iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** X **Month of publication:** October 2022

DOI: https://doi.org/10.22214/ijraset.2022.47129

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue X Oct 2022- Available at www.ijraset.com

A Paradigm Shift in IoT Cyber-Security: A Systematic Review

Prateek Verma[1], Rishabh Sharma[2], Maunil Mistry[3]
[1, 2]*B.Tech Student, Department of Electronics and Communications, Maharaja Agrasen Institute of Technology (GGSIPU)*
[3]*B.Tech Student, Department of Electronics and Communications, Vishwakarma Government Engineering College*

Abstract: ***Internet of Things (IoT) describes physical gadgets with sensors, processing power, software and other technologies that connect with other devices and exchange data over the internet. Consumers and businesses everywhere are getting used to the notion that critical data processing and analysis are now being fulfilled by a new-age technology infrastructure. The Internet of Things (IoT) is the key driver of this trend, with 64 billion IoT devices estimated around the world by 2026. IoT activities span everything from smart homes and smart cities to fitness monitoring and inventory control. This rapid growth of IoT has its own set of challenges and the first one being the threat of cybersecurity. As the Internet of Things rapidly grows, cyber threats and associated risks continue to evolve and become increasingly complex. The biggest threat to these IoT devices, which has grown exponentially in the past few years, is the Distributed Denial of Service (DDoS). This highly daunting cyber-threat is extremely dangerous because perpetrators can now use IoT devices to make the attacks more severe. The formation of botnet clusters (hijacked IoT devices) is the heart of a large-scale DDoS cyber-attack. This research work focuses on the emerging trends associated with the volume of cyber-attacks. Furthermore, the research delves into the possible solutions to tackle, and possibly eradicate the cyber-threats present in the IoT systems. The domain of Machine Learning is investigated thoroughly in order to extract the ongoing research and possible solutions pertaining to the cyber-security of IoT devices and systems.***
Keywords: Internet of Things (IoT), Security, DDoS, False Data Injection, Machine Learning, Botnet

I. INTRODUCTION

The Internet of Things (IoT) describes the physical devices or electronics that are capable of connecting to the internet and having access to the same. A complete IoT system consists of four different components: sensors/devices, connectivity, data fetching and processing, and user interface [1]. The need of these IoT devices was to exchange data over the internet, either by transmitting information or receiving information, thus leading to the development of this technology. This also resulted in a major revolution as wired connections were significantly replaced by these connected devices. A major advantage of these 'internet' enabled devices is the seamless communication between people, processes, and devices. Due to the advent of low-cost computing, the cloud, big data analytics and cost-efficient devices, IoT systems can continuously collect and store data with minimal human intervention [2]. However, every technology has its own set of limitations and drawbacks. Since these IoT devices are constantly connected to the internet, they are extremely vulnerable to exploitation and have a constant threat to security. Due to power supply constraints, these IoT devices are generally 'headless', without any overhead security features or additional software that might tackle any future security threats. This limitation did not matter in traditional devices or networks because they were completely isolated without any kind of access to the outside world. However, due to the advancement of technology, the interconnectedness of the IoT ecosystem has grown explosively. At present, more than 10 billion Internet-enabled devices are in use with at least one active endpoint [3-6].

This novel technology has made these IoT devices and industrial IoT devices a prime target for cyber criminals. Most internet-enabled devices are not designed with security in mind, and the traditional operating systems do not possess enough processing power to incorporate and carry out the additional security features. As a result, a significant increase in the cyber-attacks especially to these IoT systems and IoT networks have been observed [7-8]. These attacks can be extremely devastating for an organization since these devices can be remotely controlled or have their functionality disabled by exploiters and hackers. These IoT systems are vulnerable to hijacking which can be further used for weaponization. There have been multiple large-scale attacks pertaining to the domain of IoT in the past. These massive security threats led to over-utilization of both resources and time, which caused major operational hindrances to various organizations and industries. In August 2018, a petrochemical plant located in Saudi Arabia was hit by a deadly cyber-assault in the form of 'False Data Injection' attack. This deadly attack was carried out by hacking into the industrial network and remotely gaining access to an engineering workstation which was further infected by a customized malware [9-13].

