This hypothesis is called conditional independence in class. The most well-known and used algorithms for classification are the decision tree algorithms. A common classifier is the C4.5 algorithm [19], designed by Ross Quinlan. This algorithm is focused on an algorithm that seeks to identify a limited decision tree focused on ID3. For classification, the decision tree algorithm is developed by C4.5 can be used, it is sometimes pointed as a statistical classifier. The C4.5 algorithm was identified by the developers of Weka ML software as pioneer decision tree software that is currently the most frequently used ML workhorse of far in practice [20]. K-Mean Clustering [20] is the distribution of data sets to groups varies between the middle and the data points of the group.

The k-NN algorithm [18] is a learning algorithm groups with similar and is recognized to be extremely efficient, particularly grouping problems in specific problem domains. The most reasonable method for ML tasks is SVM [18] in regression and classification. Support Vector Machines can be extended not just to issues of grouping, but to also representation learning. Fuzzy c-means clustering [18] is a grouping approach that enables two and more groups to contribute to one piece of information. In information processing, this mechanism is also used. Numerical structures influenced by the role of a human mind are the Neural Networks [18]. In the literature [18], several neural network applications were recommended: stock market prediction, electrocardiogram, defence, character recognition, and payment history and several other instances.

For anomaly detection [21], ML techniques are widely used. Through resolving the shortcomings of knowledge base intrusion detection, they have attracted extensive interest from intrusion detection experts. Experiments performed using [22] suggest that C4.5 are most robust than k-NN. A further research on the three C4.5, Multi-Layer Perception, and SVM classifiers [18] based detection functional prototypes revealed that C4.5 is the better

approach of detection accuracy and minimal skill training; it obtained an average accuracy of 99.05 percent.

V. DDOS STRATEGY

As shown in Figs. 2 and 3, a Distributed Denial of Service attack consists of various components. Fig. 4. Shows many steps to initiate an assault on Distributed Denial of Service.

1. **Choosing the representatives:** The intruders select agents who will execute the threat. Some computers are exploited to use it as agents, depending on the type of bug present. Such systems, that have significant capabilities, are exploited by hackers so a strong attack flow can be created. The intruders tried physically gaining control of such systems in the initial years. .However, it became easier to understand such devices automatically and efficiently with the advent of advanced security products.
2. **Compromise:** The intruder utilizes agent systems security holes and weaknesses, and installs the malware program. Not just that the intruder often uses the requisite measures to defend against detection and neutralization of the installed code. According to the ability to communicate for DDoS attacks, as shown in Fig. 2, the compromised nodes, i.e. zombies between the intruder and the target, are hired by unsuspecting informant hosts from a wide range of high throughput insecure hosts connected via the Internet. The methodology for DDoS attacks is shown in Fig. 3 is more complicated due to the integration between zombies and the target(s) of the intermediary layer(s). It further simplifies the traceback mainly because of (i) the difficulty of detaching the (partial) traceback data in relation to numerous perspectives and/or (ii) the need to link a huge proportion of routers or servers. This process is automated by self-propagating techniques like the Ramen worm[23] and Code Red[23].It is relatively hard for the clients and operators of the agent systems to recognize because they've become member of a Distributed Denial of Service attack mechanism until a specialized defense mechanism is being used. An key benefit of an agent system that in memory size and capacity, the client systems are quite expense. Thus, they moderately influence the system's efficiency.

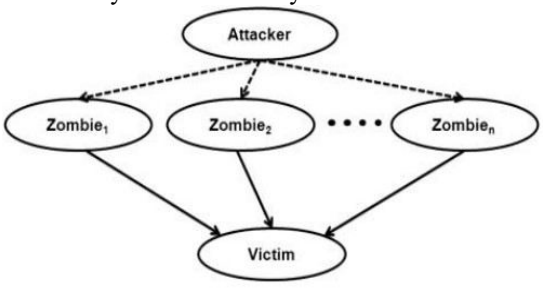


Fig. 2. Direct DDoS attack: Deliver the zombies specifically to monitor traffic to strike the attacker's host[23].

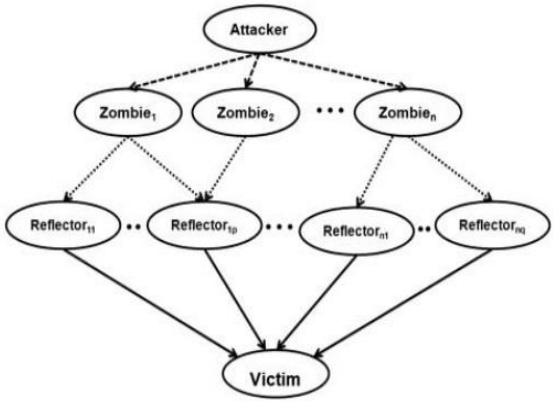


Fig. 3. Indirect Attacks: to break the host machine, forward the control traffic to zombies explicitly [23].

3. **Communication.** In order to determine the operators are operational, when and where to plan attacks or when to update agents, the intruder interacts with any quantity of controllers. Different protocols, like UDP, TCP, ICMP, may be used for certain interactions between hackers and controllers. Agents may interact with a particular controller or several controllers depending on the specification of the threat system.
4. **Attack**. The intruder causes an assault. It is possible to change the target, the period of the assault, and surgical procedure characteristics such as form, frequency, duration and node details. There are major discrepancies in the characteristics of incoming packets, it's indeed helpful for the intruder, as it affects identification.

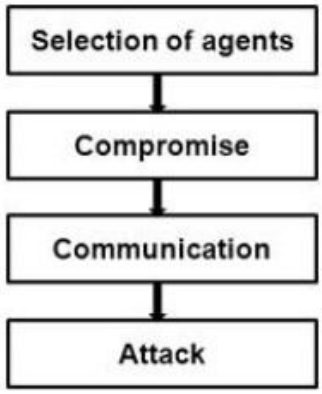


Fig. 5. Steps to perform a DDoS attack[23].

Over the recent decade, intruders and hackers have begun using an online multi-user chat platform called as IRC[23]. It's because clients are able to establish public, secret and private channels via IRC chat channels. An IRC-based DDoS attack framework is equivalent to operative DDoS attack system, apart from that an IRC network monitors the address of linked operators and controllers and establish a connection between them rather than using a controller program installed on a server. A particular user's exploration enables the detection of the channel of communication; however the identification of other users remains covered. These "web - based frameworks or IRC frameworks for multiple users have many other strategy to

ensure initiating DDoS attacks. Amongst these significant benefits are they have a high level of anonymity, are hard to track, and to provide a powerful distribution mechanism that is assured. In addition, this is no longer important for the intruder to manage a catalog of operatives, as he can easily sign on to IRC server and check all active users[23].

notifications about the condition of operators from the operator program and engage in alerting the hackers about the operators' condition. The operators are sometimes pointed to as "Zombie Bots" or "Bots" in an IRC-based Distributed Denial of Services attack.

CONCLUSION

The biggest threats to Network were Cyber attack which can cause great damage to industry and individuals. With establishment of evolving techniques like cloud computing, IoT, artificial intelligence, and hackers will conduct reduced DDoS attacks, making it far more difficult to identify and avoid DDoS attacks. A variety of Machine Learning strategies were used to classify DDoS attacks and identify DDoS attacks, like Naive Bayes and Decision Tree, etc. Today, cyber security draws significant interest from both academic world and commercial businesses. Designs of Research start to evolve and consumer applications are now accessible, based on the early studies. We provided an outline of DDoS attacks, identification strategies and DDoS strategy in this article. A comparative analysis of the latest intrusion detection reveals that almost all systems are not able to meet the all real-time system requirement.

REFERENCES

[1]. Aamir, M., & Zaidi, S. M. A. (2019). DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation. *International Journal of Information Security*, *18*(6), 761–785. https://doi.org/10.1007/s10207-019-00434-1

[2]. Christos, D., & Dimitrios, S. (2007). Network Security Current Status and Future Directions. *Wiley IEEE Press*, *6*, 572.

[3]. Cao, Y., Gao, Y., Tan, R., Han, Q., & Liu, Z. (2018). Understanding internet DDoS Mitigation from academic and industrial perspectives. *IEEE Access*, *6*, 66641–66648. https://doi.org/10.1109/ACCESS.2018.2877710

[4]. Lima Filho, F. S. De, Silveira, F. A. F., De Medeiros Brito Junior, A., Vargas-Solar, G., & Silveira, L. F. (2019). Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning. *Security and Communication Networks*, *2019*. https://doi.org/10.1155/2019/1574749

[5]. Qian Li,Meng Linhai,Zhang Yuan, Y. (2019). *DDoS Attacks Detection Using Machine Learning Algorithms* (Vol. 1009). *International Forum on Digital TV and Wireless Multimedia Communications*, Springer Singapore. https://doi.org/10.1007/978-981-13-8138-6

[6]. Tao, Y., & Yu, S. (2013). DDoS attack detection at local area networks using information theoretical metrics. *Proceedings - 12th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, TrustCom 2013*, 233–240. https://doi.org/10.1109/TrustCom.2013.32

[7]. Mousavi, S. M., & St-Hilaire, M. (2015). Early detection of DDoS attacks against SDN controllers. *2015 International Conference on Computing, Networking and Communications, ICNC 2015*, 77–81. https://doi.org/10.1109/ICCNC.2015.7069319

[8]. Dong, P., Du, X., Zhang, H., & Xu, T. (2016). A detection method for a novel DDoS attack against SDN controllers by vast new low-traffic flows. *2016 IEEE International Conference on Communications, ICC 2016*. https://doi.org/10.1109/ICC.2016.7510992

[9]. Zhang, G., Jiang, S., Wei, G., & Guan, Q. (2009). A prediction-based detection algorithm against distributed denial-of-service attacks. *Proceedings of the 2009 ACM International Wireless Communications and Mobile Computing Conference, IWCMC 2009*, *January*, 106–110. https://doi.org/10.1145/1582379.1582403

[10]. Kumar, S. (2007). Smurf-based Distributed Denial of Service (DDoS) attack amplification in internet. *Second International Conference on Internet Monitoring and Protection, ICIMP 2007*, *0521585*. https://doi.org/10.1109/ICIMP.2007.42

[11]. Kumar, S., Azad, M., Gomez, O., & Valdez, R. (2006). Can microsoft's Service Pack2 (SP2) security software prevent SMURF attacks? *Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services, AICT/ICIW'06*, *2006*, 89. https://doi.org/10.1109/AICT-ICIW.2006.60

[12]. Kim, S. J., Lee, S., & Bae, B. (2014). HAS-analyzer: Detecting HTTP-based C&C based on the analysis of HTTP activity sets. *KSII Transactions on Internet and Information Systems*, *8*(5), 1801–1816. https://doi.org/10.3837/tiis.2014.05.017

[13]. JACOB GREGOIRE; HUND, RAALF; KRUEGEL, C. H. T. (2011). Picking Command and Control Connections from Bot.pdf. *USENIX ; Login*, *36*(5), 16. https://www.usenix.org/legacy/event/sec11/tech/full_papers/Jacob.pdf

[14]. Tegeler, F., Fu, X., Vigna, G., & Kruegel, C. (2012). BotFinder: Finding bots in network traffic without deep packet inspection. *CoNEXT 2012 - Proceedings of the 2012 ACM Conference on Emerging Networking Experiments and Technologies*, 349–360. https://doi.org/10.1145/2413176.2413217

[15]. Sofi, I., Mahajan, A., & Mansotra, V. (2017). Machine Learning Techniques used for the Detection and Analysis of Modern Types of DDoS Attacks. *International Research Journal of Engineering and Technology(IRJET)*, *4*(6), 1085–1092. https://irjet.net/archives/V4/i6/IRJET-V4I6200.pdf

[16]. Alkasassbeh, M., Al-Naymat, G., B.A, A., & Almseidin, M. (2016). Detecting Distributed Denial of Service Attacks Using Data Mining Techniques. *International Journal of Advanced Computer Science and Applications*, *7*(1). https://doi.org/10.14569/ijacsa.2016.070159

[17]. Liao, H. J., Richard Lin, C. H., Lin, Y. C., & Tung, K. Y. (2013). Intrusion detection system: A

comprehensive review. *Journal of Network and Computer Applications*, *36*(1), 16–24. https://doi.org/10.1016/j.jnca.2012.09.004

[18]. Zekri, M., Kafhali, S. El, Aboutabit, N., & Saadi, Y. (2018). DDoS attack detection using machine learning techniques in cloud computing environments. *Proceedings of 2017 International Conference of Cloud Computing Technologies and Applications, CloudTech 2017*, *2018-January*, 1–7. https://doi.org/10.1109/CloudTech.2017.8284731

[19]. Quinlan, J. R. (2014). *C4. 5: programs for machine learning*. Elsevier.

[20]. Witten, I. H., Frank, E., & Hall, M. a. (2011). Data Mining: Practical Machine Learning Tools and Techniques (Google eBook). In *Complementary literature None*. http://books.google.com/books?id=bDtLM8CODsQC&pgis=1

[21]. Lane, B., Poole, M., Camp, M., & Murray-krezan, J. (2016). Using Machine Learning for Advanced Anomaly Detection and Classification. *Advanced Maui Optical and Space Surveillance Technologies (AMOS) Conference*.

[22]. smanto, H., & Wardoyo, R. (2016). Comparison of running time between C4.5 and k-nearest neighbor (k-NN) algorithm on deciding mainstay area clustering. *International Journal of Advances in Intelligent Informatics*, *2*(1), 1. https://doi.org/10.26555/ijain.v2i1.49

[23]. Bhuyan, M. H., Kashyap, H. J., Bhattacharyya, D. K., & Kalita, J. K. (2014). Detecting distributed denial of service attacks: Methods, tools and future directions. *Computer Journal*, *57*(4), 537–556. https://doi.org/10.1093/comjnl/bxt031

A Study on Natural Language Processing and Machine Learning Application areas

Suchetha Vijayakumar
Research Scholar
Srinivas University, Mangalore
such_vijay@yahoo.com

***Abstract*— Natural language processing (NLP) is meant for representing and analyzing human language computationally. It is widespread nowadays in various areas such as machine translation, email spam detection, information extraction, summarization, medical, question answering etc. Understanding and translating Human spoken language has always been a great challenge for computers. Be it in terms of translation or understanding, it has always been an incomplete and incompetent task. In the present days, advanced techniques like deep learning is said to do this task without much hassles. But still machine learning models are found to be left behind in understanding what human language really means. Natural Language Processing (NLP) and Machine Learning (ML) play a vital role in taking the above scenario further. NLP and Machine Language have already gained entry to sentiment analysis, speech recognition, text classification, machine translation, question answering etc. to name a few. During the past few years, we could also see a dramatic shift in the NLP research into large scale applications of statistical methods, such as machine learning and data mining. This paper contains few prospective areas of studies and research with respect to NLP and ML.**

Keywords—Natural Language Processing, Machine Learning, Sentiment analysis, Transfer Learning, Reinforcement Learning

I. INTRODUCTION

Natural Language Processing (NLP) combines Intelligence and Linguistics so as to make computer systems understand human languages. The main purpose here is to satisfy human desire to communicate with the computers using natural spoken language. NLP paves way to all those who are not well versed with machine specific language but still would want systems to understand the language spoken by them. Natural language processing is the study of mathematical and computational modelling of various aspects of language and the development of a wide range of systems with extended capabilities.

Machine Learning (ML) is built on Artificial Intelligence. The motive behind ML is to study and understand the data , devise methods to convert and fit the same into what is called as models that can be utlised by people. Basically the models tell the computer systems the best ways to handle data efficiently. An overview of the data that is available for study may not give a brief understanding and definitely will not help in any way to come to conclusions. Presently huge amounts of data are available in various domains which are obtained through various sources. Therefore, we have a big challenge in front of us to see to that the computer systems are able to learn on their own and also analyse the data fed to them to arrive at conclusions or decisions.

II IDENTIFIED RESEARCH AREAS IN NLP AND ML

In order to make machines understand, they need to be taught. In other words, "machine teaching" needs to implemented. Therefore, to accomplish this task we need to devise a proper framework and also feed properly formatted, relevant, cleaned data into the system or machine. NLP algorithms will be of help to us in achieving this. ML on the other hand can be helpful in finding or devising a new model to handle further tasks irrespective of how easy or complicated the task is.

In this section we focus on few important areas and applications where NLP and ML has been effectively used.

A. Text Minng , Text Analytics and Text Processing:

Text mining is a technology which is an extension of Artificial Intelligence. Here the unstructured text of various types and from various sources are converted to structured text. NLP algorithms help in achieving this transformation. The resultant structured text can be fed into various ML models for further processing and analysis. Text analytics can use unsupervised and supervised learning to identify technical terms and parts of speech while unsupervised learning can determine symbiotic relationships between them. The various text processing applications that can be done using trained ML models are Language translation, Sentiment analysis, Email spam filtering to name a few. Language translation is used to translate the input text from one language to another language. Sentiment analysis classifies the text into various categories of sentiments such as positive, negative, neutral etc from a corpus of structured text. Email

spam filtering detects and sorts unwanted mails that come to inbox.

B. Training NLP models with reinforcement learning:

Machine Learning identifies a special area called as Reinforcement Learning. Here the machines or systems will find out and determine an ideal and perfect behavior for a given context so as to get maximum performance. This type of learning ensures that the machines of systems will get used to the behavior through the feedback from its working environment. Reinforcement learning can be placed in between supervised and unsupervised learning. The reason for doing so is that there is some form of supervision taking place though not in terms of output or result specification. Reinforcement learning can get feedback from the environment only after deciding and confirming the output based on the input that is given. Reinforcement learning is applicable in various areas such as sequential decision making problems. Here the user needs to interact with the system in a sequential manner by giving sequence of inputs and receiving outputs along with the feedback for every action.

It is proven that Reinforcement learning is efficient when compared to supervised and unsupervised learning. But training models from scratch in terms of Reinforcement learning is still a difficult task as the system might become very slow and unstable sometimes. Hence experts have been working on training NLP based supervised models first and then improve the same by fine tuning with Reinforcement Learning.

C. Recurrent Neural Networks and NLP:

Recurrent Neural Networks(RNNs) is a special type of neural network. Unlike the other neural networks, in RNN the output of the preceding step is required to be fed as an input in the current step. So there arises a need to remember the previous steps output. Technically this is accomplished with the help of a hidden layer. NLP is a good case study for RNNs. If we give three words and want the system to predict the fourth word, it can be handled effectively by RNNs. Therefore, RNNs become a perfect choice for research involving sequence labelling, text/document classification, text generation, entity tagging and music generation. RNNs can be used to do the above tasks by constructing appropriate language models so as to reduce the possibilities of model getting confused with a given sequence of words. A language model encoder needs to be created first post which it can be reused so as to save training time.

D. Market intelligence monitoring:

Business organizations can use NLP and ML to play a wider role in the market. NLP algorithms can be used to track and monitor market intelligence reports so that the organizations can plan the future. Analysing various interests of customers, their sentiments, pattern of buying, pattern of spending etc can shed more light on developing and enhancing business opportunities. Customers complaints can also be tracked and solved. Also can be used to keep a watch on competitors. Financial marketing uses this technology to predict money related matters and also strategically devise ideas to resolve and plan their future. With the help of ML models and algorithms, these information can be processed in a much faster speed and also provide a suitable real time analysis of the same. As a result of this various departments and teams can only focus on higher prospects of business than spending time and resources on data collection and later on doing an analysis.

E. Fine-tuning models with Transfer Learning:

Transfer learning in NLP enables extraction and transfer of knowledge learned and gained from one source to a new destination. Using this technique, a pre trained model which is tested on one dataset can be used for a similar task in the same domain. In the recent years several Transfer learning methods and implementations have emerged thus proving its efficiency. It is also found that the new Transfer learning methods are very easy to integrate and also the performance and results are fairly appreciable. One of the well-known applications of Transfer Learning is detection of subtypes in Cancer. Another area where Transfer Learning has found its way is Simulation and it found that Transfer Learning enables many advanced ML systems in the real world.

F. Customized product recommendations:

In the present growing world of Internet, many e-commerce companies face the challenge of identifying human thought process. Many researches are carried out in order to bridge the gap between human thought process and online shopping. NLP based sorting has proved to be efficient in solving and also finding a solution for the long lasting problem of e-commerce companies. It utilises the huge amount of data generated through e-commerce websites on a daily basis and effectively organises them to generate suggestions and also search results. This is the reason why customers are able to see products of their choice

and requirements while doing online shopping. Hence NLP and ML can be used to enhance customer engagement by analysing their shopping trends, purchase behaviour, auto generated product descriptions to name a few.

G. Intelligent semantic search:

Semantic search is different from an ordinary search in which the search is extended to synonyms of the word to be searched. Google search is a classic example for semantic search. NLP and ML are used to construct and host such semantic search applications. NLP along with Natural Language understanding (NLU) does a comprehensive study regarding the core idea contained within the text and also analyses the parts of speech in the user query. ML can add on to the task of building semantic search engines when there is a necessity to build domain specific, comprehensive and low cost resources. Hence an integrated ML and NLP architecture would be of help in this scenario.

H. Intelligent cognitive communication:

Cognitive computing is an upcoming technology which is an integration of powerful technologies such as Artificial Intelligence, Machine Learning and Natural Language Processing. Some of the important features of cognitive computing are text analytics, voice recognition and analytics, image and visual analytics. These systems are considered to be self-learners. Once the initial instructions are given, these systems will start learning on their own depending on the data that is input. In other words, it is making the computers think like humans but with high computational power and memory. Therefore, NLP and ML can be integrated to enhance cognitive computing.

I. Growth in chatbots and virtual assistants:

Customer service and experience are the two important external factors for the success of any company. With the help of these two factors the companies can improve their products and at the same time can cater to customer satisfaction. Interaction with each and every customer is probably an impossible task for most of the companies. Chatbots help companies in such scenarios. These chatbots are built to obey automated rules and use technologies such as NLP and ML. These advanced technologies see to that chatbots process all types of data and also respond to any type of requests and commands. Using the capabilities of NLP and some ML, the chatbots respond in voice messages to user enquiries. Hence chatbots have these days become one of the most sought after support and service functions

J. Sentiment analysis for social media:

Sentiment analysis is also known as Opinion mining or emotion AI. This is a subfield of NLP which can identify and extract opinions from a huge repository of data that arises through blogs, reviews, social media, forums and many more. NLP can convert all these unstructured data into structured data through its algorithms and open source tools. Together with ML, NLP can do lots of other tasks such as classify the opinions as positive, negative or neutral, analyse the opinions with respect to attitude and background. The ML driven NLP helps in analysing the huge amount of text generated in the above mentioned sources in no time. Though this technique started few years back, it is becoming better day by day and also catering to the needs of analysts. NLTK (Natural Language Toolkit) is used for the purpose of analysing sentiments by processing natural languages.

III CONCLUSION

Natural Language Processing and Machine Learning are the two buzzwords heard of recent when it comes to model building and training a model. The future of this world in the current pandemic situation calls for many such model predictions and their subsequent development and implementation for the broader development and survival. It can create wonders and magic especially when both NLP and ML are used together in any one of the above mentioned area of study.

REFERENCES

[1] Prof. Alpa Reshamwala, Prof. Dhirendra Mishra, Prajakta Pawar "REVIEW ON NATURAL LANGUAGE PROCESSING" IRACST – Engineering Science and Technology: An International Journal (ESTIJ), ISSN: 2250-3498, Vol.3, No.1, February 2013

[2] Lluís Marquez and Jordi Girona Salgado,"Machine Learning and Natural Language Processing",2000

[3] Akshi Kumar , Teeja Mary Sebastian "Machine Learning assisted Sentiment Analysis" JCSI International Journal of Computer Science Issues, ISSN (Online): 1694-0814, Vol. 9, Issue 4, No 3, July 2012

[4] Jelena Luketina and Nantas Nardelli and Gregory Farquhar and Jakob Foerster and Jacob Andreas and Edward Grefenstette and Shimon Whiteson and Tim Rocktäschel , "A Survey of Reinforcement Learning Informed by Natural Language", eprint=1906.03926, 2019

[5] Kanchan M. Tarwani, Swathi Edem, "Survey on Recurrent Neural Network in Natural Language Processing. International Journal of Engineering Trends and Technology (IJETT) , V48(6),301-304 June 2017.

[6] Yunhui Guo , Honghui Shi , Abhishek Kumar , Kristen Grauman , Tajana Rosing , Rogerio Feris "SpotTune: Transfer Learning through Adaptive Fine-tuning" , Nov 21, 2018

[7] Meghana Ashok, Swathi Rajanna, Pradnyesh Vineet Joshi, Sowmya Kamath S "A personalized recommender system using Machine Learning based Sentiment Analysis over social data" IEEE Students' Conference on Electrical, Electronics and Computer Science (SCEECS) , 2016

[8] Anusha S, N Vignesh Karthik, Sampada K S "Comparative Study on Voice Based Chat Bots", INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING 6(12):172-175 , December 2018

[9] Rambocas, Meena, and João Gama. Marketing research: The role of sentiment analysis. No. 489. Universidade do Porto, Faculdade de Economia do Porto, 2013.

[10] Jagdale, Rajkumar S., Vishal S. Shirsat, and Sachin N. Deshmukh. "Sentiment analysis on product reviews using machine learning techniques." Cognitive Informatics and Soft Computing. Springer, Singapore, 2019. 639-647.

A study on Cybercrime – Identity, Fraud and Theft with special reference to logistics management

Ms. Vidya Bhat
Assistant Professor
Besant Women's College
Mangalore
9916868757
Email:vidyaganeshbhat123@gmail.com

Ms. Shaila Kamath
Assistant Professor
Besant Women's College
Mangalore
9844678768
Email:shazkamath68@gmail.com

Abstract: Cybercrime is a series of criminal activity concerning the digital network device using computer. Cybercrime attacks can commence wherever there is digital data, opportunity, bad intentions and motive. The cyber criminals may be a single user or sometimes engaged by a country to spy the progressive data. Cybercrime includes a series of crimes committed using internet and connected devices like laptop, Smartphone, Android television, games etc., This new venture of criminals has become all the rage due to extensive information technology and corrupt mind of some fugitive. It is disseminated through some virus which is intentionally designed to smash the data and may cause extensive damage to the server. This activity may be done without their knowledge to gain personal information for any financial benefits. They target computers using malicious software and exploit data and disable them. This technology driven crime has become a threat and continues to rise.

Keywords: Cybercrime, Financial, Software, Malicious, Technology

Cybercrime is a digital theft committed with the help of network to engender profit. The intentions may be to cause harm to an individual, group, nation or sometimes reputation of a person. This procedure involves creating a virus like Malware to ruin the data, steal space disc, admission to personal information, destroy data on the computer or convey the information out of the personal contact. They may even extend to illegal ends such as fraud, pilfering identity etc.,

Cybercrime is an activity which is a bane to the society where humans carry out those activities where the returns are high, risk of loss is low and the effort needed to be put also is very less compared to the returns gained. When a person uses his skill and earns huge profits, he is ready to take the risk at any extent. Criminals come up with new techniques and methods in order to cope up with the changing technology.

In the present world, people of all ages that are from the young age to the old age depend on digital technology. Kids use digital technology for playing games, teenagers use for recreational purposes like chatting and browsing and adults try to make life easier and comfortable. Almost all activities such as banking transactions and product purchase have been done through digital technology. Cybercrime has become a threat to the individuals, business and community as a whole. It has been a crime in priority for the government to tackle this issue.

Cybercrime against people can be done by collecting the personal details of the customers through fake means. Hacking is a very expensive cybercrime which involves millions of expenses. Cybercriminals want to become rich by doing little so they focus on rich people with huge business and large turnover.

Types of Cyber Crime:

1. Hacking: It is a kind of cybercrime where illegal instructions are sent to a person through which his personal information is accessed.
2. Child pornography and abuse: it is a cybercrime technique which is carried on by the criminals by sexually abusing the minor children through chat rooms.
3. Piracy or theft: This is a kind of cybercrime where a person violates the copyright and downloads movies, videos and music.
4. Identity theft: This is a common method problem faced by the people who are using internet for banking services and cash transactions. Criminals usually try to access data of a person's bank account, debit card, credit card and social security etc. This personal information is accessed to use persons account by withdrawing money or purchasing products in his name. This will be a huge loss to the victim as he loses his funds from the

account and even his credit history will spoil.

5. Computer vandalism: It is a type of cybercrime where the hard drive data is erased or login credentials are extracted in order to disrupt a flourishing business.
6. Malicious software: It is a network or software which gets access into a system with an intention of causing damage to the existing software and stealing the personal information of the victim.
7. Fraud calls or Email: This is a type of crime where a person receives calls or messages from the criminal as a bank employee. He tries to contact and collect personal details of the victim such as account number, password, OTP etc so that the criminal can access the account and withdraw all amount available in the account. It is also called as vishing or voice phishing.

Ways to tackle cybercrime:

1. Strong password: It is always better to use different passwords and user name for each account. Weak password can be easily hacked so always it is advisable to have a strong password.
2. Be social media savvy: Always ensure that the social networking profiles such as face book twitter etc is set to private.
3. Secure mobile devices: It is advisable to download applications from trusted sources and ensure that antivirus is installed in the phone or system.
4. Protecting data: The more sensitive details such as financial records and tax returns should be protected by using encryption.
5. Protect your identity online: Person should be very careful when giving his personal details such as name, address, phone number and financial information on internet. He should even make purchases from the trusted websites.
6. Protect your computer with security software: It is always advisable to install antivirus programs and firewall. Firewall is security software which controls the inflow and outflow of information. It watches the data and tries to distinguish between the safe and dangerous data.

Review of literature:

The issue of cyber security is not new but rather has developed more than a half century. The arrest of an East German spy in IBM's German by West Germany's police in 1968 was acknowledged as the first case of cyber espionage (Warner, 2012, p. 784). In 1983, high school student that was inspired by WarGames movie and called their selves as 414s successfully got inside the unclassified military networks (Ibid, p. 787)

Ten years ago, "the first real war in cyberspace" attacked Estonia and put the country into "a national security situation" (Hansen and Niessenbaum, 2010, p. 1168).

Our interpretation of cyber security will not be only informed by what we perceive to be the most significant to our daily life, but also by the view of the government and other prominent actors. The interplay of political expression to the variety of cyber threat (Cavelty, 2013, p. 105) is one of the reasons why it is difficult to approach cyber security issue.

Dewar (2014) explains that 'the goal of cyber security is to enable operations in cyberspace free from the risk of physical or digital harm' (p. 18). How country perceive the accumulation of interplays within securitization elements in cyber security issue and the attribution problem makes their cyber security strategy and policy are different each other.

Dunn-Cavelty notes that the securitization actors in cyber security are not only government as visible elite actors, but also nongovernment as less visible actors (Ibid, p. 118). She argues that these actors shape 'a reservoir of acceptable threat representations' that affects the cyber security practice (Ibid, p. 115)

Objectives:

1. To study the awareness about cyber crime and the victim

2. To study various precautions taken by user while using Internet

Scope: The study is subjected in Mangalore. The study throws light on the crimes committed using technological, and how to overcome it.

Methodology: The information presented in the research has been gathered from Primary and Secondary data. In order to collect the primary data, questionnaire was prepared to derive details on what is the opinion of technology users in Mangalore. The Secondary information has been collected through various articles from the internet.

Sample size: 50

Limitations:

- Demographic constraint is the main issue. The area of study is limited only in Mangalore
- Some Sections of the society are unaware of Cybercrime
- This is a broader topic and requires in depth study, but time period was limited.

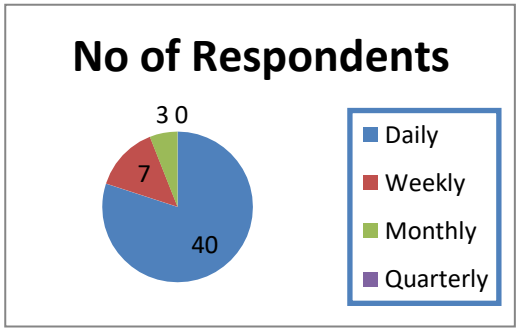
Table 1

Table showing how regularly the respondents use Internet.

Particulars	No of Respondents	Percentage (%)
Daily	40	80
Weekly	7	14
Monthly	3	6
Quarterly	0	0
Total	50	100

Chart 1

Chart showing how regularly the respondents use Internet.



Interpretation:

It is seen that 80% of the respondents use internet on a daily basis, 14% of the respondents use internet once in a week and 6% of the respondents use once in a month.

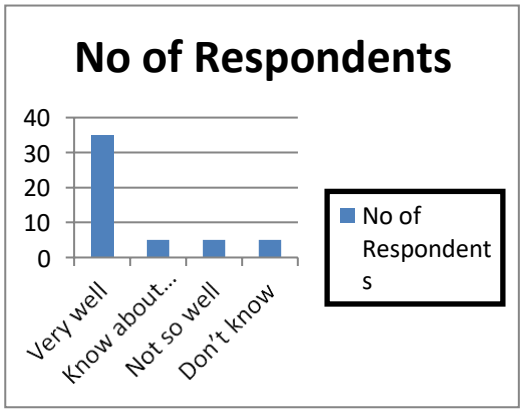
Table 2.

Table showing the awareness of the respondents about Cyber Crime.

Particulars	No of Respondents	Percentage (%)
Very well	35	70
Know about it	5	10
Not so well	5	10
Don't know	5	10
Total	50	100

Chart 2.

Chart showing the awareness of the respondents about Cyber Crime.



ISBN: 978-81-944670-1-4 | November, 2020

Interpretation:

In the above chart it is seen that 70% of the respondents know very well about the cybercrime, 10% of the respondents know about cybercrime, 10% of the respondents don't know much about cybercrime and 10% of the respondents don't know anything about cybercrime.

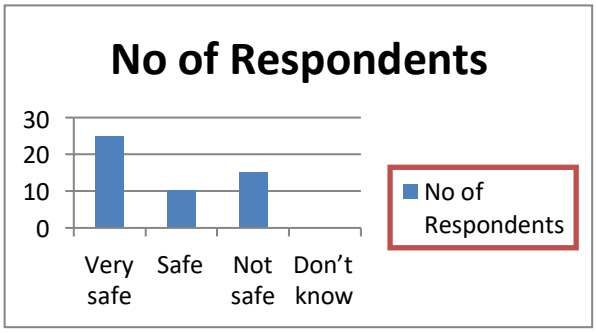
Table 3

Table showing the extent of safety the respondents feel about their information when online.

Particulars	No of Respondents	Percentage (%)
Very safe	25	50
Safe	10	20
Not safe	15	30
Don't know	0	0
Total	50	100

Chart 3

Chart showing the extent of safety the respondents feel about their information when online.



Interpretation:

50% of the respondents feel very safe about their information when online, 20% of the respondents feel safe about their information and 30% of the respondents do not feel safe about their information when online.

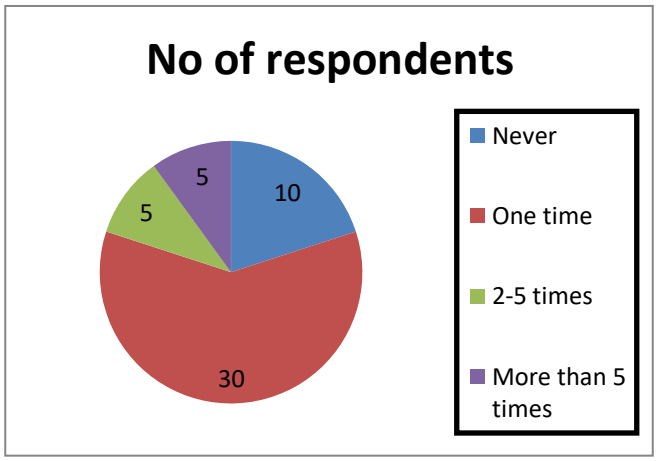
Table 4

Table showing how many times the respondents have been the victim of cybercrime.

Particulars	No of respondents	Percentage (%)
Never	10	20
One time	30	60
2-5 times	5	10
More than 5 times	5	10
Total	50	100

Chart 4

Chart showing how many times the respondents have been the victim of cyber crime.



Interpretation:

20% of the respondents never became the victim of cybercrime, 60% of the respondents once became the victim of cybercrime, 10% of the respondents became victim of cybercrime for about 2-5 times and 10% of the respondents became victim for more than 5 times.

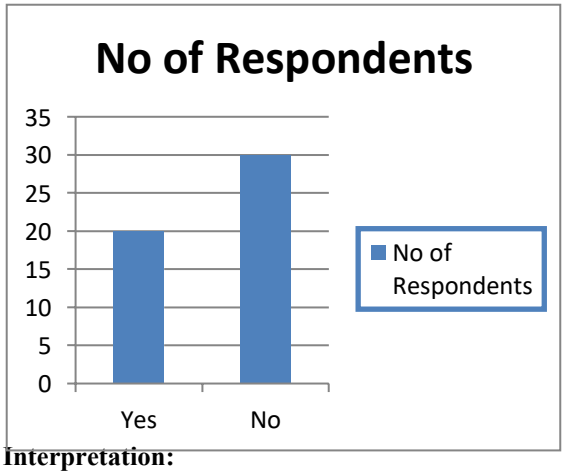
Table 5

Table showing whether the respondents have faced any cyber frauds such as social media account hacking.

Particulars	No of Respondents	Percentage (%)
Yes	20	40
No	30	60
Total	50	100

Chart 5

Chart showing whether the respondents have faced any cyber frauds such as social media account hacking.



Interpretation:

40% of the respondents have faced cyber fraudssuchas social media account hacking and 60% of the respondents have not faced cyber frauds such as social media account hacking.

Conclusion:

Cybercrimes occur at the back of the computer. Not everyone is a victim, but it's still risky. The hacker might be some thousand miles apart from the victim. The technology has paved way not to physically rob the bank, but have everything on the tips of their fingers. They need not use any lethal weapons, just everything on their fingertips. It is a treacherous offence to invade someone's privacy. On the other side, technology users should be alerted with proper guidelines like using unique password, using VPN, to update the software at regular intervals, keep the information.

References:

[1] Dewar, R. 2014. 'the Triptych of Cyber Security: A Classification of Active Cyber Defense'. 6th International Conference on Cyber Security

[2] Dunn-Cavelty, M. 2010. 'Cyber Security' in A. Collins, Contemporary Security Studies. Oxford: OUP

[3] Cavelty, M. 2013. From Cyber-Bombs to Political fallout: threat Representations with an impact in Cyber-Security Discourse. International Studies Review, 15, pp. 105-122

[4] Hansen, L. and Niessanbaum, H. 2009. Digital Disaster, Cyber Security, and the Copenhagen School. International Studies Quarterly, 53, pp. 1155-1175

[5] Warner, M. 2012. Cybersecurity: A Pre-history. Intelligence and National Security, 27 (5), pp. 781-799

A Comprehensive and Integrated Approach to Smart and Secure Remote Public Voting System

Vinayachandra
Research Scholar College of Computer Science and Information Science, Srinivas University, Mangalore, India
veeciashu@gmail.com

***Abstract*—The foundation of democracy is elections. It is through elections that people assert their voice, their opinions, and choose a person whose philosophies complement them most. Elections allow voters to select their representatives and for the people, elections are also necessary to express their anger at the ruling government. The election process is deemed successful when there is high voter participation. Sadly, countries like India are facing the problem of low voter turnout. People in rural areas move to major cities to make their livelihood, schooling, and socio-economic commitments. During elections, because of their dedication to their workplace, visiting their home city, designated polling station, and casting votes become burdensome for them. Moreover, challenges in getting to polling places and often poor weather conditions cause individuals to miss the opportunity to choose the candidate of their preference. People with disabilities and senior citizens, due to travel-related difficulties, also refuse to go to the polling station. So, many general elections record lower electoral turnout. Only when equal opportunities are given to all eligible voters to exercise their franchise such a democratic structure is considered truly democratic. The need for the hour is a feasible framework that allows all eligible voters to cast their vote by visiting any of the polling stations open across the country. This paper proposes a suitable structure that addresses most of the issues related to free, fair, transparent, secure, inclusive smart remote voting. Emerging technologies such as IoT, Cloud Computing, Edge Computing, Blockchain, and Data Analytics will be adapted to build the architecture. The only theoretical model of the system is given here, which is open for further study and development.**

Keywords—Internet Voting, Remote Voting, IoT, Blockchain, Security, Framework.

I. INTRODUCTION

India is considered the world's largest democracy with over 130 billion population and over 91 billion eligible voters. True to its spirit of democracy, the legislature and executive powers are decentralized and distributed to different administrative settings, from central to local governance. Parliament at the helm is supported by State Assemblies, Municipal Corporations, Zilla Panchayats, Taluk Panchayats, Town Municipal Councils, and Gram Panchayats. Representatives of these different administrative institutions are elected through the general election process. Though the term of these elected bodies is either five or six years, mid-term elections may necessitate on several occasions due to political or administrative reasons. Thus, in general, one in any part of India may witness major elections every six months. Because of the logistics involved in the process, it is a herculean task for the Election Commission to coordinate the election process. According to the current framework, the Commission is opening Polling Stations in designated places based on the number of eligible voters, where the voters must come and vote. All the voters must physically come to the polling station to exercise their franchise, other than the people participating in the voting process, voters belonging to certain age groups, and the state serving personnel; they have the option of postal ballot. General voters have no other option excluded from exercising their franchise. Most of the time, voters may be working in different locations of the country, or they may be traveling, or they may be studying in faraway places, losing their chances of voting because it's not convenient for them to come to the polling stations and cast vote. Also, there may be a fear of getting bullied into voting for someone who is not of their choice. The reasons may be administrative or economic [1].

As technology usage increases every passing day, transaction security is needed more. Secure Internet-based transactions include online shopping, banking, tax payment to payment of installments, and license renewal for vehicle insurance. The Internet can also be used in a certain way to allow secure elections. Many classes of people, including military personnel, overseas citizens, businessmen, physically challenged people, elderly people, sportsmen, college students, etc. will benefit from the chance to vote from anywhere. Due to its remarkable limitations, conventional paper-based ballots have become outmoded. The e-voting system is convenient as it is accurate, faster, and requires less labor compared to printed ballots. But, apart from the performance aspect, several times the implementation of e-voting machines was not as convinced as expected. The system is to be commissioned with limited control and operated by people with limited technical expertise.

Utilizing technology in voting procedures can make it quicker, more efficient, and less susceptible to security breaches. The technology can ensure the safety of every vote, better and faster and much more accurate counting & automatic tallying. The process uses minimum paper documents and is therefore environmentally friendly. But there are issues of concern als [2]. The e-voting system is vulnerable to several serious attacks from external sources. There is indeed a likelihood that anybody who has immediate access to the e-voting system can access it suspiciously. Malevolent software can steal one candidate's votes and assign them to some other. An attacker may deny officials access to the e-voting system or render an e-voting system unavailable for the Election Day voting process. This is known as a Denial of Service (DoS) attack. Such an attack is extremely complicated to identify. The standardized, simplified, adaptable system that will ensure safe and secure remote public voting is essential to formulate. This study would make a sincere effort to propose an ideal framework that addresses most of the issues related to free, equal, open, safe, inclusive, and intelligent elections[3].

II. OBJECTIVES

- To specify the need for a Remote Public Voting System
- To list the advantages and challenges of Remote Voting

 ISBN: 978-81-944670-1-4 | November, 2020

- To summarize technologies supportive for the implementation of Smart and Secure Remote Voting framework
- To recommend an ideal framework that will provide a smart and secure remote public voting system in the form of Kiosk

III. BACKGROUND & MOTIVATION

A considerable amount of research is being done on the use of technology and frameworks in the electronic voting system (EVS) in general and in remote voting in specific. The opinions or methods of system management have been widely discussed, but there is inadequate emphasis on trustworthiness. Much of the study, however, focused on the voting system and only technological aspects were taken into consideration. As these parameters are considered to be paramount, the authors must be concerned about authenticating voters, preserving voter data, and securing the voting process. The standardized, simplified, adaptable system that will ensure safe and secure remote public voting is essential to formulate. Much of the study, however, centered on the voting system and only the technological aspects were taken into account. This work is being aimed to fill the void.

IV. VOTING SYSTEM

The voting system involves three processes: registration of voters, casting of votes, and tallying of votes. Some requirements are imposed by national legislation to determine the eligibility of voters. Identified voters are then registered and issued with a specific electoral id attached to the card. The database of voters is stored in a very safe way. Vote casting is the second phase involved. This includes issuing paper ballots or electronic ballots and allowing voters to cast their votes upon identity confirmation. Finally, the votes cast are collected, counted and the result is communicated to the party concerned. The three-step election procedure is shown below in Fig.1.

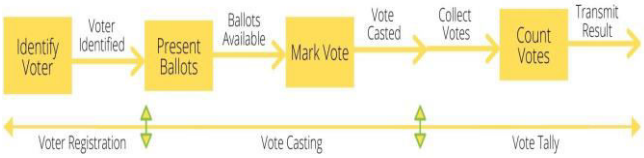


Fig. 1. Three-step election procedure

In paper ballot mode and electronic mode, the mechanism involved in the election is mostly identical. The only significant difference is that many of the works involved in the process are done manually in paper ballot mode and they are significantly reduced in the case of EVS [4].

The different options of polling station-based and remote voting systems are listed in Table-I below [5].

TABLE I. DIFFERENT OPTION OF VOTINGS

Type	Non-Electronic	Electronic
Polling station based	Paper ballots in the specified polling station	Electronic voting machines in the specified polling station
Remote	Mail voting	Internet Voting
	Proxy voting	E-Mail voting
	Paper ballots in the distance polling station	App-based Mobile voting
	Mobile ballot box	Electronic Voting machines in distance polling stations
	Paper ballots in special polling stations	Electronic voting machines in specified polling station interconnected via the Internet
	Paper ballots in a polling station outside the voter's constituency	Online polling stations

V. METHODOLOGY

Remote voting applies to all those methods that allow voters to vote, either from abroad or from within the country, from locations other than the polling station allocated to their place of residence. It comprises both electronic voting and non-digital voting mechanism. Electronic voting is voting that uses electronic means to either aid or take care of casting and counting votes. E-voting can use standalone electronic voting machines or systems connected to the Internet, depending on the specific implementation. Voting on the Internet can cover a range of Internet services, from the simple transmission of tabulated results to full-function online voting. The degree of automation may be limited to a paper ballot's digital marking, or maybe a comprehensive system of voter authentication, voting input, voting recording, data encryption and transfer to servers, and election results consolidation and tabulation. Most of these tasks must be performed by a worthy EVS while complying with a set of criteria set by regulatory bodies and must also be capable of successfully addressing strong protection, precision, fairness, swiftness, safety, auditability, accessibility, cost-effectiveness, scalability, and ecological sustainability requirements.

There three major categories of electronic voting:

- Voting system that is physically supervised by government representatives or autonomous electoral authorities
- Web-based remote electronic voting (also referred to as Internet voting) where the voter needs to submit his or her vote digitally from any location to the electoral authorities.
- The hybrid system, the combination of the two – supervised remote electronic voting via the Internet [6].

A system suitable for supervised remote electronic voting through the Internet aid is proposed in this work. For the purposes of contributions numerous research articles published in peer-reviewed journals, papers published in conference proceedings, chapters published in edited books, articles published on websites, and official manuals related to national & state election process and procedures were examined to identify different issues relating to remote voting. An attempt is made to learn about various technologies used to build a smart and stable remote voting system consisting of many emerging technologies such as the

Internet of Things (IoT), smart sensors, cloud networks, blockchain, etc. [7].