This malware was responsible for reprogramming the SIS (Safety Instrumentation System) controllers targeted for industrial use. The attack was detected in August when plant's machinery began shutting down randomly during the working hours. The malware took down the entire machinery by operating it outside of its normal parameters until it suffered serious damage [14]. The wear and tear triggered an explosion which caused physical damage to the plant's infrastructure. This malware was specifically designed to override the safety systems, in an overt attempt to cause catastrophic damage to the entire petrochemical plant. A key challenge in building security protocols for IoT is that there is a lack of standards available, due to the complexity of the IoT ecosystem and a huge number of devices supplied by a wide range of vendors. Strong cyber protection of IoT networks starts with trained cybersecurity experts who are proficient in protecting infrastructure, securing data and information, running risk analysis and mitigation, architecting cloud-based security, and achieving compliance [15-16].

The security challenges pertaining to IoT and its associated systems are tough to detect and even tougher to eradicate completely [17],[18]. However, multitudes of research work have been carried out in this field which have contributed to the positive developments in the past. Internet of Things (IoT) is a state-of-the-art concept just at the beginning of its inception, thus the traditional systems designed for internet security would not hold relevant for this technology [19-22]. Therefore, an emerging and in-demand field is being explored in order to attain the security associated with IoT. The field of machine learning is being explored intensively, with new and research methodologies adding to the safety and security of these interconnected devices. Machine learning (ML) refers to a type of Artificial Intelligence (AI) used to optimize performance criteria using example data or past experience through learning [23]. More precisely, ML algorithms build models of behaviours using mathematical techniques by utilising huge datasets. These algorithms also enables the ability to learn without being explicitly programmed. Furthermore, these algorithms use historical data as an input to predict the new output values. Machine learning is a promising technology in the field of IoT for several reasons, e.g., IoT systems produce a sheer amount of data which is required by ML which further enable the IoT systems to make informed and intelligent decisions [24-27]. ML also reduces the processing of vast amount of data every time a security breach occurs by utilising the past data that has been already processed and accounted for in the algorithm. ML is widely used for security, privacy, attack detection, and malware analysis. The ML algorithms are majorly classified into four categories: supervised, unsupervised, semi-supervised, and reinforcement learning [28].

The main contributions of this paper can be summarised in the following pointers:

1) The work takes a deep dive into the trends and patterns which led to the development of this extensive survey. The adoption and penetration of ML techniques along with the rising threats to the IoT systems is presented thoroughly.
2) The work presents an in-depth, systematic and comprehensive survey pertaining to the role of machine learning mechanisms in IoT systems and devices.
3) The work describe state-of-the-art results on Machine Learning (ML) techniques in IoT systems and networks with a specialised focus on security and privacy of the IoT systems.
4) The work focuses on the ML techniques that have replaced the traditional security measures associated with the IoT systems. The in-depth review of different research works along with the present challenges pertaining to the field of ML-IoT is also discussed.

This work has attempted to gather all the scattered research undertakings, present them in a systematic manner and reflect upon the important inferences that were generated along the way. A vast array research works were handpicked and filtered accordingly, in order to best fit in this work. The rest of the paper is divided and organized into six sections. Section II deals with the research methodology which further expands upon the selection and filtration of relevant and important research papers. Furthermore, Section III sheds light upon the altering patterns and emerging trends pertaining to the domain cyber-security in IoT systems. The Section IV puts an emphasis on the DoS/DDoS attacks while also discussing the existing solutions and future works. Similarly, Section V deals with the False Data Injection (FDI) attacks and takes a deep dive into the existing works.

II. RESEARCH METHODOLOGY

The entire study revolves around the discussion of multiple high-standard research papers pertaining to the domain of 'Security in Internet of Things', which were further used for developing inferences and drawing conclusions.

A. Initial Selection

The amount of research papers selected for the work was around 150, out of which only 50 publications were incorporated into the study. These distinguished research papers were selected exclusively by the extensive usage of various keywords. The keywords coupled with manual selection led to a unique collection of high-quality research papers.