VI. SYSEM DEVELOPMENT PROCESS

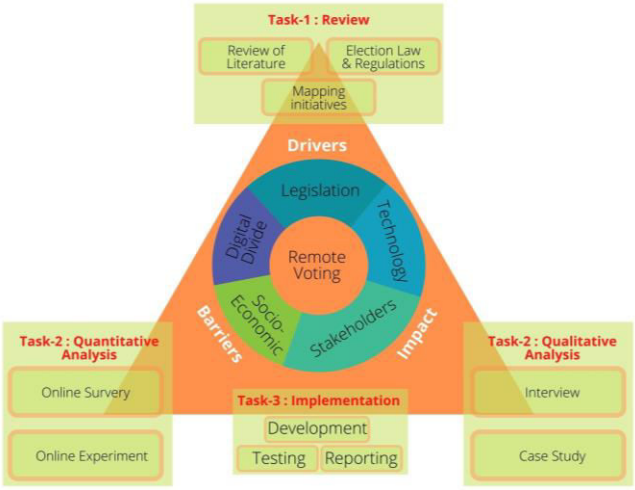


Fig. 2. Framework Development Process

As shown in Fig. 2 above, the entire work of framework development is split into four tasks. Task-1: articles published in peer-reviewed journals and chapters of books published on the subject in edited books are reviewed to find out the progress recorded to date. The national election process, regulations, and laws are also studied. Besides, initiatives taken so far on the topic are also mapped. Task-2: online survey, online experiment, interview, and case study are used for qualitative and quantitative analysis. Task-3: where the framework is applied, tested, and completed with reporting. The framework is driven by legislation and technology, impacts stakeholders, and is based on barriers to socio-economic and digital divides [8].

VII. PROPOSED FRAMEWORK

The proposed Smart and Secure Remote Public Voting System consists of evolving technologies such as Smart Sensing and Automatic Functioning Internet of Things and Edge Computing, Cloud Framework for Remote Processing and Data Storage, Blockchain for the safety and security of Sensitive Information, Internet as a Network Infrastructure, and many additional elements. The reason for implementing such a relatively involved remote voting system rests heavily on voters' perception of transparency, security, and inherently simple insider attacks [10].

The system should provide much stronger protections against viruses and intruder attacks than other systems proposed, in addition to offering a mechanism that a voter would trust. This helps voters to cast their votes from the remote place where they live, where they are working, where they are learning, and where they are getting training or treatment rather than driving to their designated Polling Station wasting precious time, resources, and energy. This also guarantees more voter turnout as it provides a fair opportunity for the eligible voters to exercise their franchise [11].

The motivation behind proposing this framework is to address the following requirements of the election process.

- ***Accuracy*:** The system must ensure that none of the voting parties will change any of the votes cast. It must prohibit anyone from deleting valid votes from the final count and also the invalid votes must not be considered in the final count. It should also be remembered that the steps taken to prevent hostile parties from muddling the mechanism do not interfere with the privacy of the electorate [12]
- ***Equality*:** The system should allow only eligible voters to vote; this means the system must ensure that only registered citizens can vote, recognizing eligibility is tested during the registration process and every registered voter can only vote once and that every vote is weighted equally [12].
- ***Privacy*:** Ensures the confidentiality of ballot contents by encrypting the ballots using a specific cryptographic technique. The method preserves privacy by preventing either the election authorities or someone else who can connect any ballot to the voter who casts it, and by not allowing the elector to show that he/she has voted in a specific way. [13].
- ***Verifiability:*** By some means, if it is possible to check whether the votes are counted correctly, then the system can be considered as verifiable. During the process of verification privacy of voters must be protected. There should be a provision to ensure the final results of count for ensuring each party that all the legitimate votes have been included in the final count.
- ***Convenience:*** A system is contemplated as a convenient one if it permits voters to cast their votes with minimal equipment or skills. This property will increase the turnout, predominantly in government elections that rely on a large number of voters, and it is not feasible to expect this huge number of voters to get into training [14].
- ***Robustness:*** The voting system must operate properly even if the system is partly failing. This must be avoided that a small coalition of electors, talliers, and other groups participating in the election to disrupt elections should not, for example, be able to write an encrypted vote in an unauthorized format, which can go unnoticed at first and then prevent the mechanism from functioning effectively.
- ***Voter Independence:*** In certain voting schemes, a voter may copy the vote of another voter without necessarily understanding the vote being copied, for example, if a vote is encrypted using the voting server's public key such encryption may be copied, such duplication of votes must be avoided by the voting system.
- ***Efficiency:*** The system must ensure that all operations are handled speedily and efficiently. Efficiency depends on the overall performance of the system, such as the number of transactions (seconds/minutes) per time and the response time to user queries [15].
- **Non-coercion of voters:** By avoiding manipulation and coercion of voters, the system must minimize

the risk of voters being forced to vote in such ways. The system must be capable of concealing voters' identification and the choices made during the vote.

- ***Fairness*:** No partial results are known before the election is closed [16].

The system includes following elements:

a. Remote access
b. Security checks
c. Encryption of data using Blockchain
d. Secured Cloud Storage

a. Remote Access: This provides a way for the eligible voters to cast their votes to their choice of candidate from a remote location. It is assumed that the Commission will install several Remote Voting Systems units in the major cities spread across India. So that whenever an election is scheduled for any of the constituency or area of India, the legitimate voter needs not to travel to his/her constituency, city, town, or village, instead, one can exercise his franchise directly by visiting one of the units installed in his/her place [17].

b. Security checks: The system ensures safety and security through Smart Voter ID Card, pre-loaded secret code, and One Time Password (OTP). Every Voter ID card consists of a unique Voter ID stored in the chip which is used to initiate the process of voting. The preliminary process generates an OTP and sends it to the registered mobile of the voter. This is used as the second-factor security [18].

c. Encryption of data using Blockchain: Once it comes to Blockchain voting, each vote will be considered equivalent to a transaction, and by using multiple Blockchains along with public key encryption, the decentralized voting process can be safeguarded while maintaining the voting process's anonymity function. The votes can be randomized in the digital ballot box more than three times in the Blockchain voting process ensuring the identity of the voters is never revealed. When the voting is closed a separate Blockchain program is installed in the digital ballot box for the counting of votes. The specific Blockchain matches the Blockchain of the public newsletter board and thus shows that the online voting system correctly worked. The Blockchain voting system combines the transaction audit trail with the public key encryption which solves the auditability problem [19].

d. Secured Cloud Storage: The cloud platform serves as ubiquitous storage for voter data. The voter database which is stored on the private cloud can be accessed anytime from anywhere. The use of md5 algorithm for encryption makes the data unintelligible to unauthorized persons. The recovery mechanism can be used to decrypt the data as and when necessary [20].

Describing the security of cloud data storage is nothing other than data protection, applications, and infrastructure involved in cloud computing. An effective type of security infrastructure will be able to identify any problems that could occur with protection and management of data and should fix the problems with protection control mechanism. The protection control should identify all sorts of vulnerabilities and fix them. They are required to protect the system against different kinds of attacks. With the help of Consensus algorithm, it becomes impossible for fraudulent nodes to connect with honest nodes. The main important advantage is that it does not require the DoS attack and also impact on mining chances from the low stake. Proof-of-work imposes such constraints or limitations in network behavior. It takes lots of potential effort to be put into action. To apply the calculations, the attacks require lots of computational resources and therefore a lot of time [21].

On the basis of requirements identified, a layered architectural framework based on block-chain technology is proposed as Fig. 3 below:

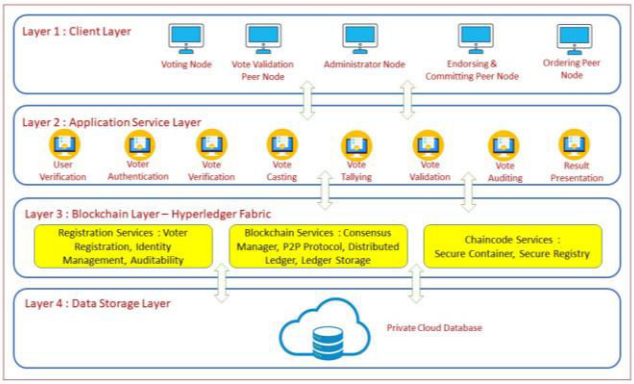


Fig. 3. Proposed Secured Remote Voting System Framework

Layer-1: Client layer This layer comprises of several electronic devices by using which uses interact with the voting system. In the blockchain infrastructure, these devices are called as peer nodes. They interact through smart contracts. In Hyperledger Fabric they are called 'chaincodes'. The peer nodes will be assigned different responsibilities such as:

- *E-Voting nodes*: These nodes are specifically used for voter identification. They also record the casting of votes as well as all blockchain transactions.
- *Administrator nodes*: For the purpose of configuration of blockchain network channels, assigning of roles and responsibilities, providing access restrictions these nodes are used.
- *Ordering nodes*: In blockchain there are some transactions which can be accessed publicly. These nodes enable the users to access public transactions related to e-voting blockchain.
- *Vote validation*: The validation of votes is the responsibility of these nodes. Added to this, they also confirm the authenticity of transaction related to the block.
- *Committing nodes*: The task of validating and committing a new block is the responsibility of these nodes.

Layer-2: Application Service Layer - It consists of many services that are available in case of an e-voting system. the type of services that can be accessed by a specific node in the blockchain are determined by the access control mechanism. This layer also takes care of various permissions for a specific node in the blockchain.

Layer-3: Blockchain Layer – This layer contains a integrated blockchain architectural framework namely Hyperledger Fabric V2.0. This is responsible for providing blockchain information systems solution. It facilitates the creation of permitted blockchain networks that have in-built functionality such as encryption and protection of privacy. The Hyperledger Fabric includes "ordering nodes" that

guarantee blockchain integrity by ensuring that the committing peer nodes are made available to only ordered blocks of an endorsed transaction before they are added to the blockchain.

Layer-4: Data Storage Layer – This layer consists of databases that are required to store the details of registered voters. It also contains the details of candidates contesting for election. This database is the one which is used as the document for authentication and authorization of voters [12][19].

VIII. BENEFITS & CHALLENGES

Benefits

- It contributes a lot to the increase in voter turnout
- It saves money, time, and hardship involved in moving from place to place to cast a vote
- It enables hassle-free voting from anywhere
- It is fast and convenient because individuals need not have to wait in lines to get their chance for voting.
- It will be more secured and well-structured than that of the manual process of monitoring, voting, and protecting the votes.
- Using pre-recorded details voters can easily log in to the system for casting their votes.
- It uses Blockchain technology for providing greater security hence it is more reliable when compared to the existing system.
- It reduces the dependence on human resources to manage the voting process
- It makes all the processes involved in voting are made fool-proof, tamperproof, and automatic
- Greater accessibility for voters with special needs [20][21].

Challenges

- The system is only useful for remote voters. Not the voters who are residing in the area in which voting is held
- A huge initial investment is required to set up units in different locations across India
- The biometric authentication process may consume a little more time as it involves several processes
- Biometric authentication some time fail to identify even genuine voter and deny the right to vote
- It is functioning largely depends on the sophisticated network infrastructure as most of the operations are cloud-based
- As almost all the processes are happening using cloud infrastructure, there may be cause to worry about the security and safety of data [22]

IX. DISCUSSION & FUTURE WORK

For India having fairly easily adopted technology originating in the West, it is only reasonable to use the accepted technological aspects to solve real-world problems. Now, concerning India's new technical buzzword, it is felt that this is undoubtedly the Internet-of-Things (IoT), the Cloud and Blockchain, and related smart aspects. To communicate and share data with other devices and systems over the Internet, the Internet of Things introduces a new world of connected objects that are loaded with sensors, programs, and other technologies. Cloud computing can facilitate the provision of on-demand computing and storage resources [23]. Blockchain is the latest technology that has the potential to have a massive effect on different industries. Voting is one of the applications benefiting from Blockchain technology implementation. Some of the reasons why the voting system needs Blockchain to include trouble-free voting, utmost confidence for voters, greater transparency, privacy for voters, cost-effectiveness, legitimacy, factual outcome, higher security, and ease for tootling. In this work, by analyzing these three versatile technologies a conceptual model that provides a framework for remote voting safely and securely was produced. The framework is made up of four different inter-linked layers that embrace different adaptive technologies. The proposed framework gives a complete knowledge of its functioning. Smart sensors are proposed to gather user data and they are processed using edge and cloud computing techniques. To ensure security and robustness Blockchain technology is proposed. This is used to encrypt using a particular algorithm that is very hard for one to decode without the key being used. In this work, only the conceptual model of the proposed framework is presented. Its development and technicalities were not discussed. In future work, this model will be implemented with a specific algorithm, technique, and hardware inferences [24].

X. CONCLUSION

Although remote voting has been used in only a few countries for national-level elections, It is a method of voting which is progressively getting explored as a way to allow voters who would otherwise find it hard to use the election process at their polling place on the day of elections. Remote voting, however, poses a range of technical challenges based on issues of protection, privacy, and confidentiality, as well as challenges to stakeholder participation and process observation. All of these must be comprehensively addressed for moving forward. In this proposed work, the possibility of kiosk-based smart and safe remote voting will be studied & analyzed and a suitable framework will be recommended. The proposed work will help every rightful voter given a reasonable opportunity to cast his/her franchise and indirectly contribute to the growth of the nation as well as his / her own.

REFERENCES

[1] Bhuyan, Dip Jyoti(2019). Effectiveness of an electronic voting machine in the electoral system of India: New opportunities and challenges. International Journal of Recent Technology and Engineering, 8(2), 192-199.

[2] Salami, H. J., Adebayo, O. S., Isah, A. O., Lawal, K. H., & Alhassan, J. K. (2019). Development of a Secured E-Voting System With OTP as Second Order Authentication. i-Manager's Journal on Software Engineering, 13(3), 7-14.

[3] Ujir, H., Sing, L. C., & Hipiny, I. (2014). A modular approach and voting scheme on 3D face recognition. 2014 International Symposium on Intelligent Signal Processing and Communication Systems, ISPACS 2014, 196–199.

[4] Shital A, P., & Praveen G, K. (2015). IRIS Detection in Voting System. International Journal for Research in Applied Science & Engineering Technology (IJRASET), 3(8), 469–473.

[5] Dixit, P. V., Phalke, S., Ubale, R., Gavali, A.B, Prajkta, S., & Aparna, S. (2015). A Biometric-Secure E-Voting System for Election Process. International Journal of Advanced Engineering and Global Technology, 3(3), 425–430.

[6] Nithya, J , Abinaya, G, Sankareswari, B., & Saravana Lakshmi, M. (2015). Iris recognition-based voting system. International Conference on Science, Technology, Engineering & Management, 44–51.

[7] Nithya, S., Ashwin, C., Karthikeyan, C., & Ajith kumar, M. (2016). Advanced Secure Voting System with IoT. International Journal Of Engineering And Computer Science, 5(3), 16033–16037.

[8] Bindia, & Aggarwal, N. (2016). NEXT GENERATION HI-TECH E-VOTING TECHNIQUES IN INDIA. International Journal of Science, Engineering and Technology Research (IJSETR), 5(1), 228–233.

[9] Khoury, D., Kfoury, E. F., Kassem, A., & Harb, H. (2018). Decentralized voting platform based on ethereum blockchain. In 2018 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET).1-6.

[10] Patil, S., Bansal, A., Raina, U., Pujari, V., & Kumar, R. (2018). E-Smart Voting System with Secure Data Identification Using Cryptography. In 2018 3rd International Conference for Convergence in Technology (I2CT)

[11] Snega, S., Saundarya, S., & Balraj, R. (2018). Highly secured electronic voting machine using aadhaar in IOT platform. International Journal of Electrical and Electronics Research, 6(2), 41-47. ISSN 2348-6988

[12] Alam, A., Zia Ur Rashid, S. M., Abdus Salam, M., & Islam, A. (2018). Towards Blockchain-Based E-voting System. 2018 International Conference on Innovations in Science, Engineering and Technology, ICISET 2018, 351–354.

[13] Yi, H. (2019). Securing e-voting based on blockchain in P2P network. EURASIP Journal on Wireless Communications and Networking, 2019(1), 1-9.

[14] Shanmugasundaram, G., Kalaimathy, A., Johnvee, M., & Pavithra, S. (2019). Perspective Analysis of Digital Voting Systems. In 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN).1-6.

[15] Mohan, M. Madhu, M. Prakash, M. Madhuseelan, and A. Kishore Kumar(2020). "Design of Secured Biometric Voting Machine. International Journal of Research in Engineering, Science and Management, 3(3),199-201

[16] Sathya, V., Sarkar, A., Paul, A., & Mishra, S. (2019). Block Chain Based Cloud Computing Model on EVM Transactions for Secure Voting. In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC) (pp. 1075-1079). IEEE.

[17] Jagtap, A. M., Kesarkar, V., & Supekar, A. (2019). Electronic Voting System using Biometrics, Raspberry Pi and TFT module. In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI). 977-982. IEEE.

[18] Rubin, A., & Park, F. (2019). Security Considerations for Remote Electronic Voting over the Internet. Communications Policy and Information Technology.

[19] Sadia, K., Masuduzzaman, M., Paul, R. K., & Islam, A. (2019). Blockchain Based Secured E-voting by Using the Assistance of Smart Contract. Springer IETE International Conference on Blockchain Technology (IC-BCT 2019).

[20] Akhtar, S. J., & Limkar, M. B. (2014). a Biometric-Secure Cloud-Based E-Voting System for Election Processes. International Journal of Electrical and Electronics Engineering Research (IJEEER), 4(2), 145–152.

[21] Komatineni, S., & Lingala, G. (2020). Secured E-voting System Using Two-factor Biometric Authentication. In 2020 Fourth International Conference on Computing Methodologies and Communication (ICCMC) (pp. 245-248).

[22] Olumide S., A., Olutayo K., B., & E. Adekunle, S. (2020). A Review of Electronic Voting Systems: Strategy for a Novel. International Journal of Information Engineering and Electronic Business, 12(1), 19–29.

[23] Kavitha, S. N., Shahila, K., & Kumar, S. P. (2018). Biometrics Secured Voting System with Finger Print, Face and Iris Verification. In 2018 Second International Conference on Computing Methodologies and Communication (ICCMC) (pp. 743-746). IEEE.

[24] Prabhakaran, G., Dharshini priya, T., Janani, N., Deepan Raj, R., & Elavarasan, P. (2018). Electronic voting machine based on fingerprint and iris authentication. International Journal of Intellectual Advancements and Research in Engineering Computations, 6(1), 135–139.

The Role of Cognitive IoT in the Analysis of Student Classroom Behavior for the Enhanced Teaching-Learning Experience

Rajeshwari M
Research Scholar, CCIS, Srinivas University, Mangalore, India
Assistant Professor, Department of Computer Science, St Philomena College
Puttur, India
Orcid ID: 0000-0001-9613-4967; E-mail: rajimuraleedhar@gmail.com

***Abstract*—A classroom is a teaching and learning space. Typically, a common classroom has a blackboard, desks, and seats. The systematic use of information technology will increase the participation of students; improve access, and resultant effect more convenient. Smart data gathered by facilities installed in the classroom teaching, researchers, and teachers can assess learning and teaching performance. Collection of data on students in education inside the classroom includes their interest, web-based digital monitoring, video surveillance data, and IoT data to study and monitor their behavior. A traditional classroom can be turned into a smart and intelligent classroom that actively listens to and analyses voices, interactions, gestures, actions, etc., by combining IoT technology with social and behavioral assessment, to complete the study about the performance of the lecturers and the engagement of the listeners. For all courses, the learning results are often based on achieving cognitive abilities and building a good framework for skilled and personal development for learners. This paper explores the role of IoT with added intelligence to analyze the various classroom behaviors of students while learning with an ideal framework model. This paper also discusses the challenges and benefits involved in such an intelligent classroom.**

Keywords*—*Cognitive IoT, Behavior Analysis, IoT, Education, Intelligent Classroom

I. INTRODUCTION

Today all people consider learning is imposed on the learner, and everyone is compelled to do the same things at the same frequency in the classroom the same day. But all students are unique and different. It's going too fast for some, too sluggish for some, in the wrong direction for others. The teacher's thoughts have to be fully visible to all students, while student thinking has to be visible and readable to the teacher. The characteristics of a student and stored in the institutional student information system are required to study his behavior, learning style, and intelligence. The characteristics such as age, gender, nationality of the student, language spoken at home and known, medium of language he completed his earlier studies, home remoteness (urban/rural) are essential to track student behavior in the classroom. Today, almost all higher education institutions live with one-size-fits-all, face-to-face teaching pedagogical styles of teachers and students [1].

Thoughts, values, emotions, experiences, and needs all have a profound effect on human behavior and contribute significantly to individual behavioral variability through contexts and individuals. Behaviour Analysis that follows a solely behavioral approach will provide little insight into the perception of the user's thoughts and cognitive behaviors, his/her motives, perceptions, values, emotions, and intentions that drive and inspire behavior. Cognitive models of human behavior are interested in explaining the mental processes and mental operations that people use when thinking, recalling, learning, or communicating with their living and non-living environment.

Teachers have an important role to play in facilitating student achievement in the classroom. The act of teaching is however a dynamic and multifaceted practice. Student success in the classroom can rely on different variables such as socioeconomic status, teaching strategies, classroom atmosphere, student-teacher relationships, student involvement, etc.

Without our realization, many of the devices around us collect information about us: items embedded in objects, mounted on human bodies, controlled by sensors, engineered for intelligence, and constructed for the capacity and communication of information gathering. All these things interact, make decisions, and share information across the cloud [2]. Various methods and techniques are now available for automated tracking and analysis of the data collected on student activities. Trying to unlock the full potential of IoT in education with tools will increase learning outcomes and make a substantial investment. IoT's heterogeneous devices can gather data from multiple sources to create a smart, automated environment using intelligent technologies such as cloud computing, learning analytics, and cognitive computing. These systems can collect high-volume data from a variety of sources, interpret them, and make decisions of their own, like humans. Cognitive IoT is indeed an IoT system that combines cognitive computing technologies such as Artificial Intelligence (AI), Machine Learning (ML), Natural Language Processing (NLP), and Big Data Analytics with data generated by connected devices and actions that these devices can perform to provide automated and intelligent human-machine interaction in any user application.

II. METHODOLOGY

The study is carried out by the use of secondary data in academic journals, books, papers, chapters, and conference proceedings. The relevant data are obtained and conceptually examined through a proper analysis of the articles. The own

theoretical model for student behavior analysis is suggested with the use of available smart classroom tools. The model is clarified in theoretical terms by its structure and its advantages.

III. Related Study

Intelligent ambient is developed using IoT to provide automated feedback on the effectiveness of lectures focused on certain metrics. A smart classroom is created with sensing and monitoring technology to explore learner behavior. This lets the instructor observe the student's reaction to the class to increase the standard of his teaching. (Gligorić, *et al* 2012) [3].

Sohsten and Murilo, in 2014 [4] suggested performance assessment between such a multi-face recognition system on a close device with Emgu CV and a variant associated with Windows Azure, calculating the number of face detection and ranking frames processed in real-time. A higher number of processed faces is better in terms of reliability.

Smart Objects (SO) include heart rate test monitors, smart lighting objects, smart locks, power meters, and different types of sensors. It is projected that the SO numbers that can be linked using IoT technology will hit 224 billion by 2022. (Al-Fuqaha et al, 2015) [5] (Manyika *et al* (2015) [6].

Hong in 2017 [7] suggested a device consisting of several wearable sensors based on the Bluetooth approach for tracking behavioral motion patterns of children with autism. It also utilized microphone cameras to capture videos and photographs for the execution of these cases.

Taking part in the classroom is a time-consuming and repetitive process. Also, students are present physically, but the problem of attending the class would not be attentive. It is monitored by an IoT solution built on the Embedded Linux board called Raspberry Pi. This collects the student image then stores it on the cloud and has been analysed periodically with the aid of the Face Recognition API. (Patil & Sachapara, 2017) [8].

Effective functions are cognitive mechanisms that enable behavioral learning to meet the objective of maturation and environmental simulation. The author studies the contact and relationship between the student and the teacher. Teachers will aim to promote the student's optimistic approach to emotional learning and cognitive processes that are important to teaching in the classroom. Teachers study the student's personality, emotions, aspirations, and expectations. This helps the teacher increase teaching performance by engaging students in conversations and classroom actions. Additional input and expertise make teaching perfect for students to achieve better learning outcomes. (Vandenbroucke *et.al* 2018) [9].

The author used a ready-to-use IoT platform with heterogeneous devices that host learning materials for customized learning. Sensing systems can capture the actions and manner in which materials are managed during contact with them. It is implemented in the form of a game enabled on mobile devices and robots. Here, pieces of information usually called atoms for the learning target to be accomplished by the student are set by the teacher. The sequence of learning activities decided by the instructor varies from that of each student. This personalization may represent a linear graph that shows the difficulties involved in learning behavior. (Spyrou *et al* 2018) [10].

Bahreini, *et.al*, in 2019 [11] Built facial expression recognition applications for emotions. In addition to allowing the identification of emotions through webcam image documents and recorded video records, facial expressions are used in real-time and continuously. It uses the FURIA algorithm to provide timely and relevant input, depending on the facial emotions of the untried fuzzy rule testing cases. The primary objective of this research was first to validate the use of webcam data for real-time and precise assessment of facial expressions in learning situations.

The IoT model includes the Brain Library in Arduino and Mindflex for managing hardware devices with no prior knowledge or experience makes it easy to use. Thus these communication models can effectively read and perceive a person's state of mind from a distance. Therefore IoT and this module have also been used together in long-distance learning. (Padhi, A., *et.al*, 2019) [12].

IV. Proposed Model

A new conceptual framework is introduced here, focused on student behavior in classroom teaching. Students are equipped with smart chairs, laptops, smartphones, CCTV monitoring, voice, and video recognition apps, and smart devices in the classroom. The Internet of Things is made up of heterogeneous devices, where data originates from a variety of classroom sources. The data generated by each object is therefore recorded, stored, and properly handled. The involvement of students in their learning process is important in classroom teaching. It is sustained at three levels: cognitive, emotional, and behavioral levels. Here all three stages are interrelated in their attributes to enhance the learning experience. Without cognitive and emotional participation, students may not display positive conduct in the classroom while studying. Then in higher education, students take the subjects as their duty to write the exams and finish the course without knowledge. Therefore, focus and behavior in classroom activities are needed to enhance motivation and knowledge skills. This good commitment to the classroom is the responsibility of the teacher. He/she must prepare teaching plans, techniques, course materials, and approaches to make them more interactive in their learning process. On the other hand, smart technology, such as cloud computing, ML, learning analytics, concentrate on how learning data can be collected, analysed to enhance learning and teacher decision-making skills. The new smart classrooms are helping in this direction.

Behavior can be conveyed by bodily signals in real life. These signals can be classified into two major types: behavioral signals such as facial expressions, hand and body movements, speech or written text, and physiological signals such as changes in heart rate, pulse, breathing, eye gaze, pupil size, skin temperature, etc. Students can display different behavioral signals on different types of teaching techniques, tools, resources, and their contents, subjects, etc. At the very same time, physiological signals are not conveyed and

displayed explicitly by their body parts but are interpreted by their behaviors and participation in the classroom. Here, therefore, teachers use technologies to evaluate all forms of student behaviors to enhance teaching methodologies inside the classroom. Fig. 1 below demonstrates a system that uses IoT devices with cognitive computing technology to evaluate student behavior in the classroom teaching-learning process.

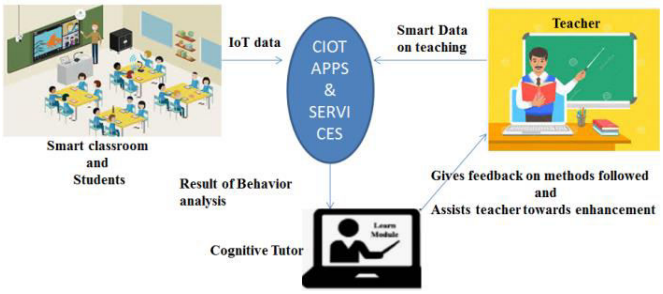


Fig. 1: Proposed CIoT Model used inside the smart classroom

The proposed CIoT model continuously monitors IoT data from the smart classroom to create smart knowledge and decision-making on student actions while learning. The activity was observed by tracking human facial emotions through live video capture using an IoT-enabled camera and image processing using facial emotion recognition software. This camera can also record every movement of the body or actions of the students inside the classroom. The Emotion values obtained have been analysed and collected in a cloud. Behavioral insights are sensed by proper wearable and non-wearable camera-based sensors to monitor learners' needs, behaviors, and reactions against teaching techniques used within the classroom. Continuous monitoring and analysis are characterized by several parameters: (a) Various teaching methods and techniques used in the classroom; (b) Subject-specific interest; (c) Schedule; (d) Disruption factors; and (e) Type of learning. Each parameter can produce a huge amount of sensitive data and provide useful insights. As a result, the majority of data is stored in the cloud-provided database and must be processed using cognitive computing technology to make meaningful decisions. This procedure is performed periodically to monitor student activity and to detect the best teaching-learning classroom practice that will improve student performance and the learning process.

a) *Methodologies and strategies used inside the classroom:* A good teacher will always seek to promote students by using proper teaching tools. Such teaching approaches used are ICT-based classes, e-resources, interactive textbooks, cognitive tools, conventional teaching methods, group studies, questions and answers, seminars, debates.

b) Subject-specific interest: Each university identifies possible courses and subjects that include language, arts, business, science, and other subjects. Students can select their course, which includes some compulsory subjects and elective subjects. But students need to study all multiple streams of subjects included in the course, whether they are interested or not.

c) *Schedule:* Timetables are always prepared for every subject, teacher, and course. The schedule of subjects would have an important role to play in creating interest in students. Each subject will have schedule issues such as a stream of subjects, scheduled time (morning, noon, evening, etc., and the instructor and these factors together will determine student behavior.

d) *Disruption factors:* Students and the environment will always be disrupted if they are not involved in learning. This would also trigger a shift in the actions of interested students. The schedule also often makes it important to cause distractions within the classroom.

e) *Type of learning:* Learning is a skill. It increases student understanding, interest, and behavior. Types of learning are also important for enhancing their enthusiasm for education and learning. Several types of learning are available, such as interaction-based learning, blended learning, cognitive learning, rote learning, receptive learning, and flipped learning.

These factors provide guidance on his methodologies, resources, and methods used in the classroom by building a new cognitive tutor to assist the physical instructor. Using this information teacher will improve his teaching actions and make students participate in their studies with better performance levels.

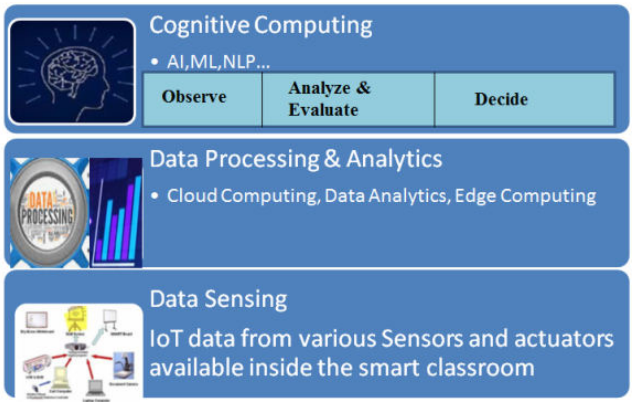


Fig. 2: Structure of CIoT apps and services

Fig. 2 shows the layout of the CIoT Apps and the solutions used by the instructor in the classroom to evaluate student behaviors. It is composed of 3 layers that cover all the technologies needed to analyse student behavior from sensing to decision-making. The functionalities of those layers are as follows:

- Data Sensing is responsible for automatically collecting raw data using IoT sensors that are installed within the classroom. Sensors may be smart wearable devices (such as smartwatches, neck-mounted smart devices, cloths, or other body parts) or non-wearable devices (such as Q-sensors, bio-sensors, etc.) built within the setting. Usually, wearable sensors such as accelerometers provided with a variety of applications

 ISBN: 978-81-944670-1-4 | November, 2020

are used to recognise their activity, different motions, and postures used during learning. Accelerometers use acceleration due to gravity to monitor and classify gestures and postures such as the tilt of a certain body part. In classroom practices, these accelerometers, together with gyroscopes, are used to feel their behavioral rhythm. For the identification of their presence and motion, non-wearable sensors such as infrared sensors are used. Other non-wearable sensors are often used to track the psychological and behavioral signals of pupils, such as ultrasonic sensors, pressure sensors, vibration sensors, video-based sensors, low-resolution thermal sensors, and audio or sound sensors.

- The Data Processing and Analysis layer will collect and start processing the raw data generated by the data sensing layer. Sensors detect vast quantities of data here, some of which are organised and others not. Such heterogeneous bulk data is either stored on an edge server or a cloud server. Then the processing of data at the edge or cloud would be considered. Several technologies such as ML, Big Data Analytics, Cloud computing, and edge computing and algorithms are needed in this processing and analysis phase.
- The Cognitive Computing Systems layer contains ML algorithms and techniques for extracting patterns and understanding movements of the face and body. To recognise various facial expressions (such as tired, yawning, fear, disgust, excitement, anger, unhappy, etc and situations, teaching techniques, materials used by the instructor to validate the hidden actions, task generation, and situation awareness measures are included. In general, the method of face recognition includes several steps: face localisation, registration, extraction of features, and classification [13]. To localise a face in an image or video, detection and segmentation algorithms will be used. These ML learning algorithms help make precise decisions by identifying face boundaries, the face is present or not (placing a binary label in an image or video for each pixel, color, or texture information), facial curvatures such as nose tips, eyebrows, ears, etc. The key points such as geometry, intensity, and labeled sections of the face can be identified and encoded by face registration. To discover the closest mapping between the model and the detected features, regression, and other machine learning techniques are used. ML with a proper training set such as face appearance (intensity, color, etc or geometric details (such as sleeping postures, face orientation, etc.) is used to extract important features. To detect the facial expressions of their emotions, trained face recognition algorithms (static or dynamic) can be used. Clustering is an algorithm for unsupervised learning that is used to detect several expressions beyond the predefined initial set. Dynamic face recognition involves discovering geometric features of face parts, Recurrent Neural Networks, and clustering to be used where students continue to alter their facial expressions.

The patterns of behavior are thus extracted, categorized, and remembered. These algorithms are required to make decisions and infer knowledge of user activity, actions, and involvement based on the context used by the environment of the teacher and classroom. The instructor has to carry out this review process weekly/monthly on teaching techniques used within the classroom. For each teaching style, the data after analysis is regularly recorded and the best approach will be chosen and implemented within the classroom to improve learning efficiency, outcomes, and behavior [14]. Therefore, all these various AI disciplines together evaluate the student's actions towards enhancing progress and suggest the instructor in the form of feedback.

V. Benefits And Challenges

In classroom teaching, fundamental behavior analysis is often needed. It has barriers as well as advantages. Until applying teaching methods to learners, it would be the duty of a teacher to evaluate and further refine the implementation. Since the teacher always expects the best performance from his strategies from the students.

A. Benefits

- Study engagement: Student behavior is very much needed to involve students either emotionally or cognitively. The learning efficiency will be improved by this interaction.
- Better learning experience: The most significant attribute of a human being is behavior. The other dependent variables, such as interest, mood, and quality of learning, will often be determined.
- Reflection: Thoughts of teachers are expressed on students by observing their actions. Teachers will also assess the best possible instructional approaches to be applied in the classroom.
- Individual evaluation: Examination of behavior can enhance personalized learning, as individual behavior is essential. The assessment of human emotions helps the teacher to find the best possible methods for teaching.

B. Challenges

- Hardware and software requirements: In the behavioral analysis, the most advanced hardware and software with networking technology are needed to achieve the most precise results.
- Expensive: multiple sensors, modern high volume technologies, and heterogeneous data are needed for behavioral analysis. The cost of implementation would also be high.
- Maximum preprocesses: Using the machine, analysis, and decision are enabled. It takes more and more training to act like a human being. It needs more preparation time than execution.

VI. Conclusion

In the classroom setting, the behavioral examination of the students is suggested automatically using their facial

expressions, hand movements, and body postures. This proposed framework also suggested potential cognitive computing algorithms used with the aid of a cognitive teacher to assess his behavior towards learning at each level of recognizing facial expressions and body movements. This paper also lists the challenges and advantages. The more effective automated data sensing, interpretation, and decision-making systems find the best intelligent approach is discussed in this paper.

The future study includes a review of the Pre-Post examination to be carried out on a group of students available within the classroom. It also involves testing the link between the behavioral involvement of the students and their success in the examination. The proposed approach can also be re-engineered and evaluated in numerous other fields for its applicability, such as media, athletics, and healthcare for user engagement evaluation.

REFERENCES

[1] Moreira, F., Ferreira, M. J., & Cardoso, A. Higher Education Disruption Through IoT and Big Data: *Springer International Publishing AG 2017*, *1*, 389–405, 2017.

[2] "The Internet Of Things For Educators And Learners." Accessed November 16, 2020. https://er.educause.edu/articles/2016/8/the-internet-of-things-for-educators-and-learners.

[3] N. Gligorić, A. Uzelac and S. Krco, "Smart Classroom: Real-time feedback on lecture quality," *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, Lugano, pp. 391-394,2012.

[4] Von Söhsten, D., & Murilo, S. Multiple face recognition in real-time using cloud computing, Emgu CV, and Windows Azure. International Conference on Intelligent Systems Design and Applications, ISDA, 9, 137–140, 2014.

[5] Al-Fuqaha, Ala, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari, and Moussa Ayyash.. "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications." *IEEE Commun. Surv. Tutorials* 17 (4): 2347–76, 2015.

[6] Manyika, J., Chui, M., Bisson, P., Woetzel, J., Dobbs, R., Bughin, J., & Aharon, D. "Unlocking the Potential of the Internet of Things". McKinsey Global Institute, Accessed November 3, 2020. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world

[7] Min, C. H. Automatic detection and labeling of self-stimulatory behavioral patterns in children with Autism Spectrum Disorder. In 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 279-282). IEEE, July 2017.

[8] Patil, P., & Sachapara, V. "Automatic Attendance Marking, Attention and Facial Expression Analysis System Using IoT". International Journal of Innovative Research in Science, Engineering and Technology, 6(2), 1804–1812, 2017.

[9] Vandenbroucke, L., Spilt, J., Verschueren, K., Piccinin, C., & Baeyens, D. "The Classroom as a Developmental Context for Cognitive Development: A Meta-Analysis on the Importance of Teacher–Student Interactions for Children's Executive Functions". Review of Educational Research, 88(1), 125–164, 2018.

[10] Spyrou, E., Vretos, N., Pomazanskyi, A., Asteriadis, S., & Leligou, H. C. "Exploiting IoT technologies for personalized learning". In 2018 IEEE Conference on Computational Intelligence and Games (CIG) (pp. 1-8). IEEE, August 2018.

[11] Bahreini, K., van der Vegt, W., & Westera, W. " A fuzzy logic approach to reliable real-time recognition of facial emotions. Multimedia Tools and Applications", 78(14), 18943-18966,2019.

[12] Padhi, A., Babu, M. R., Jha, B., & Joshi, S. "An iot model to improve cognitive skills of student learning experience using neurosensors". Springer Briefs in Applied Sciences and Technology, 5(2), 37-50. Springer, Singapore, 2019.

[13] El Mougy, A. " Character-IoT (CIoT): Toward Human-Centered Ubiquitous Computing. In Character Computing" (pp. 99-121). Springer, Cham, 2020.

[14] Jassim, E. K., & AL-Hemiray, E. H. " Cognitive Internet of Things Using MQTT Protocol for Smart Diagnosis System". Iraqi Journal of Information & Communications Technology, 2(3), 30–37,2019.

A Proposed Framework to enable Intelligence in an IoT based Classroom Environment from a Deep Learning based Multimodal Perspective

Lakshaga Jyothi .M, Research Scholar, Department of Computer Science and Engineering, Vinayaka Mission's Kirupananda Variyar Engineering College, Vinayaka Mission's Research Foundation (Deemed to be University), Salem, Tamilnadu, lakshagamecse06@gmail.com

Dr. Shanmugasundaram R.S., Professor, Department of Computer Science and Engineering, Vinayaka Mission's Kirupananda Variyar Engineering College, Vinayaka Mission's Research Foundation (Deemed to be University), Salem, Tamilnadu, rsssslm32@yahoo.com

Abstract. Smart Classrooms are going to be the future trend-set in the uprise of future technologies like the Internet of Things and Artificial Intelligence. Nowadays, these technologies are gaining their reach in all the corners of a diverse set of fields. Most of the educational institutions have started giving a chance to these technologies to serve them for their better functioning in some or other way. But then, due to setback, these technological implementations in the educational sector is still in the early stage. The success of any technological advancements lies in its adaptation to serve the society in the primary concern. Due to the miracles of Deep Learning principles in solving any tasks related to Computer Vision and Natural Language Processing problems, this paper has been taken as a jumpstart in studying the possible applications of Deep Learning to suit the education culture. But fortunately, Deep Learning systems can be made possible to meet IoT based technologies for its sustain in the long run. This paper has also investigated the potential applications of combining Computer Vision and Natural Language Processing tasks to serve an educational classroom setup. Such amalgamation has not reached the eyes of many researchers. Many researchers have designed an intelligent classroom in a context using single modalities either using standalone Computer Vision or Natural Language Processing problems or vice-versa. Hence to fill this gap, this paper has proposed an approach that can be integrated into an IoT system resulting in Intelligent classroom architecture.

Keywords — Internet of Things, Computer Vision, Multimodal data fusion approach, Deep Learning, Smart Mirror

I. Internet Of Things in Higher Education

Higher Education: Though, the higher education sectors have started integrating IoT systems in the educational space. Various research shows that there is no full-fledged solution for integrating such IoT systems in the academic sectors. Diverse nature of these systems and their challenging factors like Space, Cost, Availability, etc. Need to be taken into consideration for its implementation. Nowadays, automation has taken charge of taking care of some daily routine tasks to work smarter. When it transforms from smart to Intelligence, its context means a little different, as the former means quick delivery of the content automatically, and the latter is delivery of the content fast as well as efficient performing some computations as we desire. Visual cues from a Classroom environment include taking care of Teaching Activities, Classroom Activities, Administrative activities, Student Engagement activities, Academic Performance Activities, and Maintenance Activities.

II. Related Works

To make things more intelligent, it has to be equipped with a possibility to take produce actionable decisions resulting from better valuable insights. This paper

investigates how we could put forth Intelligence to a classroom environment to use AI algorithms for resource-constrained devices such as IoT devices, mobile devices, etc. As said by Mark Weiser [1], “The deepest technologies are the ones that fade away," emphasizing how computational processing will disappear when more and more information is being generated in the coming future. The prime goal of any IoT device is the possibility to grow connections among as many devices as possible anytime, anywhere, and anyplace[2].

Classroom: In a classroom, teachers, as well as students, are considered to be the main stakeholders. Various new pedagogical methods can be equipped with IoT to induce teaching through IoT technologies [3]. Few teaching approaches or styles were useful for IoT implementations for a smart classroom [4-9]. A smart learning paradigm with relevance to ubiquitous learning methods for student-centric learning experiences [10-11]. Learners engage through richer contexts using AR / VR to support smart and intelligent technologies [12]. To capture e-content or lecture videos in a classroom environment[13]. Collaborative e-learning through projects with IoT [14], Student-centered learning methods with projects via problem-based learning [15]. With IoT, student's attendance can be maintained based on their facial recognitions [16]. Learning analytics can be performed on the presence of students as well as teachers in the classroom setup [17], Using neuro or biosensors, tracking the neurological parameters, emotional stress states, etc. [18], Place sensors inside the classroom to take audio feeds, temperature records, wearable sensors to track and study multimodal data [19]. Monitoring classroom using IoT based on student engagement via their concentration levels [20], Feedback responses on the quality of lectures provided by the lectures [21]. Smart mirrors with AI [22] to work as an interactive mirror to show curriculums, college notices, class timetables, etc., for ease of work in the classroom setup. Classrooms with tel-education methods [23], student-teacher classroom relationship enhancement styles [24], transforming any classroom-technology working on control to transform smart into intelligent spaces as “intelligent classroom” via ubiquitous computing and ambient intelligence [25].

III. INTELLIGENT CLASSROOM

A. Background Researchers:

The current traditional classroom setting, where cameras can be introduced, and the camera feed videos to be examined with the principles of Computer Vision technology to constitute an intelligent classroom. Every other researcher has shown interest in creating an intelligent classroom with the venture to use Computer Vision Technology to understand the visual semantics of the classroom environment. [26] RFID associated with Ambient Intelligence with cameras to study the normal activity/states of the user adaptable for the user needs. An Intelligent campus (ICIoT) [27] that can run sensor nodes and WSN to manage power consumption with 16.7 Percent saving . Augmenting Mixed Reality in the classroom, the classroom experiences can be promoted or enhanced [28], The robust energy management system with INTEL GALILEO, and Z-HOME for an intelligent classroom [29]. Mega 2560 and Blynk has been used to manage energy efficiency inside the classroom [29], in [30] the author tied to track the classroom activities, with the cameras installed in the classroom. The author at [31] tries with a conceptual model to deliver a multiagent based framework to support classroom automation. Based on the SCADA infrastructure, the classroom is constructed with the IoT architecture to implement intelligent classroom [32], in [33] the author tries to track the face, feature and lip movements to incorporate machine learning algorithms to monitor student activities in the classroom [34], Deep Learning and the osmotic computing-based classroom was constructed, and researched [35], emotionally aware AI-based classrooms can be constructed as in [36].

B. Deep Learning and Computer Vision for Intelligent classroom experiences

i. Attendance monitoring

In [37], the Smartphone-based CNN approach to take attendance based on their known faces in the classroom []. in [38] with the use of SOTA techniques, attendance management system was constructed on the recognition of 98.67% accuracy with LFW dataset and 100% accuracy in the classroom dataset. With web-based attendance management systems along with CNN and MySQL for XAMPP web server [39], in [40] researcher tried to modulate a methodology to build face datasets on considerations with two quantitative and qualitative research datasets for marking student attendances in the classroom setup.

ii. Environment monitoring

As in [41], use cameras to monitor the environmental conditions and the activities of the students to promote some quality educational set up for the students.

iii. Emotion Recognition System

A Mobile oriented Cloud Hybrid Architectures (MOCHA) [42] has been used with deep learning focused on recognizing emotions with gestures and facial expressions.