Keywords
- Security challenges in IoT systems
- Denial of Service (DOS) in IoT
- Machine Learning and IoT Security Survey
- ML-based intrusion detection system for IoT
- ML methodologies to prevent FDI in IoT systems
- Machine learning for security of IoT systems
- False Data Injection (FDI) in IoT systems
- ML techniques to prevent DOS attacks in IoT

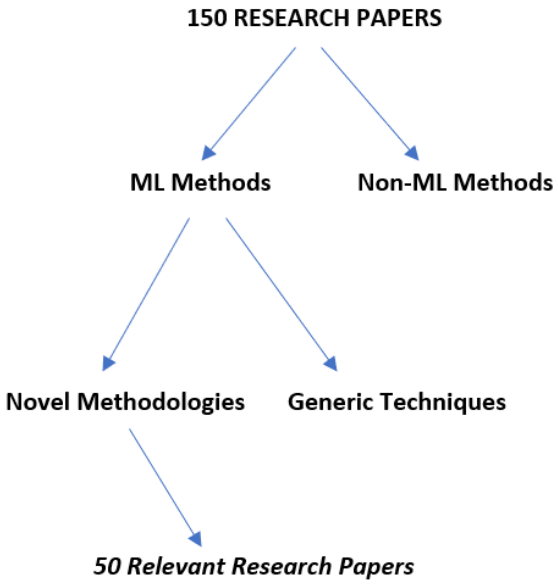


Fig. 1. Classification of Papers

B. Filtration

As shown in Fig,1, the selected research works were filtered in such a way that only those papers were included in the study which were associated with machine learning techniques in order to curb the challenges pertaining to the domain of IoT systems. On the other hand, the papers containing non-ML methods and generic techniques were completely discarded from this work.

C. Relevant Work

After the process of filtration, 50 relevant research papers were obtained. These filtered research works contain both the machine learning techniques along with the novel methodologies associated with them. These novel research methods will be studied further and discussed in this case study. The research works that are accounted into this case study belong to the time period of 2016 – 2022. Furthermore, since this case study also sheds light onto the past cyber-attacks, multiple sources and an extensive set of survey papers were also taken into consideration.

III. EMERGING TRENDS

Since the inception of this technology, Internet of Things (IoT) devices and systems have been under a constant threat of getting breached and exploited [29]. These devices have been subjected to a variety of different cyber-attacks. Furthermore, the attacks carried out on these systems can be categorically summarized. These cyber-attacks are classified under four broad categories: Eavesdropping, Man-in-the-Middle, Routing, Denial of Service [30].

1) *Eavesdropping:* This type of attack occurs when a hacker incepts the data that is being transmitted between two devices. This allows the exploiter to modify or delete the transmitted data. Eavesdropping, also known as sniffing or snooping, relies on unsecured network communications. For instance, an unsecured network affecting multiple baby monitors could expose hundreds of thousands of live devices, potentially allowing a hacker to drop in view a camera's video stream.

2) *Man-in-the-Middle:* The concept of this attack is that am hacker intercepts a communication between two systems or devices. This is a dangerous attack because the attacker poses as the original sender. As the attacker has inserted itself into the original communication, it can trick the recipient into thinking that a legitimate stream of data is being received. For instance, the healthcare records are stored electronically which can be shared among patients and doctors. A cyber criminal can gain access to those data repositories and inject false data to misled the diagnosis and treatment
3) *Distributed Denial of Service:* This is a cyber attack in which an attacker attempts to flood a server with internet traffic to prevent users from accessing connected online services and websites. DDoS attack aims to overwhelm the devices, services and network of its intended target with fake and illegitimate traffic, rendering those services inaccessible and useless to the genuine users. The sheer volume of the fake requests and pings makes a DDoS attack much harder to mitigate. The most recent and the largest attack in history occurred in February 2020 to none other than Amazon Web Services (AWS). The ramifications include a sharp decline in legitimate traffic followed by lost business and damage to reputation.

This study is extremely important from the standpoint of security threats to the IoT systems. After analyzing the selected research papers and other survey papers that are associated with this field, intriguing and unique conclusions were obtained. The obtained inferences display a shift in the trend of attacks that are being carried out on the IoT systems and devices. The time period that is associated with this case study spans over 17 years (2005-2022). These 17 years are broken down into clusters of 5-7 years such that conclusions can be determined with ease.