To study [43], non-verbal cues of the students and their performances in the classroom-based 350 students and 71 percent precision with the collected data of the Gold Standard report on comparison with Cohen-Kappa precisions.

iv. Edge / Fog computing for IoT devices

IoT based smart classroom [44] construction has been focused on using neural networks with fog microservers and edge computing tools in the classroom for an osmotic computing-based classroom setup.

v. Affective computing

Study on the affective states on the students in the classroom [35], especially in the e-learning environment based on spontaneous and posed datasets with 83 and 76 percent accuracy in detection as well as classification.

vi. Machine translation

Koren speaking students have been helped with web-based machine translation tools to study their action inside the classroom using Google translate / Navel translate [45]

vii. Self Annotation classroom reports

The author in [46], By recording the audio versions based on single, multiple, and no voice dataset in the classroom set up to study the time consumption and self-reports generations and annotations automatically in the classroom

viii. **Behavior analysis system**

To monitor the activities of the teachers and the students in the classroom [47] by leveraging 1800 frames of 6 videos and 10-20 participants in the classroom for an effective student behavior monitoring system

IV. PROPOSED MULTIMODAL PERSPECTIVE

Due to the recent breakthroughs in multimedia processing, it is explored to deal with the perspective of combining the tasks or approaches or different types of data sources that can be termed as "multimodal data sources." Based on the theory of 'semiotics,' where the study is to relate the signs and its meaning by interpreting them in terms of words, phrases, or sentences based on the semantic representations. This kind of visual content retrieval is supported by distributed semantics. Multimodal data sources such as video, audio, text, etc., can be approached with some cognitive perception.

The proposed approach works on the basis of constructing a classroom setup conceptually that can be helpful to build and implement the same in further research work based on multimodal learning paradigms. As described in Figure.1, the architecture was proposed based on the applications of Deep Learning techniques on fusion with Computer Vision and Natural Language Processing systems enabled through IoT systems in a Classroom Environment. The scenario considered for this setup consists of different modules from a faculty perspective since very few types of research have been conducted to study the teachers present in the classroom. 1. Attendance Management System 2. AI-based Smart mirrors for interactive information displays with regards to classroom information 3. Lip Reading systems to recognize the cues based on different modalities such as Audio signals as well as video/image signals viz. Lip Reading, etc. based on light-weight neural networks on multimodal data fusion approach such as combining an Image feature extractor using some Convolutional Neural Network (CNN) model and then Audio feature extractor to work on time series or sequential data such as Recurrent Neural Network (RNN), Long short term memory (LSTM), etc. Finally, integrate the multimodal data fusion approach to aggregate those two networks to make a prediction on the movement of lips of the user or speaker in the classroom setup on use-cases such as visual passwords, visual speech recognitions for word, or sentence level predictions.

In this paper, it is taken as a step towards implementing such combinations in a classroom set up for the first time since lip reading has been investigated in many scenarios but not experimented in an educational use-case like classroom setup. This may need to be studied from various factors such as all the students in the classroom or the teacher in the classroom. Initially, the teacher can be taken as a primary model for dataset collection and experimentation on the same setup, which needs minimal parameters for experimentation. On the basis of this conceptual framework, implementation work is in progress. The educational sector is not an exclusion since Deep Learning is growing its path in all the fields of research for the betterment of human society. Making Deep Learning possible for the IoT systems considering its computational limitations on working with light-weight neural networks or the embedded neural networks equipped to an IoT device, either it is a Raspberry Pi or a Mobile device. In the mere future, computers will seek to explain and interpret real-world situations as we can see, hear, and respond to.

i. Initial questionnaire study:

Based on the scenario, initially, a questionnaire has been constructed to support my work on implementing IoT technologies in a classroom environment with few questionnaires on the considerations and feedbacks from a collective set of pupils from the student background, later it will be collected from the teachers perspective related to using IoT based cameras inside the classroom environments. Out of 50% responses from the participants, more than 50% have responded positively to support my research work in using IoT technology in the classroom setup. Below is the category distribution of participants for this study.

Table 1. The above table mentioned outlines the participant's distribution for this questionnaire study

Questions	Participants	Positive responses
1. Do you think adopting educational technology will transform today's teaching and learning process?	50	35
2. Do you think current technological advancements are beneficial to the student's community?	50	46
3. What is your idea of using IoT technology in the Higher Educational sectors?	50	30
4. What is your idea about using cameras inside the classrooms, and specify your suggestions	50	31
5. Do you think cameras inside the classroom will benefit those the most?	50	Admins and staffs (35 collectively)
6. Face recognition cameras can be useful inside the classroom	50	34
7. How do you want the cameras to be used effectively inside the	50	More than 35% for attendance

classroom?		and lecture recordings in the classroom
8. Do you think the cameras will affect the student community emotionally	50	30
9. Are you aware of Face recognition technologies	50	41
10. If confidentiality maintained, do you believe face recognitions cameras will help the education sector	50	33
11. Using cameras will improve students attention and engagement inside the classroom	50	35

Table 2. The above table mentioned outlines the questionnaires and the positive responses out of 50 participants.

ii. Preliminary Testing:

Each module has to be preliminarily studied, for which initial work has been started to sustain my progress in the research work during this pandemic situation.

Module 1. Attendance management system. Datasets were collected from a total of 5 participants with 200 datasets each on a sum of 1000 images per each participant. Further data pre-processing, training, and testing phase is in progress.

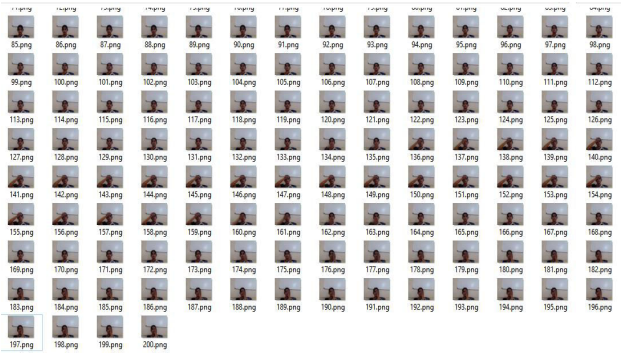


Figure 1. Sample dataset collection of a participant of 200 samples

Module 2. Smart Mirror system. A smart mirror module has been installed before the classroom setup and experimentation on to show regular updates like display

details, temperature, daily news, etc

Figure 2. Smart mirror set up with all the necessary installation done and started with the regular updates.

Module 3. For this module, preliminary is not yet completed. Utilize the Face videos and then extracting the mouth region to experiment on the Multimodal data fusion approach (multiple modals such as speech or text) to learn the data and its representations in a classroom setup. But the initial pipeline is constructed as follows: Collect Mouth region datasetfrom the existing face videos, then preparation of correct dataset for the algorithm implementation, to train the dataset and find the correct parameters for the model, test the same and check for implementation in a convolution based neural network and similarly, the same follows with collecting audio data and pass them through the recurrent based neural networks and integrate them in a single network and predict the lip movements and the words spoken.

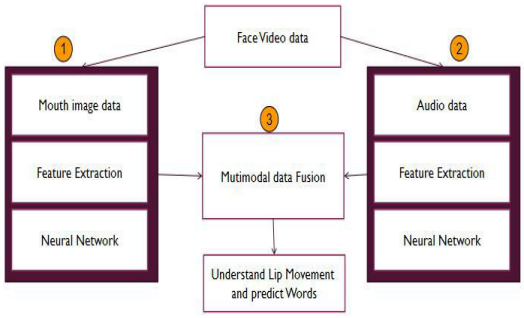


Figure 3. Workflow pipeline for the proposed multimodal data fusion approach

iii. Conceptual design:

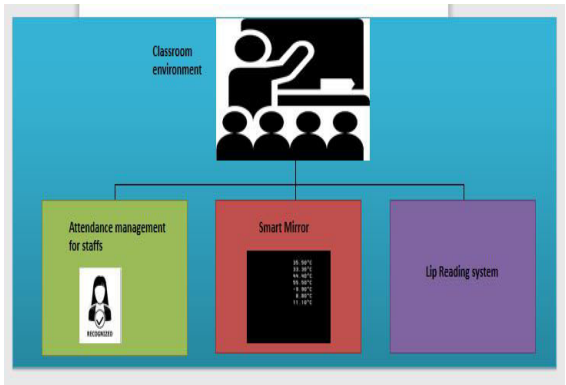


Figure 4. The architecture of the conceptual design

V. FUTURE WORK

Based on this conceptual setup, implementing the above design for formal classroom experimentation is in progress. The breakthrough in Deep Learning to solve complex problems and meet IoT technologies led to the beginning of this research work. This research tries to propose a conceptual framework to enable intelligent classrooms by integrating Deep Learning for IoT devices, thereby filling the gap to support educational classrooms employing these hybrid systems. This framework was thus helpful to progress the research work to the next stage of concept implementation, and the result is in progress. The future work investigates the possibility of including Augmented Reality / Virtual Reality (ARVR) inside the classroom with the above technologies. This could be taken as the first step to making the implementation more successful in the classroom scenario. Considering this research as a starting point, it can be possible to focus on building intelligent classroom infrastructure. To provide enhanced performances of the educational workflows in the coming future.

ACKNOWLEDGEMENT

Our sincere thanks to our institute Vinayaka Mission's Kirupananda Variyar Engineering College, Salem, Tamilnadu, for extending the facilities for my research through the Centre for Research and Development (CRD).

VI. CONCLUSION

In this paper, we have discussed the applications of IoT in Higher Education. There are few previous types of research conducted to build an intelligent classroom based on a different context. This paper has been organized so that it initially investigates the glimpses on the applications of IoT in Higher Education and its Classroom setups. Secondly, the significant applications of Deep Learning-based Computer Vision scenarios for a Classroom Environment as approached by many researchers. Thirdly, it discusses the styles of creating an Intelligent Classroom from their context. Finally, the proposed approach consisting of different modules to work on the perspective of how the multimodal data fusion approach works for a classroom environment. The leading objective of this research work is to explore how traditional classrooms can be added with advanced technologies like Deep Learning. The future generation classrooms can be a (DLeIC) Deep learning enabled IoT Classroom to lift the educational space into a new dimension. The future of these IoT based deep learning systems seems to amuse as more connecting things will be communicating besides less human intervention resulting in a new era of the higher education sectors and its environment that will thrive to survive. This implementation work can be considered as experimentation that can lead to gain more values from the more profound analysis of the classroom semantics for better actionable insights. This proposed architecture reveals that these state-of-the-art technologies can induce the developers, researchers, and scientists in new directions to enable Intelligence in a Classroom Environment. Based on this proposed architecture, concept implementation for a proof-ready solution to build systems can give rise to intelligent classrooms. The future of Deep Learning-based IoT systems seems to be more promising for the near future. Thus, it can be concluded that future classrooms will become an IoT connected environment for both the staffs and the students.

REFERENCES

1. Shahla Gul, Muhammad Asif, Shahbaz Ahmad, Mahammad Yasir, Muhammad Majid, M.Sheraz Arshad: A Survey on Role of Internet of Things in Education. International Journal of Computer Science and Network Security, Vol. 17, no. 5, 2017.
2. Keyur K Patel, Sunil M Patel, T.: Internet of Things- IoT, Definition, Characteristics, Architecture, Enabling Technologies, Application and Future Challenges. IJESC International Journal of Engineering and Computing, vol. 6, Issue no. 5, ISSN: 2321 3361,2016.
3. Tikhomirov V., Dneprovskaya, N., Yankovskaya E.: Three Dimensions of Smart Education, 2015.
4. Fitzgerald, R., A Smart Teaching System, 2013.
5. Moazeni, S., Pourmohammadi, H. Smart teaching quantitative topics through VARK learning styles model. IEEE Integrated STEM Education Conference, 2013.
6. Tilton., W., Adult professional development: can brain-based Teaching strategies increase learning Effectiveness? Fielding Graduate University, ProQuest, 2011.
7. Tredowski, TN., Woods, AM.: Seven Student-Centered Principles for Smart Teaching in Physical Education, Journal of Physical Education, Recreation & Dance. vol. 86, no. 8, pp. 41-47, 2015.
8. Zhu, ZT, Yu, MH, Riezebos, P.: A research framework of
9. Smart education. Springer Open Online Published, Smart Learning Environment (2016).
10. Gwak.D.: The meaning and predict of Smart Learning. Smart learning Korea Proceeding, Korean e-Learning Industry Association, 2010.
11. Kim T., Cho J.Y., Lee B.G.: Evolution to Smart Learning in Public Education: A Case Study of Technologies for Networked Learning. IFIP Advances in Information and Communication Technology, Springer, Berlin, 2013.
12. Kim, S., Song, S.M., & Yoon, Y.I.: Smart learning services based on smart cloud computing- Sensors, vol. 11, no. 8, pp. 7835-7850, 2011.
13. Shivaraj Kumar T.H, Sriraksha T.A, Noor U Saba.: An IOT Based Secured Smart e-Campus, International Journal of Humanities and Social Science Invention, vol. 6, no. 3, pp. 88-93, 2017.
14. Mercatus, 2015
15. Shrinath, Vıkhyath, Shivani, Sanket, Shruti: IOT Application in Education, International Journal of Advanced Research and Development, vol. 2, no. 6, 2017.
16. Chang. C.H.: Smart Classroom roll caller system with IOT architecture, Proceedings of 2nd International Conference, IBICA, pp. 356-360, 2011.

17. A.Alghamdi, S.Shetty: Survey toward a smart campus using internet of things. Proceedings 2016 in IEEE 4th International Conference, Future Internet of Things and Cloud, FiCloud, 2016.
18. Kusmin, M., Laanpare, M., Saar, M, Rodriguez-Triana, M.J.: Smart Schoolhouse as a Data-Driven Inquiry Learning Space for the Next Generation of Engineering. IEEE EDUCON – Global Engineering Education, 2017.
19. Bajaj V., & Pachori, R. B.: Detection of human emotions using features based on the multiwavelet transform of EEG signals. Intelligent Systems Reference Library, 2015.
20. A.Rythivaara.: Collaborative classroom management in a co-taught primary school classroom. vol. 53, pp. 182-184, 2012.
21. Chew. C.B.: Sensors enables Smart Attendance System using Nfc and Rfid technologies. International Journal for New Computer Architecture and their Applications. vol. 5, no. 1, pp. 19-28, 2015.
22. Shreyansh Khale., Aditi Sathe, Rugveda Salunke, Shweta Nathan, Amit Maurya, Smart mirror, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume.8, Issue-2S11, September 2019.
23. Uskov V.L., Bakken J.P., Pandey A.: The Ontology of Next Generation Smart Classrooms. İn: L.Uskov V., Howlett R., Jain L. (eds) Smart Education and Smart e-Learning, Smart Innovation, Systems and Technologies, vol. 41, Springer, 2015
24. Andrew W. Wright.: RFID Classroom Management System. Master thesis, Industrial and Manufacturing Engineering, California Polytechnic State University (2011).
25. R.Josè, H.Rodriguez, N.Otero: Ambient Intelligence: Beyond the Inspiring Vision. Journal of Universal Computer Science(J.UCS), pp-1480-1499 (2010).
26. Rabie A. Ramadan, Hani Hagras, Moustafa Nawito, Amr El Faham and Bahaa Eldesouky; "The Intelligent Classroom: Towards an Educational Ambient Intelligence Testbed" Intelligent Environments (IE), Sixth International Conference, 2010.
27. Ping Zhang and Jianzhong Wang; "Management of Intelligent Campus Wireless Sensor Networks Based on Runtime Model", Journal of Computer and Communications, pp-22-31, 2015.
28. James DOOLEY, Vic CALAGHAN, Hani HAGRAS, Micheal Gardner and Daniyal AL_GHAZZAWI; "The Intelligent Classroom: Beyond Four Walls", Presented at the Intelligent Campus 2011 (iC'11), Nottingham 26th July 2011.
29. Anish Gupta, Punit Gupta and Jasmeet Chhabra; "IoT based Power Efficient System Design using automation for classrooms" 3rd International conference on Image Information Processing, IEEE Computer society, ISBN: 987-1-5090-0148-4, 2015.
30. Yasodhran R, Karthick S, Prince Roy and HariKrsihnan V; "IoT based Classroom Automation using Arduino", International Journal of Trend in Scientific Research and Development (IJTSRD), Vol.2, Issue.2, ISSN: 2456-6470, Feb 2018.
31. Nafhath Rasheeda Rafiq, Saida Fatima Mohammed, Jitendra Pandey and Ajay Vikram Singh; "Classic from the Outside, Smart from the Inside: The Era of Smart Buildings" 2017 6th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), Sep. 20-22, 2017.
32. Jose Aguilar, Priscila Valdiviezo, Jorge Cordero and Manuiel Sanchez; "Conceptual design of a Smart Classroom Based on Multiagent Systems", International Conference. Artificial Intelligence, 2015.
33. Huang, L.S., Su, J.Y. and Pao, T.L.,A context aware smart classroom architecture for smart campuses. Applied Sciences, 9(9), p.1837, 2019.
34. Shashi Pal Singh, Ajai Kumar, Archana Singh and Kartika Jain; "Smart and Intelligent Next Generation Classrooms over Cloud", Proceedings of APSIPA Annual Summit and Conference 2017.
35. Pacheco, A., Cano, P., Flores, E., Trujillo, E. and Marquez, P., 2018, October. A smart classroom based on deep learning and osmotic IoT computing. In 2018 Congreso Internacional de Innovación y Tendencias en Ingeniería (CONIITI) (pp. 1-5). IEEE.
36. Kim, Y., Soyata, T. and Behnagh, R.F., Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education. IEEE Access, 6, pp.5308-5331,2018.
37. Karnalim, Oscar, Setia Budi, Sulaeman Santoso, Erico D. Handoyo, Hapnes Toba, Huyen Nguyen, and Vishv Malhotra. "Face-face at classroom environment: Dataset and exploration." In 2018 Eighth International Conference on Image Processing Theory, Tools and Applications (IPTA), pp. 1-6. IEEE, 2018.
38. Sarkar, Pinaki Ranjan, Deepak Mishra, and Gorthi RK Sai Subhramanyam. "Automatic attendance system using deep learning framework." In Machine intelligence and signal analysis, pp. 335-346. Springer, Singapore, 2019.
39. Sutabri, Tata, Ade Kurniawan Pamungkur, and Raymond Erz Saragih. "Automatic Attendance System for University Student Using Face Recognition Based on Deep Learning." International Journal of Machine Learning and Computing 9, no. 5 (2019).
40. Budi, Setia, Oscar Karnalim, Erico D. Handoyo, Sulaeman Santoso, Hapnes Toba, Huyen Nguyen, and Vishv Malhotra. "IBAtS-Image Based Attendance System: A Low Cost Solution to Record Student Attendance in a Classroom." In 2018 IEEE International Symposium on Multimedia (ISM), pp. 259-266. IEEE, 2018.
41. Tew, Yiqi, Tiong Yew Tang, and Yoon Ket Lee. "A study on enhanced educational platform with adaptive sensing devices using IoT features." In 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), pp. 375-379. IEEE, 2017.
42. Kim, Y., Soyata, T. and Behnagh, R.F., Towards emotionally aware AI smart classroom: Current issues and directions for engineering and education. IEEE Access, 6, pp.5308-5331,2018.
43. Gupta, Sujit Kumar, T. S. Ashwin, and Ram Mohana Reddy Guddeti. "Students' affective content analysis in smart classroom environment using deep learning techniques." Multimedia Tools and Applications 78, no. 18 (2019): 25321-25348.
44. Ashwin, T. S., and Ram Mohana Reddy Guddeti. "Affective database for e-learning and classroom environments using indian students' faces, hand gestures and body postures." Future Generation Computer Systems (2020).
45. Briggs, Neil. "Neural Machine Translation Tools in the Language Learning Classroom: Students' Use, Perceptions, and Analyses." Jalt call journal 14, no. 1 (2018): 2-24.
46. Cosbey, Robin, Allison Wusterbarth, and Brian Hutchinson. "Deep learning for classroom activity detection from audio." In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3727-3731. IEEE, 2019.
47. Ngoc Anh, Bui, Ngo Tung Son, Phan Truong Lam, Phuong Le Chi, Nguyen Huu Tuan, Nguyen Cong Dat, Nguyen Huu Trung, Muhammad Umar Aftab, and Tran Van Dinh. "A computer-vision based application for student behavior monitoring in classroom." Applied Sciences 9, no. 22 (2019): 4729.

Need and Challenges of Online Voting

K. M. Kiran Raj
Computer Science and Information Science
Srinivas University
Mangalore, India
kiranraj224@gmail.com

***Abstract*—Choosing representatives using elections is the core task of any democratic structure. The election method increases accountability, ease and protection with improvements from paper ballots, mechanical lever machines, punch cards, scanned paper ballots to Direct Electronic recording machines. India, with more than 910 million voters spread across 1 million polling stations, is the largest democratic country. Even with the use of E-voting through VVPAT machines, the turnout of India in recent general machines is only 67.47%. Many new techniques or practices were suggested to address this, such as mobile voting, block-chain-based systems and internet voting, but it could not be implemented even with several benefits. Security and confidence needs to be further strengthened in order to enforce it in practise.**

Keywords—Ballot, Direct electronic recording machine, E-voting, Mobile voting, Block-chain.

I. INTRODUCTION

Election is important part of any democratic system through which right to select representatives is given to every individuals. Election is used from ancient Greece, Rome, Medieval and Vedic period to select rulers. Many unfair methods Electoral frauds were conducted [1]. Election process is perfected through years of experience & hard work. Different organizations or system follow different election methods which is convenient for them. India follows democratic system with 542 parliamentary constituencies, 908.7 million electors, 1 million polling station with 12 million polling officials. India uses EVM with VVPAT to conduct election which enhanced credibility & transparency. In 2019 Lok Sabha election had the highest voter turnout of 67.47% [2]. New technologies like blockchain, bio metric verification is proposed to make election more secure.

II. TYPES OF VOTING SYSTEM

A. Paper Ballot

The voter will cast vote by marking against the printed ballot paper where name & party symbol is printed. the Postal ballot and ePostal ballot is provided for the individuals unable to arrive in polling station. The Marker or pen will be used [3] [4]. Some of the involving problems are

- The marking of ballot outside provided box is considered invalid.
- Any other forms of mark on the ballot paper is considered invalid.
- Counting is done manually which is error prone, time consuming.
- Requires huge labour work & ballots
- In Postal ballot / ePostal ballot the ballot should arrive to returning officer in provided time period.
- Unauthorized assistance may be provided in marking the ballot [5] [6].

B. Punch Cards

In punch cards the voters will punch holes against the candidate name or the number representing the candidate. The punched cards will be fed to the ballot box [7]. Some of the problems are [8]

- The cards can be damaged during the punching process.
- The template / cards must be aligned properly or the voting will be invalid, or chances of voting to another candidate increases.
- Jamming of Punch card readers.

C. Lever Machines

Levers are labeled against the candidates names. The voter will cast his vote in secret. The counters of the selected candidates is incremented. The problems with lever machine are [8] [6]

- The Defects are found in odometer counting machine.
- Lever machines are expensive to move, store and maintenance is complex.

D. Optical Scanning System

The voters are provided with machine readable ballots where the voter will completely darken the circle against the candidate name. The ballots will be collected and fed to machine [9]. OMR / OCR / ICR system will be used to tabulate the votes. Some of the problems are [8]

- Writing anything on ballot paper considered as over-voting.
- Mark Sense Ballots are larger and requires more time for processing.
- Invalid votes because of not shading the provided oval shapes completely

E. Direct Recording Electronic Machine

In India, Brazil, Belgium and Venezuela most of the voters used DRE System. The Electronic device is used to store the votes. Voters will press the button against the name of the candidate and party symbol. To enhance transparency and credibility Voter Verified Paper Audit Trail (VVPAT) is used along with Electronic Voting Machine (EVM) [2].

III. NEED FOR ONLINE VOTING

India follows largest democratic system with were representatives are selected through election. In the 2019 Lok Sabha election is conducted for 542 constituencies across 1 million polling station, in 908.7 million registered electors

 ISBN: 978-81-944670-1-4 | November, 2020

613 million voters casting their voting right including 2.28 million postal ballots. Fig. 1. Lok Sabha Voter Turnout % from 1951 - 2019 provides the voter turnout % it can be seen that the highest turnout of voting is 67.47%. From the last 3 elections the voter turnout % is gradually increasing.

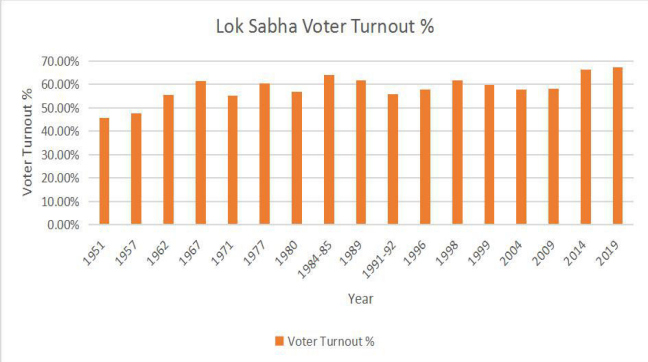


Fig. 1. Lok Sabha Voter Turnout % from 1951 - 2019

The some of reason for the adopting online / internet voting is listed below [4] [2]

A. *Voter Turnout*

Even though there is huge amount of registered electors, the progress in voter turnout is slow. In 17th Lok Sabha election highest voter turnout as 67.47%. That points to 32.53 % of voters did not turn up to cast their vote. The Election commission and government is trying to make improvement by conducting various awareness programs but improvements are very slow.

B. *Time Consuming*

2019 Lok Sabha election is conducted for 39 days from 11th April to 19th May 2019 in 7 phases and results were declared on 23rd May 2019. Since India is a diverse nation with huge landscape the election cannot be conducted in single phase and will be conducted in several phases.

C. *Expensive*

VVPAT is used along with EVM at every polling station to remove any forms of mistrust from the voters mind. 2.33 million ballot units, 1.635 million control units and 1.74 million VVPAT machines were used. These machines and its maintenance is complex and expensive.

D. *Requirement of Physical Presence*

India has voter turnout of 67.47% mainly because voter can only cast his vote from his constituency only. If voter is working in other constituency or in other country he/she will not be able to cast vote. Only through physical presence at constituency voting can be done.

E. *Less Information / Postal Ballot*

The voter are neglecting the important process of the democratic system even though government and Election commission is trying to provide it through different awareness programs.

The postal ballot can be used by officials, government employees and members of armed forces. Many voters not having information regarding it because of it's complexity nature.

F. *Manual Involvement*

The machines and the postal ballots are handled manually because of which they are error prone. Around 12 million officials were used for election process.

IV. Internet / Online Voting

To overcome the disadvantages / problems of existing system Internet Voting is used. The voter will cast his / her vote through the internet, where voter will login to the website through computer or mobile and will cast vote by ticking boxes [10]. Estonia is the only country that is able to provide Internet voting. There are 3 different forms of internet voting depending on how and where the votes are casted.

A. *Polling Site Internet Voting*

The voters will be physically present at polling station and they will cast vote using the machines provided by the officials. Authentication of voter is done similar to other system.

B. *Kiosk Internet Voting*

The voters will cast their votes from the machines provided from the officials. The machines can be present at public places. Authentication of voter will not be under election officials.

C. *Remote Internet Voting*

The voter will be able to cast vote from any remote place like home, workplace etc. Using personal devices like mobile or computer devices. The machines and authentication of the voters are not under the control of the election officials. Remote internet voting in not secure like polling site internet voting and kiosk internet voting.

Some of the advantages of internet voting are listed below:

1) Increase in Voter Turnout: Physical presence of the voter is not required at the polling station. Voter can vote from any place due to which individuals not present at constituency or not able to come to polling station because of health issues are able to vote.

2) Time Efficient: Less time is required for counting and the election can be held in a single phase, due to very less manual requirement

3) Less Expensive: Maintaining & purchase the EVM, VVPAT or Ballots is expensive which are not required in internet voting. The election officials required for smooth processing of election is also less.

V. Internet / Online Voting In Estonia

Estonia is a small country located in the northeastern Europe formed of 1500 islands and islets, with population of 1.3 million [11]. Estonia is one of the country with flexible digital ecosystem. One of the first country to use internet voting from 2005. The 44 % of the voters use internet voting. The internet voting will be opened in advance before the election day. The internet voting can only be done in

provided time period i.e. 10 days before voting to 4 days before voting. The process of internet voting is listed below [12].

- Voter identification is done by using ID card / Mobile ID when login to the system. The ID card will be having unique ID signature. When its PIN is entered verification code will appear on Mobile screen.
- After login to application candidate and party symbol according to voters will be displayed.
- The voter can choose any candidate and vote for them.
- The confirmation of candidate can be done by the user by entering the PIN followed by appearance of verification code on mobile screen.

The voter can vote many times as required but only the last vote will be considered and remaining votes are removed.

A. Reasons For Success of Internet Voting in Estonia

1) Digital Flexibility: 99% of public services are provided to the citizens through online which made voter to adopt to internet voting very easily. Technically advanced country and provided smart cards to voters from 2 decades.

2) Population: Estonia has small population of only 1.3 Million where active voter is just over 0.5 million which makes it easier to handle the data that is generated in the election process. Providing new information, validation and authentication can be easily managed in very less time.

3) Trust: Estonia did not directly implemented internet voting using Mobile / Application but did it using providing ID card with digital signature and voting was possible through it by using card reader. Which gained the trust of the voters.

4) Politically Neutral State: Estonia is politically neural country, less enmity with other counties because of which other countries will not try to affect the voting system of it.

VI. Challenges Of Internet / Online Voting

Even though Internet / online voting seems to be tempting it is only successful in Estonia. Other countries have not implemented it mainly because of several security issues. Some of them are listed below [13].

A. Risk to Vote in Transmission

After the voting is done vote needs to go to election office server from computer. Vote will be going through hubs, routers if suspicious person get access to these network infrastructure the votes can be altered or destroyed without changing the digital signature or users information.

B. Malware on Device / Impersonation Attack

The voters device like mobile or computer may be infected with malware, trojan or viruses. Which may effect the voting or attacker may easily damage voters vote through them with the help of existing problems in voters device. Voters information and credentials can be stolen from the attacker and cast vote by impersonating as voter.

C. Denial of Service Attack

The server can be flooded with the request intentionally or unintentionally causing denial of service. When the denial of service is caused the server may be crashed because of unable to process the incoming request. The server may get slowdown drastically and other voters will not be able to cast votes.

D. Anonymity

One of the main feature of election is anonymity. The vote should be completely anonymous where no one should able to find out who have voted for so that no one can bride or threaten the voter. Internet voting fails to provide secrecy since it can be done from any place.

E. Trust

Other main feature of election is trust where system needs to be transparent and secure. The system should be able to convince everyone irrespective of technology. The voter needs to completely trust software and hardware is secure. The voter also needs to trust the transmitting votes through internet is safe. Internet voting is unable to provide it.

VII. Conclusion

Election is important any democratic system. Election can be conducted by many ways like Paper Ballots, Punch Cards, Mechanical Lever Machines, Optical Scanning System, Direct Recording Electronic Machines and Internet Voting. Democratic system provides different options for the voter to cast votes. Every system has its own advantages and disadvantages but since Direct Recording Electronic Machines have less disadvantages DRE is used by many countries. The government and election commission is trying to increase voter turnout but its increasing at slow rate. To increase voter turnout internet voting can be used but because of the security issues not implemented by many countries. The Security and new technologies must be added to be used in real time.

Reference

[1] Wikipedia contributors. (2020, November 11). Election. In Wikipedia, The Free Encyclopedia. Retrieved November 13, 2020, from https://en.wikipedia.org/w/index.php?title=Election&oldid=988195344

[2] Election Commission of India. (2019). 101 INNOVATIONS & INITIATICES. ECI Publications.

[3] Massachusetts Institute of Technology, & Lab, E. D. + S. (n.d.). Voting technology. Retrieved on November 10, 2020 from https://electionlab.mit.edu/research/voting-technology

[4] Kadam, T. (2020). Online Voting system. International Journal of Engineering Trends and Technology (IJETT), 37(5), 273–276. https://doi.org/10.2139/ssrn.3589075

[5] Wall, A., Olivier, L., & Ngidi, S. (n.d.). Challenges to Voters. ACE project. Retrieved on November 12, 2020 from https://aceproject.org/ace-en/topics/vo/vog/vog04/vog04a

[6] W. Jones, D. (2001, May 22). Problems with Voting System and Applicable Standards. The University of Iowa. Retrieved on November 10, 2020 from http://homepage.cs.uiowa.edu/~jones/voting/congress.html

[7] ACE Project. (n.d.). Punch Cards. Retrieved on November 5, 2020 from http://aceproject.org/ace-en/topics/et/eth/eth02/eth02b/eth02b1/mobile_browsing/onePag

[8] Selker, T. (2004). Old Voting Technologies: Problems and Improvements. 1–5.

[9] Laanela, T., & Green, P. (n.d.). Optical Scanning Systems. ACE Project. Retrieved on November 10, from http://aceproject.org/ace-en/topics/et/eth/eth02/eth02b/eth02b2

[10] ACE Project. (n.d.). Inernet Voting. Retieved on November 5, from http://aceproject.org/ace-en/topics/et/eth/eth02/eth02b/eth02b4/mobile_browsing/onePag

[11] Stranga, A., & Misiunas, R. J. (2020, November 17). Estonia. Encyclopædia Britannica. Retrieved on November 3, 2020 from https://www.britannica.com/place/Estonia

[12] Valimised. (n.d.). Stages of i-voting in voter application. Retrieved on November 2, 2020 from https://www.valimised.ee/en/internet-voting/stages-i-voting-voter-application

[13] Greenhalgh, S., Goodman, S., Rosenzweig, P., & Epstein, J. (2018). Email and Internet Voting: The Overlooked Threat to Election Security. ACM Computing Surveys. https://www.acm.org/binaries/content/assets/public-policy/jtreportemailinternetvoting.pdf

Corpus Analysis on one of the Philosophical works of Acharya AbinavaGupta and Acharya Sri Adi Sankaracharya

Thejash M N
Department of Social Science and Humanities
Srinivas University

Mangalore, India
thejash.cssh@srinivasuniversity.edu.in

Abstract **– This research paper documents a Philosophical corpus study on the work of Acharya AbinavaGupta and Acharya Sri Adi Sankaracharya.**

The document in the corpus is prepared using Sanskrit to English translated verses in 'Tantraloka' a work of AbhinavaGupta and English translated verses of 'Vivekachudamani' a work of Adi Sankaracharya. As a result, the corpus reflects some of the most significant texts on the magnificent work of Acharya AbhinavaGupta and Acharya Sri Adi Sankaracharya

The data or verses are manually copied into excel. Data is prepared by removing the special characters, numbers, and stop words using python modules like Stop words and regular expression. The data is then lemmatized and the frequency of the words are studied

The paper provides: (1) the spread of the corpus (2) the Jupyter notebook interface that helps in coding and analyzing the corpus using the NLTK (Natural Language Toolkit), spacy module, and word cloud (3) Graphical representation of words using word cloud and (4) a set of HTML visualizations using scattertext to navigate the corpus

Scattertext visual basically gives an idea of what are the most discussed text in the verses of Acharya AbhinavaGupta on the Y-axis and Acharya Sri Adi Sankaracharya on the X-axis and the diagonal overlap is something that are common texts across 2 Philosophical verses

Keywords - Corpus, term frequency, Graphical representation, Corpus Visualization

I. INTRODUCTION

The study of the greatest Philosophical work of Acharya AbhinavaGupta was entitled "Tantraloka", who lived in Kashmir (924-1020 CE).
ABHINAVAGUPTA was very intellectual and a great scholar. With Bharatamuni and Anandavardhana, AbhinavaGupta is one of the greatest scholars in Indian aesthetics during the 10^{th} – 11^{th} Century AD. His philosophies represent a great success of achievements in ancient India. There are many splendid works on Philosophical texts, work on Abhinava-Bharati which represents art and dance which is most popular to date, Poetical works, devotional hymns, and religious works. The thirst for knowledge in AbhinavaGupta was uncontrollable.

Another most famous work that set forth Vedanta Philosophy is of Acharya Sri Adi Shankaracharya titled "Vivekachudamani". ADI SANKARACHARYA was an Indian philosopher, born in a quiet village on the banks of Curna(also called Purna and periyaru) river in Kerala. Born in 778CE, but the truth is that direct evidence about Adi Shankaracharya's birth is practically non-existent. Adi Shankaracharya works can be broadly divided into 3 sections

1. Bhashyas (Commentaries)
2. Prakriya(Prakarana) Granthas (Concept of Vedanta)
3. Stotras (Hymns and meditation)

In this paper, we present a corpus that contains the digitized version of all the available translated verses and the application to query and analyze the corpus. This paper describes the corpus of the work of Acharya AbhinavaGupta entitled "Tantraloka" and works of Acharya Sri Adi Sankaracharya entitled "Vivekachudamani", the available Jupyter notebook interface to code and analyze the corpus and add HTML visualization to locate the corpus and view different word frequencies

The collected philosophical verses are English translated from Sanskrit and each verse is documented in excel cleaned and then loaded into Python environment for frequent word analysis and the data is represented using Scattertext and word cloud modules.

*Scattertext is used to visualize how two groups of words differ from each other.
*Word Clouds are used to visualize the depiction of selected words.

.

II. CORPUS DESCRIPTION

Corpus is a collection of text which are used to do hypothetical testing and statistical analysis.
There are 794 corpora documented in excel for Acharya AbhinavaGupta and 580 corpora for Acharya Sankaracharya and these corpus consists of 24033 tokens and the data are captured from 2 philosophies "Tantraloka" and "Vivekachudamani". The document "Tantraloka" corresponds to both ritualistic and philosophic aspects and

the document “Vivekachudamani” corresponds to Vedanta theory and practice.

Table 1: Number of corpus per Acharya

Acharya	Corpus
AbhinavaGupta	794
Sri Adi Sankaracharya	580

III. DATA PREPARATION

Here in data preparation, the data is manually entered into Excel and is reviewed and checked for any inconsistency and incorrect data. Data is then loaded into Jupyter notebook as a Pandas Data Frame. Using NLTK and regular expression library the data is cleaned by removing the stop words, special characters, and numbers. The data is then lemmatized.

IV. CORPUS METHODOLOGY

In this section, we document the Python Interface using Jupyter notebook that helps to analyze the corpus. Python Interface using Jupyter notebook integrates the excel file verses described in the previous section with NLTK, spacy, and word cloud module to facilitate corpus querying and data analysis.

Step 1: Import necessary Python libraries for the study and load the research raw data into a Pandas data frame.

Step 2: Before the analysis, common words that would look to be of less value. Selected words are excluded from the vocabulary entirely such as ‘the’, ‘them’, and ‘they’. Also removing special characters and numbers are necessary as these provide no value in our analysis. We use NLTK and regular expression to remove the stop words, special characters, and numbers

Step 3: Data is then pushed to the spacy module which is specifically designed to understand large volumes of text.

Step 4: Process Natural Language text through Spacy en_core_web_sm which a small English model is trained on the text.

Step 5: Create a scattertext corpus of the dataset

Step 6: Get term frequency and scaled f score

Step 7: Scatter plot is created to visualize the associations between high and low frequency of the text.

We made the Y-axis category "Acharya AbhinavaGupta", and name the category "AbhinavaGupta”. We named the other category “Sankaracharya” on the X-axis. Finally, we transfer the visualization to an HTML file which is can be viewed through the Chrome browser.

V. CORPUS VISUALIZATION

Below are the visualizing term associations of Acharya AbhinavaGupta and Acharya Sri Adi Sankaracharya using the word cloud. The words that are important appear more regularly. Bigger the word that appears more frequently used. The HTML visualizations are created using the Python Scattertext module which gives a view of how each corpus differ between Acharya AbhinavaGupta and Acharya Sri Adi Sankaracharya.

Fig.1.

Fig.1. Graphical representation of Top 100 words which shows the distribution that is most associated with “Tantraloka” of Acharya AbhinavaGupta using Word Cloud

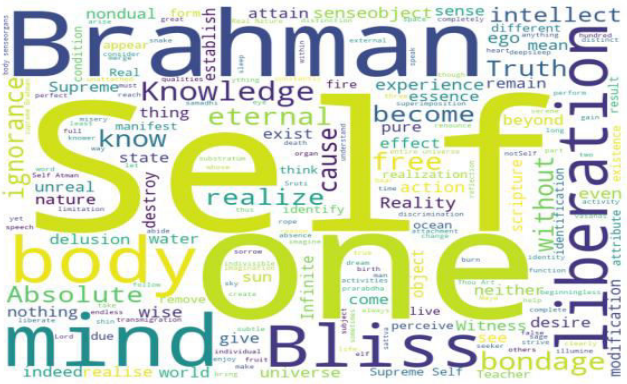


Fig.2

Fig.2. Graphical representation of Top 100 words which shows the distribution that is most associated with "Vivekachudamani" of Acharya Sankaracharya using Word Cloud

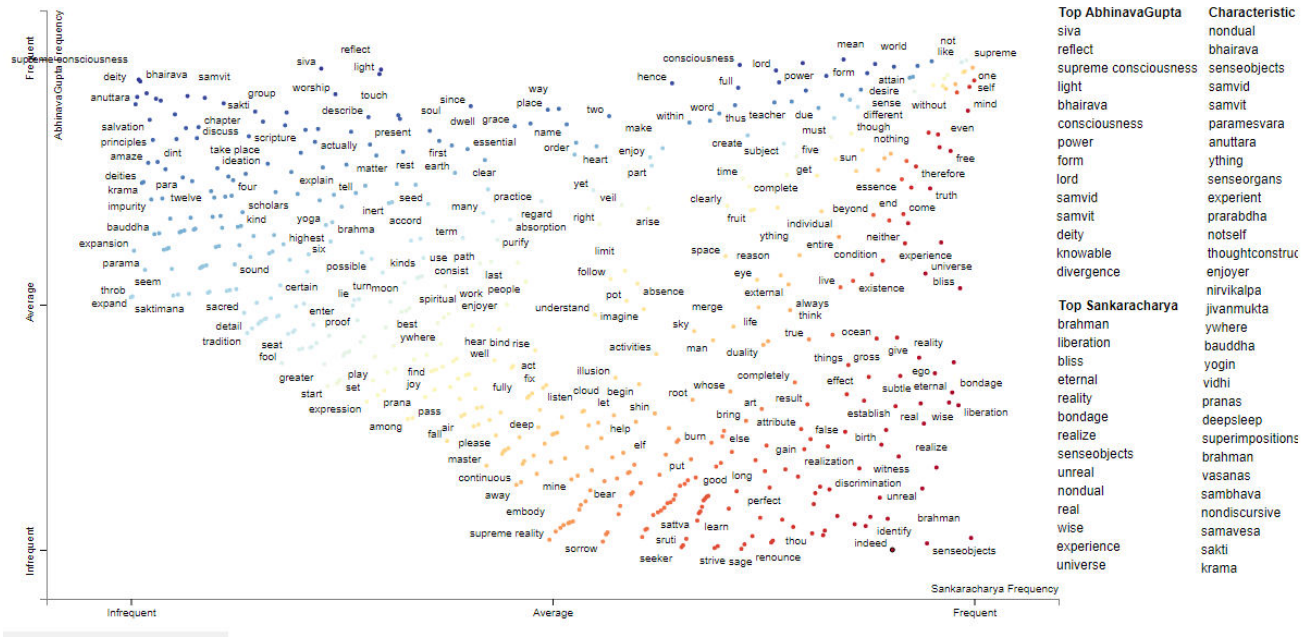


Fig.3.

Fig.3. The link between the corpus text, the Jupyter-notebook Integration, and the HTML visualizations.

The words or the texts are colored-coded by their association. Texts that are more associated with "Tantraloka" by Acharya AbhinavaGupta are blue, and text more associated with "Vivekachudamani" by Acharya Sri Adi Sankaracharya is red.

In Figure 3, we can see how word usage in the written Philosophical text by Acharya AbhinavaGupta and Acharya Sri Adi Sankaracharya. Each point corresponds to the usage of a word, where the higher up the word is on the y-axis, the more it was used in "Tantraloka" of Acharya AbhinavaGupta and the further right was used in "Vivekachudamani" of Acharya Sri Adi Sankaracharya. Words used often, appear in the top right-hand corner, while rarely used words, like "expression" or "master" appear in the lower-left corner.

The location of words shown correlates with the frequencies in this work. The most continuing used words are displayed next to the second most continuing used words. Only important or outstanding words are visible in the plot. We can closely view which words correlate to which points if we move the mouse. Click on each word and these words disclose how words are used in the context.

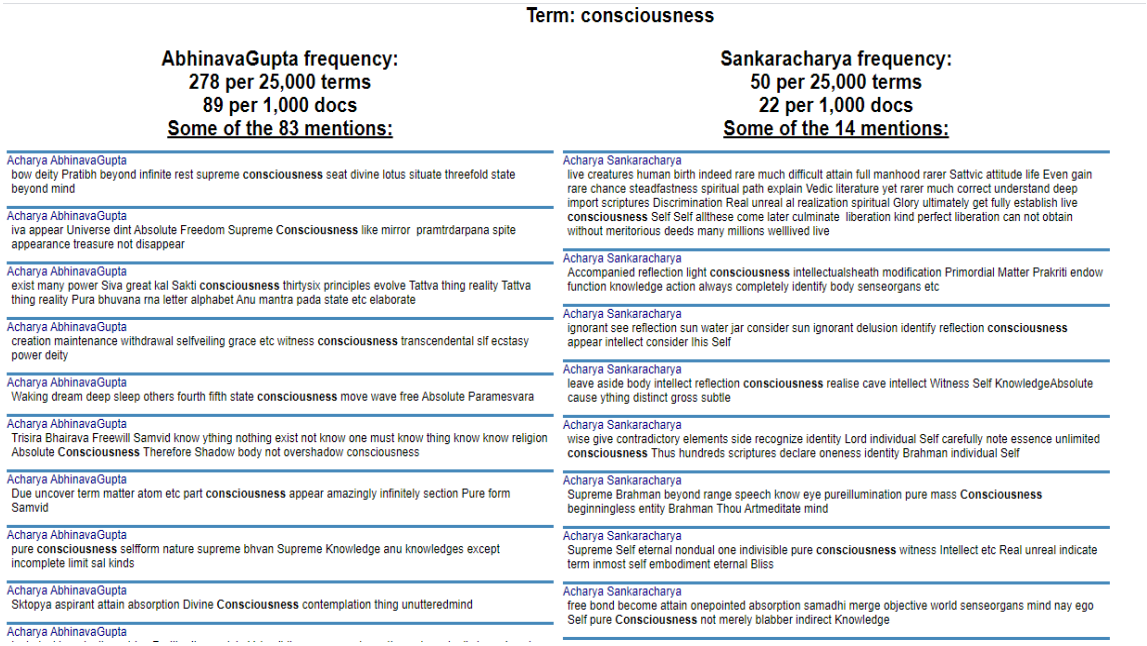


Fig.4.

Fig.4. Most relevant visualization of the word “Consciousness”. Here in “Tantraloka” of Acharya AbhinavaGupta contains the word “Consciousness” 278 times and “Vivekachudamani” of Acharya Sri Adi Sankaracharya contains 50 times.

VI. DISCUSSION

The corpus analysis of Acharya AbhinavaGupta in “Tantraloka” verses talks majorly about Siva, supreme consciousness, light, and bhairava. Acharya Sri Adi Sankaracharya in “Vivekachudamani” verses mainly talk about Brahman, liberation, bliss, non-dual, and realize. The diagonal overlap is the common word used in the philosophies of Acharya AbhinavaGupta and Acharya Sri Adi Sankaracharya are supreme, desire, mind, world, etc. The result can be further used for comparative study, sentiment analysis, etc.