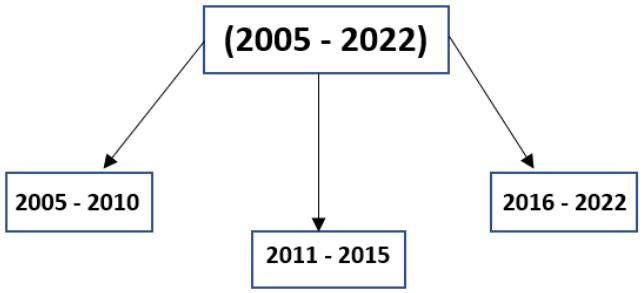


Fig. 2. Classification of Time Periods

According to Fig. 2, a time span of 17 years has been categorized into three different time periods. This has been carried out for the ease of conducting the case study. With the increasing rate of development related to the IoT systems and a widespread application of this technology, certain types of security threats have risen dramatically in the past five years. The majority of attacks that were carried out on these systems were related to the Denial of Service (DoS) and Man-in-the-Middle, specifically the False Data Injection (FDI) attacks. Due to the ease of carrying out these attacks, these fatal security threats are gaining traction in the domain of Internet of Things. FDI and Dos/DDoS attacks are extremely deadly as they can disrupt the entire networking capabilities of the IoT systems. A distributed Denial of Service (DDoS) attack is a subcategory of the general Denial-of-Service (DoS) security threat. In a DoS attack, the hacker uses a single internet connection to barrage a target with illegitimate requests, with the objective of causing downtime. However, DDoS is larger in scale as it utilizes thousands (even millions) of connected devices to fulfill the goal. Furthermore, even a short downtime can induce havoc on the entire operation and additional resources would be required to restore and reboot the associated systems.

As observed in Fig. 3, the solution landscape pertaining to the field of IoT cyber security has dramatically changed in the previous seventeen years. During the early phase (2005-2010), the traditional solutions showcased a dominant share in the total number of solutions existing to mitigate the cyber-threats. These traditional methods accounted for a variety of conventional solutions such as Honeypots, Detection Systems, Firewalls, Manual Intervention, etc. Furthermore, as Internet of Things progressed in the following years, the ML/DL methods took up almost half of the volume share of the existing solutions. By the end of 2021, traditional methods and techniques were pushed back to almost a quarter of the existing share of cyber solutions. On the other hand, Machine Learning and Deep Learning methods were bring picked up and these techniques were used extensively to solve the threats in the field of Internet of Things. The machine learning methods are comparatively very efficient to the traditional methods because they can both detect and control the attack without any human intervention.

The ML approach solves the issue of identifying unknown devices on a network and protecting the IoT infrastructure by automating the scanning and management of IoT devices across the entire network. These techniques are capable of scanning all the connected devices and shutting down ongoing attacks automatically before the IT teams are aware of it. In the case of known vulnerabilities, such as DDoS, ML compares the present network behavior with behavior pattern from the existing history of attacks and takes protective action.

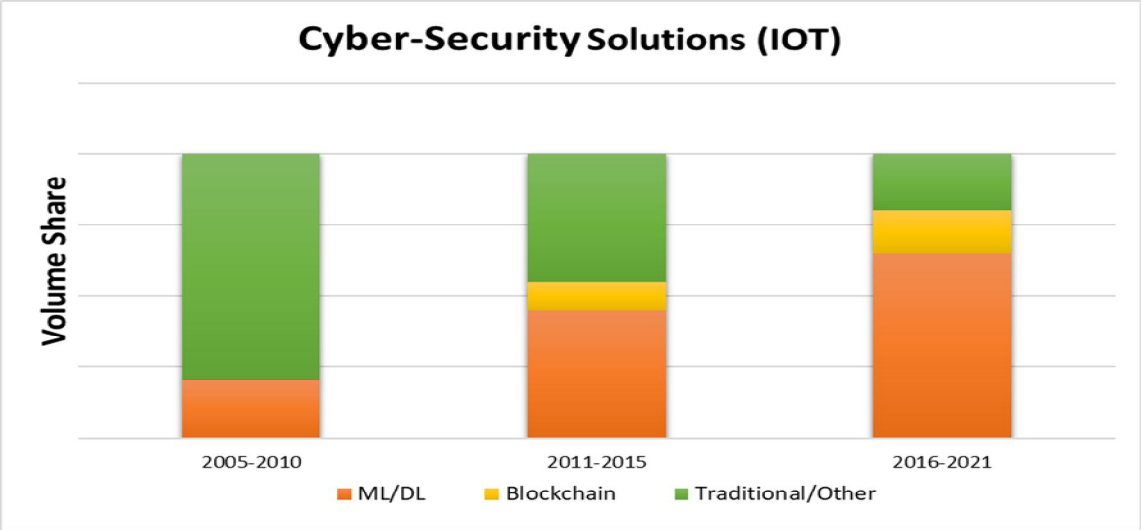


Fig. 3. Trend of Security Threats (IoT)