VII. CONCLUSION

We have presented a corpus on the Philosophical work of Acharya AbinavaGupta and Adi Sankaracharya. This corpus consists of the English translated verses from the books written by Acharya AbinavaGupta and Adi Sankaracharya. The corpus includes the digitized texts of 794 verses of AbhinavaGupta in Tantraloka and 580 verses of Sankaracharya in Vivekachudamani. Secondly, we have introduced a Python interface (Jupyter notebook) that integrates the texts using NLTK, spacy, and word cloud modules and corpus are analyzed on the frequencies on the texts written by Acharya AbhinavaGupta and Acharya Sri Adi Sankaracharya

In this study, we use HTML visualization to view the different frequency of words used in the analysis. The visualization search can be used to search a particular word to know the frequency and the context.

ACKNOWLEDGMENT

The author is very grateful to Acharya AbhinavaGupta and Acharya Sri Adi Sankaracharya for the grace and blessings to complete this research paper.

My family who supported me with love and understanding

Dr. Lavina D’Mello, Dean Social Science and Humanities, Srinivas University, and all the Academic Professors of Srinivas University for the support and guidance

REFERENCES

[1] Chaterjee, G. (2008). Sri Tantraloka – Text with English translation, AbhinavaGupta. Varanasi

[2] Chaitanya, P. and Dhiman, S. revised and edited with notes and an Introduction (2012). Sri Sankara’s Vivekachudamani: Devanāgari Text, Transliteration, Word-for-Word Meaning, and a Lucid English Translation. Tamilnadu

[3] Devy, G. N. (2002). Indian Literary Criticism. (pp.61). Hyderabad

[4] Vasiliev, Y. (2020). Natural Language Processing with Python and SpaCy – A Practical Introduction. San Francisco

[5] Mahadevan, T.M.P. (1976). Ten saints of India. (pp. 88-89). Bombay

[6] Varma, P. (2018). Adi Sankaracharya. (pp. 3). Chennai

[7] Central Chinmaya Mission Trust Bombay. Sankara- the missionary. (Chapter 5). Jamshedpur

[8] Álvarez-Mellado, E. (2020). A Corpus of Spanish Political Speeches from 1937 to 2019. Waltham

Detecting Botnets in Network Traffic using machine learning strategies

Sangeetha Prabhu
College of Computer Science and Information Science
Srinivas University
Mangalore, India
sangeethaprabhu.ccis@srinivasuniversity.edu.in

***Abstract*— over the last period, there is a significant research initiative to establish new methods that can efficiently and effectively classify botnets. As a result, different techniques of identification were described depending on multiple technology and different dimensions of botnet phenomena. In recent news reports, botnet was the principal vector in bringing various cyber crimes. Although a massive amount of botnet analysis and identification research was being conducted, various issues remain unanswered, including developing detectors that interact with new types of botnets. This paper explores contemporary methods of Botnet identification using machine learning to assess the traffic associated with Botnet. The paper offers an in-depth look at current detection approaches by analyzing the algorithm bot-related mechanisms and how various machine learning strategies were modified to obtain botnet-related information. It compares current identification techniques with their features, successes and limitations and puts particular focus on innovation with the performance assessment strategies and procedures. The analysis also describes limits, obstacles and possibilities to use the computer training to recognize botnet traffic and to build machine-based botnet detection systems.**

***Keywords*—** Cyber Security, Botnets, Botnet Detection, Traffic Analysis, malicious activity, Intrusion Detection

I. INTRODUCTION

Botnets are a disruptive network-based attack spec today as they require the use of massive organized hosts for brute and subtle assaults. Those wide numbers of hosts are organized into so-called zombies or bots, which are then managed from a distance. A selection of bots forms what is called a botnet when managed by a single infrastructure (C2). By offering an indirect level, the attack host is isolated by a layer of zombies from its victim and the attack itself is randomly skewed from the assembly of the botnet. Both the combined bandwidth and the botnets' scope derive their strength from size. By large distributed denial-of - service attacks, Botnets can cause major network failures, and the possibility of such disruption can pay businesses big ransom charges. They are accounted for a large number of viruses on an entire internet.

Botnets are often used to collect medical, business, or government confidential information for sale on a flourishing organized crime market. They are a sustainable resource and green. The use of algorithms for learning machines to recognize patterns of malicious traffic is one of the newest developments in networking botnet detection. The key premise of computer-based learning approach seems to be that botnets establish unique structures within traffic and these structures can be identified efficiently using ML algorithms [1]. As other types of botnet detection technologies, such as honeynet and signature-based methods, have been found to be incredibly inept, machine learning algorithms have been growing and popular in botnet detection.

The use of irregular network traffic techniques by machine learning algorithms is effective as it does not allow use of pre-built network signatures and therefore can identify new and unknown botnets [2]. However, researchers have found that signatures are unhelpful tools for avoiding attack, as the latest botnets were equipped with highly advanced firmware updates and evasive strategies.

II. CONCEPT OF BOTNET

Botnets have become the key subject to security issues on the Internet almost since 2003[2]. The number of attacks [3–6], misuses of digital identity and compromised computers [7] has inspired researchers to establish improved detection techniques. Botnets have created serious security issues on the Internet and, due to its continuous development concerning their architecture, protocols and manner of attack; they have increasingly become evasive [8]. This part of the review will X-ray conceptual models behind botnets and machine learning to provide a clearer understanding of the issue.

The botnet explains a network of infected hosts/devices that run software robots and are monitored by humans through one or more control systems. The infected hosts are named as "Bots" the person one, which manages the botnet is named "Botherder" and control devices are pointed to as "Botmaster"[9]. Botnets, according to [10], are a collection of compromised computers that carry malicious software that is used for large-scale attacks through various cyber network elements.

Hoang and Nguyen in [2] claim how each botnet component is considered a bot. A bot is a malicious program written by a bunch of developers that helps themselves monitor and control-compromised computer systems. Bots vary from several Ransomware in that they have been unique in its own way and fitted with contact networks to seek guidance and config files from the navigation systems. Botnets regularly alert their operating condition to their control systems. Botnet connectivity is pointed as the C&C nodes. Through the C&C channels the botmaster transfers the instructions to the bots. The security study [10] reveals that approximately 50 percent of internet traffic is linked to botnet operations, including spamming and security breaches. Grizzard et al. [11] have defined the key goal of the botnets as follows.

a. The dispersion of information: This occurs by delivering SPAM messages, refusing service threats and providing unauthorized networks with fake information.

b. Information retrieval: Botnets manage information processing through identification, financial data and password and relationship data. Information retrieval.
c. Information Processing: Information processing is conducted by processing password information to reach additional hosts. Information processing.

III. THE BOTNET PHENOMENON

Botnets are dynamic and sophisticated phenomena, spreading numerous advanced C&C tactics and destructive practices. The attackers often outfit their botnets with a wide variety of durability functions [12–15] which are developed specifically for making it challenging and even difficult to recognize them. The knowledge of botnet operations is important for developing new detection techniques and for skilled reflection on modern detection systems. Botnet phenomenon's complexity is better grasped by life-cycle of botnets, Command & Control communication channel & techniques for botnet flexibility and reliability. Three key facets of the botnet phenomena are presented further in the succeeding sections.

A. Botnet life-cycle

The function of Botnet can be dealt with by evaluating the botnet life-cycle that is a series of functional stages found during operations with the botnet. The identification addresses specific phases of the life-cycle of botnet by using specific botnet activity methodologies in these phases. For this purpose, it is important for the effective study of current botnet detection work to consider the botnet life-cycle. Various writers like Silva et al.[16], Feily et al.[17] identified the botnet life-cycle as state-setting. Zhu et al.[18] These researchers have similarly described the botnet life-cycle and separated the botnet activity into three stages: the infection, the communications and the attack process. Although the definition of the author's different operating stages varies, the life-cycle of botnet is as shown in Fig. 1.

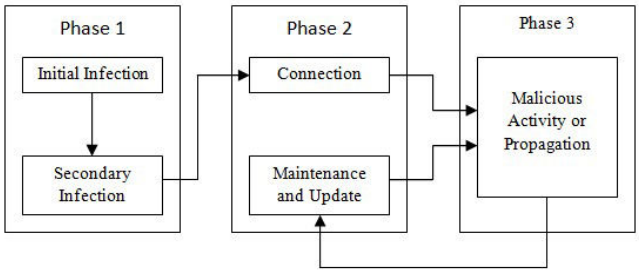


Fig. 1. Botnet life-cycle

The infection stage is the first stage of botnet life-cycle, where compromised computers become zombies in a certain botnet. The infection phase is subdivided into two sub-phases namely initial and secondary infections. The original infection may occur in numerous forms, e.g. through unintended malware updates, by uploading corrupted files via email messages, malware spreads from contaminated removable disks. The initial contamination may occur. The main role of the loader is to help get the bot malware. When the primary infection is successful, the sub-phase of the secondary infection begins when the loader installs the malware from the external network location onto the compromised computer. You can download the bot malware binaries by using different protocols, including, for instance, FTP, HTTP/HTTPS or P2P transfers.

The coordination phase is the second phase o botnet life-cycle. This process consists of multiple modes of operation involving communicating infected machines and C&C servers. The contact process involves communication on receiving orders, alerts, and information on the current state of bots from the botmaster. The contact includes many operating modes: initial connexion attempts with the Command & Control server during a successful infection process, link attempts by the bot following reboot of the affected computer, intermittent connexion attempts to report the status of an infected machine, and C&C server link attempts to change malware code or spread. The Command & Control channel can be installed in many ways, so communication between Command & Control server and zombie is achieved.

The third stage of the life-cycle of botnet is marked as an assault stage since it involves a bot operation to enforce the destructive agenda of assailants. Zombie computers will launch DDoS attacks during an attack, launch SPAM email campaigns, conduct-stolen identity delivery, click-fraud deploy exploit credibility networks and surveys, etc. [19–20]. The systems may also enforce dissemination processes such as searching for compromised computers and spreading malware during this operating time. The second and third stage is mechanically connected such that, if a vulnerable device is corrupted, they are typically altered one by one. However, the various phases of life-cycle of the botnet can follow different times so the length of the operation can vary based on the bot camp.

B. C&C channel

Command and Control Channel are the primary carrier of botnet to identify functionality. The Command and control channels are a contact channel between computers affected & botmaster. The attacker uses this channel to issue bottlenecks and receive input from the affected devices [21–23]. The C&C channel allows a large number of Bots to be remotely coordinated, which introduces consistency in botnet operations by enabling malicious botnet code to be modified and upgraded. The Command and control channel is also seen as a core component of the botnet phenomena and thus an important means for botnet identification. Over the last few years, the communication infrastructure of C&C has developed rapidly. Consequently, the C&C channel [21–25], [11], was introduced by multiple restrictions in relation to protocols and network design. The botnets can be categorized based on the C&C network topology as clustered, decentralized or hybrid network architecture.

The radar of botnet identification and neutralization systems is one of the key objectives of the operation. For this purpose, the attackers outfit their botnets with various durability strategies to provide stealth and robustness. The resilient techniques implemented on a network level are designed to provide secrecy and completeness of communication, botmaster anonymity and robustness in the effort of the Command and Control channel. The misapprehension of existing protocols and the introduction of customized communication, and the encoding of touch networks is one of the major factors of hiding information with C&C. The protection and integrity of the

correspondence are ensured using these approaches, thus essentially defeating detection strategies that are focused on traffic content for detection.

C. Resilience techniques

One of the key goals of botnet operation is to run the botnet identification and neutralization systems under the radar. As a result, intruders empower their botnets with various capability techniques aim of supporting reliability and robustness of process. Established on the device layer, stability strategies are built to verify the credibility and anonymity of contact, the morality of the botmaster and its efficiency of the fast-acting C&C channel. Among the most essential part of securing the confidentiality of C&C communication is the blurring of existing and developing custom routing protocols and the encryption of the communication channel. Using these approaches, the confidentiality and privacy of information are protected, essentially defeating surveillance strategies reliant on the quality of the traffic payloads to be identified. However, the use of encrypted communication networks and abstracted communication protocols may be deemed suspicious and may be seen as a catalyst for further traffic analysis. Fastflux and Domain Generation Algorithm (DGA) [12] are other widely used methods that provide immunity to botnet activity.

The core idea behind Fastflux is to provide different IP addresses connected to a singular top-level domain name in which IP addresses will be swapped and out of an excitation wavelength by shifting DNS records. Fastflux [14] is commonly seen in botnets to mask malware & phishing delivery websites behind the fastest-changing server of compromised computers that operate as proxy servers. This protects the anonymity of Command & Control servers while offering highly efficient harmful services. It should be noted, moreover, that the use of Fastflux produces a particular heuristic botnet that can be used for the efficient identification of botnets [26].

IV. BOTNET ARCHITECTURES

Anwar et al.,[27] Categorizes botnet architecture into three categories specifically; centralized decentralized and hybrid architectures.

a. Centralized architecture: With centralized botnet architecture, the botmaster handles all bots throughout the botnet from a single central hub referred to as the command and control server. In this all the bots directly connect to a C&C server so all obtain instructions from it. As once C&C server has been recognized, such a botnet can be removed very easily.
b. Decentralized infrastructure: In decentralized botnet architecture, no one computer manages the bot in the botnet. There are many commands and control servers that are linked and interact with the bots. With this sort of botnet architectural design, every bot on the botnet is a control and command server, as well as a zombie (bot). Detecting this form of botnet is challenging since it has no centralized authority.

Hybrid Architecture: It is a mixture of both centralized and decentralized architecture. It specified that there are two types of robots with hybrid architecture, namely, the customer's bot and the commander's bot. Tracking and identification of botnets with hybrid architecture is more difficult than with centralized and decentralized architecture.

V. BOTNET DETECTION

When botnets become dangerous, researchers and security experts used various methods and strategies to tackle the problems. The detection method determines how well the idea works such as behavioral detection [28] or signature [29]. Numerous strategies are based on various methodologies. ML-based detection techniques consist of using both approaches. Other methods used to detect bots are anomaly detection and DNS.

A. Problems with Existing Botnet Detection Systems

Researchers have made multiple attempts to establish mechanisms for botnet identification. However, issues have been reported in the current botnet detection methods, and the following paragraph outlines these potential threats and opportunities for improvement in the near term.

Lia and Chang[15] have developed a peer-to - peer (P2P) botnet tracking system utilizing data mining technique. The P2P botnet identification technique relied on traffic management input layer and data mining techniques to evaluate the behavior of the network. System disadvantages are like; it works only within the network background of the locality and will have to be repeated to the ISP level to identify the P2P botnet in a wide network. Second, the presence of NAT technology makes it difficult for the system to distinguish P2P flows. The researchers suggested a large-scale network for stronger and more stable botnet identification.

Silva et al, [16] formed a standard features extraction metric for botnet attack detection. The research further treated a similarity between both the generic feature selection and genetic algorithm. The framework had some drawback it included the growth of very high false positive rates and contained a very large dataset that were too difficult to evaluate.

Shin et al, [30] designed a peer-to - peer (P2P) botnet detection system by manipulating the strength of P2P enemies against them. The machine deployed port point retrieval, but found it hard to identify specialized encrypted data.

Santana et al, [31] provided a detailed description of performance parameters of two machine learning techniques namely Random forest and multi- layer perception in malware detection by using data analytics. The methodology recommended specific or unique strategies for detecting attacks instead of a generalized model. They did not capture the performance assessment of their framework when background traffic and/or highly unbalanced data were unmarked.

Hoang and Nguyen [2] proposed a botnet identification m echanism based on domain name query of machine learning. Reviewed the efficacy of the solution using many machine learning algorithms and experimental studies found that Random Forest Algorithm provided the maximum overall recognition performance over 90 percent. The framework did not perform the impact of Domain name device on identification; suggested processes of the conceptual scheme with bigger sample size to help examine the impact of the DNS feature on enhancing detection accuracy.

Mathur et al,[32] studied and explored five separate classification activities to obtain some of the more appropriate for botnet identification. The logistic regression classifier was found to be best fitted for intrusion detection systems. One to the downsides of this strategy is that regular network utilization is changing rapidly in the new cyber world and, occasionally, this structure wouldn't be capable of distinguishing between benign and malicious data. A continuum of growth is proposed as neural networks with deep learning training are used to account for evolving botnets

B. Problems with Existing Botnet Detection Systems

The basic premise behind machine-based learning methodologies is that botnets generate distinct characteristics of traffic and actions within the application server and these patterns can be identified using some Machine Learning Algorithms [33–34].

The sub-field of artificial intelligence is Machine learning that seeks to develop and analyze data learning mechanisms. Training in this sense means the capacity to understand dynamic trends and to make trained decisions on the basis of previously seen evidence. In machine learning the real hurdle is to make the assumption of information acquired from a small range of previous interactions possible in needed to create it informative for the new. To address those issues in the field of Machine Learning develops a range of methodologies which contextual influences from precise information and observations based on reasonable qualitative and quantitative analytical concepts. Machine learning focuses on ideas and results from several areas, such as control theory, biology, Artificial intelligence, statistics, philosophy, cognitive science, and knowledge theory. Produced MLAs are the foundation of various functions, starting from image recognition to speech recognition, forecasting, pattern classification, gaming, optimization techniques and nanotechnology. At the same period, major advances in ML theory and algorithms have promoted machine learning as the primary means of exploring information from massive amounts of data existing in numerous application fields. A few of the evolving problem domains are botnet identification, which relies on the ML algorithms to identify the bot-related patterns.

On the basis of the algorithm's desired result ML algorithms are categorized into two general areas:

1. Supervised learning
2. Unsupervised learning

Supervised learning is a group with well-structured ML algorithms that produce a mechanism that maps inputs to the expected output [35]. Such algorithms are conditioned by samples of inputs and their expected results, which is then used to simulate the output for any possible inputs. This algorithm is used to grade the entry data for a given class and for regression that predicts a real - valued result.

Unsupervised learning is the form of ML algorithms in which test set comprises a set of inputs without any corresponding outputs. The primary objective in unsupervised learning issues is to explore a set of related instances inside the input data, in which it is termed clustering, to understand the determinants of information within the input data, referred to as density analysis.

In the scenario of anomaly-based botnet identification, ML comprises the average of categorizing or grouping traffic by using unsupervised and supervised ML algorithms. Network traffic is evaluated on both packet and flow level in which unique variations of traffic are obtained. Derived traffic functions identify traffic, which finalizes the specific server throughout the particular traffic flow or the network.

In supervised learning case, ML for botnet identification can be incorporated as expressed in fig. 2a. The supervised machine learning algorithm is first equipped using data sets, establishing a function, which maps inputs and expected results. The function, which is often related to as a framework deploy to distinguish inputs from testing results. All training and evaluation data must be correctly preprocessed to use in Machine learning algorithm. Pre-processing is formulated by a pre-processing phase component that retrieves the characteristics from information available and picks, which would be included in the MLA. Selecting the perfect characteristics is the most difficult challenge of the massive implementation of MLA. Characteristics should be selected in such a way as to allow them to obtain guided botnet protocols. Often common supervised MLAs included in botnet identification are Decision Tree Classifiers, SVM, Bayesian Classifiers, ANN, etc.

Compared to the supervised learning case, the unsupervised learning scenario involves in use for grouping bot-related observations. The principal advantage of unsupervised ML algorithms would be that they do not get to be instructed in advance. Unsupervised intrusion prevention ML algorithms are configured as seen in Figure 2b. These methodologies pre-process the knowledge available by retrieving and selecting the characteristics and using the unsupervised MLA to combine inferences, comparable to each other, within the same cluster. The significant issues for the successful execution of this type of learning scene are the selection of suitable features and the persistence of the clusters. Some unsupervised learning methods used to identify botnets are Hierarchical clustering, k-means, and X-means.

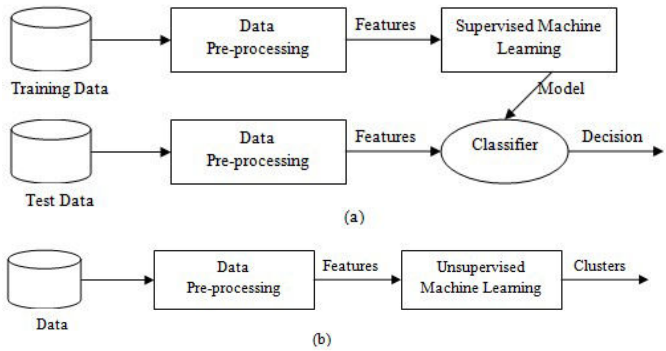


Fig.2. Frameworks for botnet detection: a) supervised learning framework and b) unsupervised learning framework

The situations described for the implementation of botnet identification MLAs are clearly a simpler example of the botnet detection methodologies based on ML. Real-life pre-processing data configurations usually involve additional, advanced processing to retrieve features that could relate experience-directed botnet algorithms. Simultaneously with the scenarios seen in Figure 2, many of the modern machine-based learning methods implement identification across

multiple stages, using a mixture of various MLAs or by applying MLAs in an integrated fashion. More comprehensive and precise, adaptable and scalable identification can be achieved in this way.

CONCLUSION

The new developments in network-based botnet identification with the use of Machine Learning Algorithms s to recognize harmful traffic patterns. The primary principle of ML -based methodologies are that botnets establish distinct structures within the network activity and these structures can be effectively identified using MLAs. Modern machine-based monitoring systems must intend at data-adapting, on-line, and effective identification. The framework must be inclined to keep up with the evolving trends of botnet activity and should start operating on-line to provide appropriate detection and to meet the demands of immediate botnet dissemination. The techniques must be time-consuming and computation time powerful so that they can be easily deployed on network interfaces, secure wider network scopes and provide more in-depth visibility of the traffic caused by botnets.

REFERENCES

[1]. Stevanovic, M., & Pedersen, J. M. (2015). On the use of machine learning for identifying botnet network traffic. *Journal of Cyber Security and Mobility*, *4*(2–3), 1–32. https://doi.org/10.13052/jcsm2245-1439.421

[2]. Hoang, X. D. (2018). Botnet Detection Based On Machine Learning Techniques Using DNS Query Data. *Future Internet*, 1–11. https://doi.org/10.3390/fi10050043

[3]. Braverman, M., Williams, J., & Mador, Z. (2008). Microsoft Security Intelligence Report. *Microsoft Security Intelligence Report Microso*, *June*. http://book.itep.ru/depository/ethernet/Microsoft_Security_Intelligence_Report_2009.pdf

[4]. Wilson, C., & Botnets, C. (2008). Cyberterrorism: Vulnerabilities and policy issues for congress. In *Foreign Affairs, Defense, and Trade Division, United States Governemnt, CRS Report for Congress* Retrieved from https://fas.org/sgp/crs/terror/RL32114.pdf on 06/11/2020

[5]. Stock, B., Göbel, J., Engelberth, M., Freiling, F. C., & Holz, T. (2010). Walowdac - Analysis of a peer-to-peer botnet. *EC2ND 2009 - European Conference on Computer Network Defense*, 13–20. https://doi.org/10.1109/EC2ND.2009.10

[6]. Zhang, Y., Xiao, Y., Ghaboosi, K., Zhang, J., & Deng, H. (2012). A survey of cyber crimes. *Security and Communication Networks*, *5*(4), 422–437. https://doi.org/10.1002/sec.331

[7]. Microsoft. (2010). Microsoft Security Intelligence Report. *Microsoft Security Intelligence Report*, *10*, 1–19. https://download.microsoft.com/download/6/0/5/605BE103-9429-4493-898B-E3D50AB68236/Microsoft_Security_Intelligence_Report_volume_10_July-Dec2010_English.pdf

[8]. García, S., Zunino, A., & Campo, M. (2014). Survey on network-based botnet detection methods. *Security and Communication Networks*, *7*(5), 878–903. https://doi.org/10.1002/sec.800

[9]. Mahmoud, M., Nir, M., & Matrawy, A. (2015). A Survey on Botnet Architectures, Detection and Defences. *International Journal of Network Security*, *17*(3), 272–289. http://ijns.jalaxy.com.tw

[10]. Chigozie-Okwum, C., & Ajah, I. A. (2019). Botnet Identification Using Machine Learning Techniques: A Survey. *2nd International Conference on Education and Development*, *July*. https://www.researchgate.net/publication/334284867

[11]. Grizzard, J. B., Sharma, V., Nunnery, C., Kang, B. B., & Dagon, D. (2007). Peer-to-Peer Botnets: Overview and Case Study. *HotBots*, *7*(2007) Retrieved from https://www.usenix.org/legacy/event/hotbots07/tech/full_papers/grizzard/grizzard_html/ on 06/11/2020

[12]. Zhang, L., Yu, S., Wu, D., & Watters, P. (2011). A Survey on Latest Botnet Attack and Defense. *2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*. https://doi.org/10.1109/TrustCom.2011.11

[13]. Nazario, J., & Holz, T. (2008). As the Net Churns : Fast-Flux Botnet Observations Tracking Fast-Flux Domains. *2008 3rd International Conference on Malicious and Unwanted Software (MALWARE)*, 24–31. https://doi.org/10.1109/MALWARE.2008.4690854

[14]. Holz, T., Gorecki, C., Freiling, F. C., Rieck, K., & Networks, F. S. (2008). Measuring and Detecting Fast-Flux Service Networks. *Network & Distributed System Security Symposium (NDSS)*. http://www.isoc.org/isoc/conferences/ndss/08/papers/16_measuring_and_detecting.pdf

[15]. Liao, W. (2010). Peer to Peer Botnet Detection Using Data Mining Scheme. *2010 International Conference on Internet Technology and Applications*. https://doi.org/10.1109/ITAPP.2010.5566407

[16]. Silva, S. S. C., Silva, R. M. P., Pinto, R. C. G., & Salles, R. M. (2013). Botnets: A survey. *Computer Networks*, *57*(2), 378–403. https://doi.org/10.1016/j.comnet.2012.07.021

[17]. Feily, M. (2009). A Survey of Botnet and Botnet Detection. *2009 Third International Conference on Emerging Security Information, Systems and Technologies*. https://doi.org/10.1109/SECURWARE.2009.48

[18]. Zhaosheng, Z., Zhi, J. F., Guohan, L., Phil, R., Yan, C., & Keesook, H. (2008). Botnet research survey. *2008 32nd Annual IEEE International Computer Software and Applications Conference*, 967–972. https://doi.org/10.1109/COMPSAC.2008.205

[19]. Bailey, M., Cooke, E., Jahanian, F., Xu, Y., & Karir, M. (2009). Al survey of botnet technology and defenses. *Proceedings - Cybersecurity Applications and Technology Conference for Homeland Security, CATCH 2009*, 299–304. https://doi.org/10.1109/CATCH.2009.48

[20]. Li, C., Jiang, W., & Zou, X. (2009). Botnet: Survey and case study. *2009 4th International Conference on Innovative Computing, Information and Control, ICICIC 2009*, 1184–1187. https://doi.org/10.1109/ICICIC.2009.127

[21]. Marupally, P. R., & Paruchuri, V. (2010). Comparative analysis and evaluation of botnet command and control models. *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, 82–89. https://doi.org/10.1109/AINA.2010.171

[22]. Zeidanloo, H. R., & Manaf, A. A. (2009). Botnet command and control mechanisms. *2009 International Conference on Computer and Electrical Engineering, ICCEE 2009*, *1*, 564–568. https://doi.org/10.1109/ICCEE.2009.151

[23]. Dittrich, D., & Dietrich, S. (2008). P2P as botnet command and control: A deeper insight. *3rd International Conference on Malicious and Unwanted Software, MALWARE 2008*, *June*, 41–48. https://doi.org/10.1109/MALWARE.2008.4690856

[24]. Wang, P., Sparks, S., & Cou, C. (2008). An advanced hybrid peer-to-peer botnet. *1st Workshop on Hot Topics in Understanding Botnets*, *7*(2), 2. https://www.usenix.org/legacy/event/hotbots07/tech/full_papers/wang/wang_html/

[25]. Zhang, Z., Lu, B., Liao, P., Liu, C., & Cui, X. (2011). A hierarchical hybrid structure for botnet control and command. *Proceedings - 2011 IEEE International Conference on Computer Science and Automation Engineering, CSAE 2011*, *1*, 483–489. https://doi.org/10.1109/CSAE.2011.5953266

[26]. Perdisci, R., Corona, I., Dagon, D., & Lee, W. (2009). Detecting malicious flux service networks through passive analysis of recursive DNS traces. *Proceedings - Annual Computer Security Applications Conference, ACSAC*, 311–320. https://doi.org/10.1109/ACSAC.2009.36

[27]. Anwar, S., Binti, J., Fadli, M., & Inayat, Z. (2014). A Review Paper on Botnet and Botnet Detection Techniques in Cloud Computing. *ISCI 2014 – IEEE Symposium on Computers & Informatics*, *October 2015*, 28–29

[28]. Yu, X., Dong, X., Yu, G., Qin, Y., & Yue, D. (2010). Data-adaptive clustering analysis for online botnet detection. *Third International Joint Conference on Computational Science and Optimization*, *1*, 456–460. https://doi.org/10.1109/CSO.2010.214

[29]. Livadas, C., Lapsley, D., & Strayer, W. T. (2006). Usilng Machine Learning Technliques to Identify Botnet Traffic. *Proceedings. 2006 31st IEEE Conference on Local Computer Networks*, 967–974. https://doi.org/10.1109/LCN.2006.322210

[30]. Shin, S., Xu, Z., & Gu, G. (2012). EFFORT : Efficient and Effective Bot Malware Detection. *2012 Proceedings IEEE INFOCOM*, *i*. https://doi.org/10.1109/INFCOM.2012.6195713

[31]. Santana, D., Suthaharan, S., & Mohanty, S. (2018). What we learn

from learning - Understanding capabilities and limitations of machine learning in botnet attacks. *ArXiv Preprint*. http://arxiv.org/abs/1805.01333

[32]. Mathur, L., Raheja, M., & Ahlawat, P. (2018). Botnet Detection via mining of network traffic flow. *International Conference on Computational Intelligence and Data Science (ICCIDS 2018)*, *132*, 1668–1677. https://doi.org/10.1016/j.procs.2018.05.137

[33]. Masud, M., Khan, L., & Thuraisingham, B. (2011). *Data mining tools for malware detection*. CRC Press.

[34]. Dua, S., & Du, X. (2016). *Data mining and machine learning in cybersecurity*. CRC press.

[35]. Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, *160*(1), 3-24.

Future of Mobile Application Development in a Post Pandemic World

Thomas C G,
College of Computer Science& Information Science,
Srinivas University,
Pandeshwar, Mangaluru – 575 001
jothomas13@gmail.com

***Abstract*—Mobile Application Development is a continuously evolving Technology that has had an impact on almost everybody on the planet. Right from Kids to Senior Citizens, all are bound to a Mobile Device that has become a part of the person itself. During these COVID times, the usage and the demands of mobile apps have been steadily growing and an array of other factors that have changed the trends of a typical mobile app development. To be informed and aware of what is needed in the market at the current scenario, itself is a vital one for all the mobile app developers. In order to lead in the market by giving a better service and experience to the clients, developers, App Sellers, Content Creators and Entrepreneurs who would like to take their product(s) to the next level, should be looking into continuous research on patterns and finding out what the tech leaders are working on. This paper discusses in detail about various app development trends that will take over the future that we look forward post the COVID pandemic.**

Keywords—COVID-19, Mobile app development trends, post pandemic future, IoT, 5G, Wearable Devices, m-commerce, AR, Block chain

I. INTRODUCTION

Mobile Applications have been impacting our lives constantly from the moment they were introduced via smart phones to us. People spend more time online primarily through mobile phones rather than any other devices. In the recent past, they have been extremely popular and useful for almost everybody in the world. There are a lot of entrepreneurs and industries that have utilized this opportunity to make profit as well as provide considerable service to mankind. Apart from that, we cannot deny that the Mobile App Development Industry is one of the fastest growing industries and it is not expected to slow down anytime in the future. With technology constantly changing, we can be assured that a significant change will be coming to the way we use our Smart Phones and Mobile Applications.

II. FUTURE TRENDS IN MOBILE APPS

We witness a huge investment in the Mobile App industry by tech giants in recent times and even small businesses have understood the dire need of Apps in their business processes. The Mobile App Development industry has been significantly boosted by the dawn of advanced technologies like IoT, AI, ML, AR, VR etc.

Let us see some of these technologies and how they will impact our future post this pandemic in detail.

2.1 Artificial Intelligence & Machine Learning

AI and ML have been in the field of Mobile Apps almost from the beginning but we are just seeing the tip of the iceberg with these advanced technologies. When a person says AI in Mobile automatically, we think of Siri, Alexa or Google Assistant. But the use cases for these technologies go furthermore. AI based applications vary significantly based on for what purpose we intend to use them. It may be any one of the following or more:

1. Speech Recognition
2. Image Recognition
3. Face Detection
4. Text and Image Classification
5. Sentiment Analysis
6. Predictive Analysis

The list might be extended in the near future. It can even be a smart filter that is applied while we take a photo using the camera app or even the adjustment of photo settings before we click. Automatic Face Detection can no longer be called a unique feature, rather it has now become a standard. As we are able to recognize Image, voice etc. and process natural language (still developing, not perfect), bringing in AI powered features in mobile apps will soon become a de facto standard.

During this Pandemic, AI has indeed played a major role. While we are facing a worldwide health crisis, AI has helped to track the spread of the virus, identify high-risk patients, predict mortality risk by analyzing the data available. It can help us fight this virus in various other ways like, population screening, help, suggestions and notifications on infection control. It also has the potential to improve the planning, treatment and outcomes of a patient.

Main Applications of AI during the Pandemic:

1. Early detection and diagnosis of the infection
2. Monitoring the treatment
3. Contact tracing of the individuals
4. Projection of cases and mortality
5. Development of drugs and vaccines
6. Reducing the workload of healthcare workers
7. Prevention of the disease

Having these technologies by our side, we can face this pandemic and move forward into a normal world, sooner than we would have without the help of AI & ML. The possibilities for Mobile Applications that apply AI in custom App Development are limitless. AI not only makes the apps we use smarter, but also improves the performance of that app at every level. Let it be backend development or frontend UI/UX, AI will surely change the way of App Development and will continue to be an inevitable part of it.

2.2 Blockchain and Decentralized apps

Blockchain is much known for being the technology that is running Crypto Currency, like Bitcoin. A blockchain created a decentralized database that cannot be hacked and is fraud-resistant. It is used for securing payments, accessing networks by generating tokens used in authentication and it is not possible to modify the database to get access. Even though it is commonly seen from a crypto currency perspective, tech giants are aiming to increase blockchain in the enterprise sector. Custom Mobile applications that are blockchain based are already kick started. DApp or a Decentralized Mobile App is an App that is owned by no one, and almost not possible to shut down nor have a downtime. It is expected that Mobile App Industry will also make the App decentralized, and do what Bitcoin did with money.

Blockchain Technology has helped to tackle the current pandemic by simplifying clinical trial processes for vaccines and drugs, raise public awareness and have all donations/fundraisings be transparent and a reliable data tracker. It is quintessential for us to have a platform to track this virus transmission as many of the current systems are vulnerable.

Use Cases of Blockchain Technology during COVID:

1. Clinical Trial Management
2. Medical Supply Chain
3. User Privacy Protection
4. Data Aggregation
5. Contact Tracing
6. Donation Tracking
7. Outbreak Tracking

Blockchain based tracking systems will validate COVID Data from varied sources and help us to moderate the spread of false data and modified data. WHO is using Blockchain technologies to convey data about the ongoing pandemic, called MiPasa. This uses Distributed Ledger Technology, that will help with early detection and identifies carriers and notifies hotspots. It is said to be a fully private app, that shares information with only authorized personnel like health officials. Post this Pandemic, organizations will take advantage of blockchain in so many use cases and use it to speed up and improve their products and services

2.3 Internet of Things

IoT is a network of physical devices that are embedded with sensors, electronics and software which are interconnected. It was a strategy working well with production and heavy machines, but recently it has extended to us as smart watches, smart phones, smart home devices etc. It is not a new concept but the outpour of mobile users around the world and across sectors have created ample amount of opportunities. We are now accustomed in using this technology to improve our day-to-day life. Almost half the devices in our home are now controlled by us via our Smartphones. We can adjust the AC, Camera, Fridge, Washing Machines, Ovens, Stoves, Door Locks, etc. of our Home with an app on our device, remotely from anywhere.

During the pandemic, when we turned to technology for our needs, little did we know that the impact of those technologies will continue beyond COVID. Drones and Contactless payments that help our online shopping, especially for food and medicines has reached places that would have taken time to get used to technologies. IoT enabled Automation systems helped workers of manufacturers and businesses to deliver services. When a pandemic like this comes, IoT based protocols like BLE, NFC, RFID, GPS, Wi-Fi are there to provide solutions to the challenge of early detection, isolating and contact tracing. During these times, we are in need of this type of communication to safeguard us from the spread of the pandemic. The following are some of the ways in which IoT has contributed

1. Telehealth Consultations
2. Digital Diagnostics
3. Remote Monitoring
4. Robot Assistance

These have changed the way people have been interacting with systems. A lot of things which we were not sure of whether it will happen or not has happened within a short span of time. It is evident that IoT will continue to dominate the market and also it is THE future of App Development. Almost all the apps would have to integrate the smart features and it would not just be a feature but a mandatory one. Moreover, the mobile world will be forced into cross-platform development.

2.4 Augmented Reality/Virtual Reality

AR/VR were popular in gaming platforms. However, it is not the case anymore. There are a lot of use cases for AR/VR in our ever-changing world and Tech Giants are already innovating on this front. We have seen the use of AR apps that help us decide what furniture would fit best in our rooms, what dress would look good on us, which lens frame would fit my face well, etc. These are just the beginning for a plethora of apps are yet to be released with AR features. While Virtual Reality takes us into another world, Augmented Reality enhances our experience in our own world. It has a great potential to be used in various industries such as, Mobile AR disruption, AR in marketing & advertising, healthcare and manufacturing.

AR/VR during this pandemic made us feel closer to one another, took us beyond the limit of social distancing and connected us to our colleagues and the entire world. While the pandemic made our work from home challenging, the benefits of these technologies becomes clearer. Even before the pandemic, people from this field have demonstrated their value across industries such as Health Care, Education,

Media, etc. It is high time that we look up to AR/VR to see how they are helping our world. Educational tools that are powered by AR/VR are taken by a large range of Educational Institutions around the world. They can support students remotely, take classes in a Virtual Room, etc. A lot of Historical Museums such as the Louvre, The Vatican Museum, etc all offer a Virtual Tour of their collections. As the days goes by, the use of AR/VR will become a Legitimate Study Tool and also a valued Educational Resource. The areas in which AR/VR can be a valued asset are:

1. Healthcare
2. Education
3. Banking, Financial Services & Insurance
4. Employee engagement in organizations
5. Field service management

We know that after this pandemic is over, things will never be the same again and people are getting more tech savvy by the day. AR/VR will also be one of those tools that we will not say adieu any time soon. Even if not all fields adopt this technology, definitely health care and education will take it forward. AR adaption is one of the top trends in app development in the current scenario.

2.5 5G Technology

5G Technology is the next level in Mobile Technology after 4G/LTE. It is expected to have 10 times decrease in latency and a boosted network efficiency and traffic capacity. It offers transfer speed that is 100 times faster and it is expected to boost up the level of Internet based apps. App Developers must be aware of the upcoming rise in speed and upscale their products to match those speeds. It is expected that there will be a lot more people would be using 5G in the upcoming year and it will sky rocket soon. It will ultimately increase the functionality of all mobile apps that will allow the developer to add new features without affecting the performance of the app. It is expected that the developers and app resellers should have a 5G network while testing and developing.

While 5G might not be reaching us as fast as we expect, the increase in virtual work and online communications demand that we need a better network coverage and advanced internet connections, even in remote places for our work from home to be a success. There will be a huge demand for devices such as mobile phones, laptops to collaborate and work from anywhere. This technology might also lead us to a better video conferencing for work or education. It is completely essential that the service providers must create the right 5G experiences and COVID is just the push that the network operators needed to move their experimental phase soon to product phase.

2.6 Wearable & Foldable Devices

Wearable Devices have steadily been on the rise ever since its inception. It is not only a considerable amount invested in the wearable market, in the near future, wearables will be as how we refer to a smart phone today. Even though all wearables now are revolving around the smart phone which needs to be in its proximity, the future holds wearables that are directly connected to the internet and are self-sufficient.

One of the hottest trends is to develop wearable smart devices and apps for them, because of their role in Health Care. This Pandemic has also brought to light that smart watches and other wearable devices can be used to study and track the health parameters and also data that supports the diagnostic process. Even though the complete potential of these wearable devices with mobile applications are not been unleashed, we already know that they are really helpful for medical and fitness regimes. These kinds of devices are also used in AR. Smart Glasses are already available in the market, which will be used for navigation, fixing mechanical problems or piloting a drone.

We all have the idea of an age-old flip phone, without even the concept of touch being introduced. Those foldable devices have started trending once again. They fold to compress or expand the screen size based on our preference. Even though it might seem like a great feature for the end user, from the perspective of an App Developer, It is vital that all these are also kept in mind while developing or updating an App. The coming years after this pandemic there might be a huge demand for foldables, which means app developers must plan accordingly.

COVID has given a complete boost for Tech as well as Life Science Companies, in spite of the recession that this pandemic caused, companies that are in the manufacturing of ventilators and other breath support devices have got a huge surge in the demand. Same goes for Wearable Device Manufacturers, because the number of customers who started using a wearable device to monitor themselves for COVID symptoms and also for contact tracing and social distancing. It is right to say that wearable devices will have their market even post this pandemic. The following are some of the wearable devices that have been aiding the health workers to tackle this pandemic:

1. OURA: A Smart Ring to Detect Early Covid-19 Symptoms
2. VITALPATCH: Monitoring Covid-19 Patient Drug Response
3. SAFESPACE: Social Distancing for the Workplace
4. TRACETOGETHER TOKENS: Singapore's Contact Tracing Tech
5. PULSE: Open-Sourced Anti-Face- Touching Tech

It is safe to say that wearables are going to take the App Development paradigm to another level and will be leading as the devices are offered in a much cheaper rate.

2.7 Beacon Technology

Beacons send signals to other smart devices that are nearby using small wireless transmitters that use Low Energy Bluetooth Technology. They are a significant development in location-based technology and also proximity marketing. The Beacon is a simple device that has a CPU, Radio and batteries and it repeatedly broadcasts a unique identifier, which is usually picked up by a mobile and marks a specific place in our environment. Once it has been connected, it will do the

function that has been programmed. Beacons is not a new technology, but they are trending as they enable the connection between online and offline worlds. They are used to understand customer behavior and finding out patterns in movements. It helps the retailer to optimize the location of the products for a better sale. Beacon Technology is cheaper compared to other viable technologies. Its market is said to be in Travel, Tourism, Real estate, Retail, Health Care, Education and Financial Institutions. It provides better student and teacher relation with powerful communication tools.

Compared to other pandemics we have faced, a significant advantage that we have over this COVID crisis is that we have a plethora of advanced technologies to make a solution to this situation more efficient and real-time. Beacon Technology has been most effective on preventing the rise of COVID. It is indeed difficult to do contact tracing to track where the users have visited, like how they travelled, which path they took, etc. The solution was to convert your smartphone into a beacon and it makes the geo-fencing of prone and safe areas possible. In the future, after this pandemic ends, Beacon technology will still enhance and be utilized for various of the aforementioned purposes and apps need to leverage this technology to conquer the market.

2.8 Mobile Wallet & Mobile Commerce

Almost every industry has been leveraging Mobile apps to increase their revenue since there is a lot of money to be made in this arena. Mobile e-commerce is an attractive feature and almost every company is launching an app to increase sales. It has come to a level where we can say that you need an m-commerce app to stay in the business. Every simple business to tech giants have taken leverage of the fact that it is essential for an app to be there to make their brands successful. With Mobile Payment Solution App Development on the rise, with security and transaction behavior among concerns, the usage of a Mobile Wallet is steadily on the rise. A hassle-free payment method is something that all customers want to see in their frequently used mobile apps.

During the COVID crisis, a lot of retailers were forced to go digital and make their income through online platforms. We can see the transition of people taking up mobile wallets like PayTM, Amazon Pay, etc. have been hiking exponentially. The stores which were strictly cash only have changed their perspective to go Online and use an App to scan and pay for their products. Right from a simple grocery shop to a portable sales vendor, all have started using an App to collect and monitor their sales and income. Even post this pandemic, the use of Mobile Apps to pay for our purchases are not going to go away and are here to stay. Hence it is essential for an App Developer to always be ready to integrate and use any of the Mobile Commerce Gateways or even a complete Mobile Wallet. It will be the future of any and all sales that is going to happen online.

Currently more than 70% of total ecommerce sales comes from Mobile devices. Apps play a vital role in the current and future of mobile commerce. Post 2020, Mobile wallets and implementation of a Payment Gateway, that offers highest level of secured encryption will become indispensable in all kinds of mobile apps.

III. CONCLUSION:

COVID has turned everyone's lives upside down in which Mobile Apps have appeared as a savior that ensures a brighter future ahead. Mobiles are becoming an absolutely essential tool for a lot of things including work. It's quite evident from what we all have gone through via remote work, education, social media, healthcare, financial, retail, etc., apps have kept a number of businesses running during the crisis and satisfied customers' demands. The current pandemic has forced every business to re-consider outdated communication tactics, technologies, and systems. Even smaller businesses are taking up mobile apps and in a world that was pushed into quarantine, apps are no longer considered as a cost, rather as an investment. We need to realize that it is a meager investment that gives bountiful returns. Integrating these modern trends & technologies into mobile apps can take customer comfort to new vistas. This pandemic has frozen digitalization and the need to digitize has grown immensely. We are going to witness a lot more businesses with a mobile app like never before. If they would consider the recent trends discussed in this paper, before developing an App, we can achieve greater heights and take our pandemic hit world into a new normal world.

REFERENCES

[1] Raju Vaishya, Mohd Javaid, Ibrahim Haleem Khan, Abid Haleem, "Artificial Intelligence (AI) applications for COVID-19 pandemic", Diabetes & Metabolic Syndrome: Clinical Research & Reviews, Volume 14, Issue 4, July–August 2020, Pages 337-339

[2] Marbouh, D., Abbasi, T., Maasmi, F. et al. Blockchain for COVID-19: Review, Opportunities, and a Trusted Tracking System. Arab J Sci Eng (2020). https://doi.org/10.1007/s13369-020-04950-4

[3] WHO. Rolling updates on coronavirus disease (COVID-19). World Health Organization. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/events-as-they-happen. Accessed 25 May 2020

[4] Wearable Wonderland: How tech is tackling COVID-19. Medical Technology, Issue 32, October 2020. Accessed 10 November, 2020

[5] A. Haleem, M. Javaid, Vaishya, "Effects of COVID 19 pandemic in daily life" Curr Med Res Pract (2020), 10.1016/j.cmrp.2020.03.011

[6] D.S.Ting, L.Carin, V.Dzau, T.Y. Wong, "Digital technology and COVID-19" Nat Med (Mar 2020), pp 1-3

SRINIVAS UNIVERSITY

Srinivas Nagar, Mukka– 574 146, Surathkal, Mangalore, Phone: 0824-2477456
(Private University Established by Karnataka Govt. ACT No.42 of 2013, Recognized by UGC, New Delhi & Member of Association of Indian Universities, New Delhi)
Web: www.srinivasuniversity.ac.in, Email: info@srinivasuniversity.edu.in
Administrative Office: GHS Road, Mangalore-01, Phone 0824-2425966

Srinivas University IEEE SU SB SAMAGRA GNANA STB11770 IEEE Student Branch

IEEE Mangalore Sub Section

Srinivas Publication
(A Unit of Srinivas University)

978-81-944670-1-4