As presented in Fig. 4, the inferences obtained after the initial analysis display a sharp increase in the Denial of Service (DoS) attacks. As observed in the visual, the quantity or the volume of attacks has increased multi-folds in the past seventeen years. As a result, the cyber-security solutions for IoT systems have also increased in number. This sudden increase can be associated with the availability of botnet clusters which can be accessed and controlled remotely. A botnet is a network of computers and devices infested with a certain malware this is under the control of a single attacking party. The scale of a botnet clusters (millions of bots) enable the attacker to perform large-scale operations that were previously impossible with malware. The nodes of these IoT systems, such as the sensors attached to the end of an IoT system or even a smart bulb which is connected to the internet, can be hacked and controlled over the internet. Since DDoS and FDI attacks account for more than fifty percent of the total cyber-threats (Fig 4.), these threats will be studied in a detailed fashion such that mitigation of these attacks can be achieved. If these threats can be eradicated, the domain of cyber security (IoT) will become highly secure and extremely easier to operate. Furthermore, if these existing threats (FDIA, DoS/DDoS) are mitigated, the field of cyber-security can further evolve and focus on the other underlying problems.

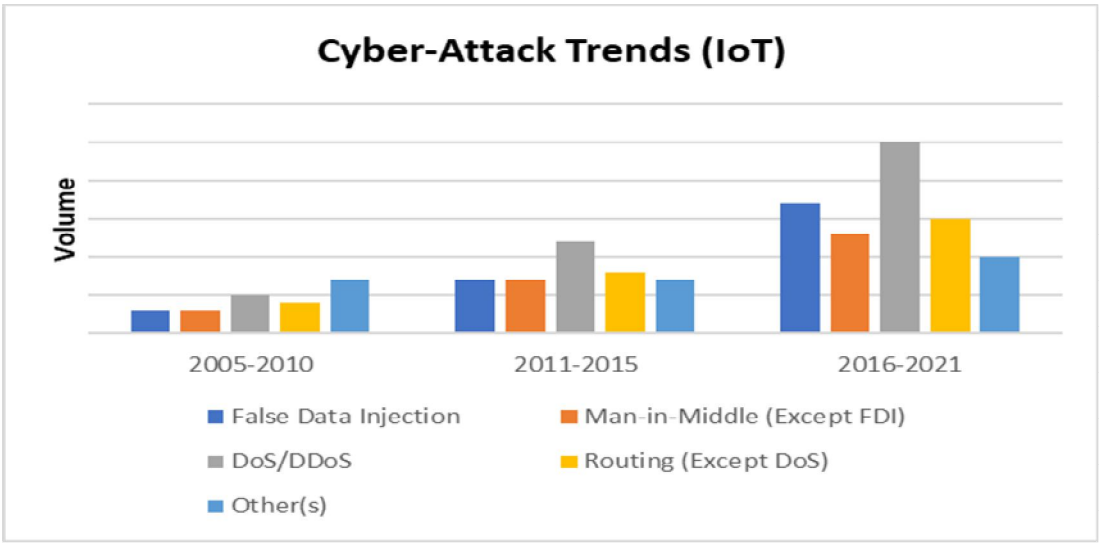


Fig. 4. Trend of Security Threats (IoT)

IV. DISTRIBUTED DENIAL OF SERVICE (DDOS)

Denial of Service (DOS) attack is targeted at shutting down a machine or network, making it inaccessible for the intended audience. They accomplish this task by flooding the target with a huge volume of traffic, or sending information that initiates a crash [32,35]. This extremely dangerous and highly sophisticated cyber threat is prevalent in the domain of Internet of things (IoT) as well. Thus, suppression and further eradication of this cyber threat is the need of the hour. In this section, different machine learning models have been studied which are further discussed in depth [38-40]. The motivation behind using machine learning solutions lies in the performance metrics associated with the deployed models. These performance metrics provide a standard which can be further used to compare different ML models in order to determine the efficiency and accuracy [42].

A. Performance Metrics

1) *Accuracy* – It refers to the percentage of correctness of the ML model
2) *Precision* – It refers to the percentage of the results which are relevant
3) *Recall* – It refers to the percentage of total relevant results correctly classified by the algorithm
4) *F1 Score* – It elegantly sums up the predictive performance by combining two competing metrics: precision and recall

In [33], a honeypot approach has been utilised which uses machine learning methods for malware detection. In the proposed solution, a framework that is based on honeypot technology has been developed. This framework catches several malware installations attempts (log files) into the IoT network. These log files were then used as input for the devised machine learning model. This solution has a dual focus: detection of malware responsible for DDoS attacks and categorizing the malware into different family classes. The proposed solution is present only in theory and the actual development needs to be carried out in real-time. This solution is inspired by ThingPot, a virtual IoT honeypot capable of emulating different IoT communication protocols and proper IoT environment. Due to the limited overhead in IoT networks, the ML classifiers can be applied only on the router level. The transformation of log files into the required format can be carried out using bash scripts on Linux. For carrying out the machine learning model, ML tools such as Microsoft Azure, MATLAB, etc. These approaches can be utilised for real-time machine learning detection framework.

In [34], the authors have proposed a method that converts the network traffic data into a visual image format, which was further used to train a CNN model. This particular model is used because it works efficiently with computer vision datasets and visual formats. During the pre-processing phase, data cleansing was carried out with conversion of cleaned data into three-channel images. After training and splitting the test data, ResNet18 model was used which consists of 18 layers (10 convolution layers, 8 pooling layers). For building the required model, different parameters need to be set than the original ResNet model. After the trained model with best accuracy is saved, the performance of the proposed technology is measured using the performance metrics. These metrics include: Precision, Recall, Accuracy and F1-Score. In the case of binary classification, the accuracy rate for detecting the DoS and DDoS was 99.99%. Furthermore, in the case of multi-class classification, the proposed method achieved an accuracy of 87.06%. This unique and novel methodology beats the accuracy rates of a state-of-the-art solution for detecting these types of attacks. This CNN model exhibited 9% more precision as compared to the state-of-the-art methodology. On top of that, this model achieved a 21% higher Recall (true-positive rate) and 17% higher F1-Score.

As discussed in [36], the research work proposed an efficient ML model that consisted of SVM, KNN, and RF for the classification of DDoS attacks using a Banking Dataset. Multiple training parameters such as classification accuracy (%), precision (%), recall (%), and F1-score (%) were taken into consideration for measuring the performance of different ML models. The most efficient ML model is analysed and accounted for in this research. In this study, the raw data is obtained from an open-source platform. After carrying out the pre-processing, the data was split into 70% training set and 30% testing set. The dataset is pre-processed for increasing the relevancy for the ML classifier. Data pre-processing techniques include (i) Removal of Socket Information, (ii) Removal of White Spaces and (iii) Data Normalization. The results section demonstrates the performance of each model when applied to a certain dataset. Three ML models were utilised in the banking dataset for enabling a fraud detection model. Comparative results indicate that SVM (Support Vector Machine) has achieved the best time complexity with 99.8% accuracy, 99.07% precision, 98.32% recall and F1-score of 98.5%. On the other hand, the other two models displayed lower accuracy levels: 'Random Forest' with an accuracy of 97.5%, 'K-Nearest Neighbour' model with an accuracy of 98.74%.

In [37], a ML-based detection system is developed that is responsible for recognition of different DDoS cyber threats. In this work, a heterogenous gateway has been achieved to gather sensory data to propose a machine learning-based DDoS attack detection system for tackling the malicious data passing through the IoT gateway.

The features of different DDoS – ICMP flood, SYN flood, UDP flood have been studied. Decision trees have been utilised for training and real-time detection of the attacks in the proposed IoT system. These decision trees detect whether the passing packets are abnormal or benign. The IoT gateway captures the sensor data packets for pre-processing which would be fed into the training data of the ML model. This research work has been implemented in real-time. For a realistic experiment, eight smart poles were equipped with a Raspberry Pi 3 having internet capabilities. After collecting the sensory dataset, the collected data is split into training samples (70%) and testing samples (30%). In conclusion, the decision trees model has achieved over 97% in both the accuracy and F1-score. The real-time detection model is highly efficient as it detects a DDoS attack packet in less than 0.5 seconds. A SDN controller is also incorporated into the work which is responsible for storing blacklisted MAC and IP addresses. Whenever, abnormal packets are detected, the IP addresses and MAC addresses of malicious devices are sent to the SDN controller.

In this research work [41], a novel hybrid Intrusion Detection System (IDS) for detection of DDoS attacks has been developed. This novel IDS was compared with multiple state-of-the-art algorithms and ML methods which are already utilised in the domain of cybersecurity. The methodology of the proposed IDS incorporates a sorting algorithm that is used for feature selection. Furthermore, the output data from the sorting algorithm is fed into the deep learning model for classification of the DDOS attacks. This deep learning model utilizes a CNN layer followed by a max pooling layer, which is succeeded by a LSTM layer. During the experimentation phase of the work, the CICIDS2017 dataset has been employed for the study work. This dataset contains the latest data network attacks emulating the real-world network data. The dataset is split into training and test sample in the ratio of 90:10. According to the work, the set of fifteen number of features obtained has achieved the best values of accuracy (98.78%), precision (99.03%), recall (99.35%) and F1-Score (98.48%). When this novel methodology is compared to other ML models, the precision value of the proposed model outperforms the rest of the models.

In [46], the authors have devised two methods with very distinct features. The first research provides a solution by implementing a Hybrid-IDS (Intrusion Detection System) which takes into account the features of both an anomaly-based IDS and a signature-based IDS. This system works on a dual nature, for instance, if an attack passed the IDS network sensors without detection (in case it is a new IP), and reached the signature-based detector without detection, then it does not match any of the signature-based attacks stored in (KAS-DB), the behaviour of the attack is already traced and carefully monitored by the anomaly-based detector. Furthermore, if an abnormal behaviour is detected then the IP address will be blocked. The second research revolves around the usage of a deep learning model to tackle the DDoS attack. This method utilizes a LSTM neural network because this neural network is capable of solving the gradients problem. Three different training models are used in the research work with the dataset being split into 80% training set and 20% test set. The first used model consists of one LSTM layer, the second used model consists of two LSTM layers and the final used model consists of three LSTM layers. The result from the first model displays an accuracy of 91.54%, furthermore the second model displays an accuracy of 96.74%. The final LSTM model displays the highest accuracy of around 99.19%.

In [47], the authors have described a Deep Learning (DL) based botnet attack detection that relies on the network flow. To determine an efficient DL model, multiple experiments were conducted on standard benchmark data sets. The main objective of this study is to classify the given connection records as either benign (normal) or attack and also to classify the attack types into their corresponding classes. This work also compares the performance of ML and Deep Neural Network (DNN) models for botnet detection. The training phase incorporates the process of manually adding labels to the records which will be utilised by the ML algorithms and DNN's. These records are differentiated based upon the corresponding botnet families. This study displays the usage of Scikit-learn (Machine Learning algorithms) whereas Deep Learning models are implemented using TensorFlow coupled with Keras. To understand the characteristics of the datasets, the t-distributed stochastic neighbor embedding (t-SNE) visualization technique is adopted. To further understand the working of this study, the connection records have a two-fold distribution. These include classification of the records into different attack categories and classification of records into different botnet families (Mirai, BASHLITE). From the experiments conducted, it can be observed that Decision Tree (DT) is an optimal algorithm. However, DNN performed better than the DT classifier. As a result, both the DT and DNN models can be used for botnet detection which can be deployed in real-time IoT networks. Furthermore, SVM models also displayed a good performance, however, the training and testing time associated with this model led to the downgrade of this ML classifier.

As presented in [51], the main objective of the study was to develop a IoT malware detection system that is termed as EDIMA (Early Detection of IoT Malware Network). EDIMA employs machine learning algorithms for traffic classification, a packet vector database and a policy module. The performance of this system is verified through a testbed structure and the associated results are procured. This proposed solution consists of machine learning algorithms running at the user access gateway which detect malware activity based on their scanning traffic patterns.

The entire EDIMA architecture is a set of very specific ML algorithms that are directed at capturing and detecting the malware associated with the IoT network. In this study, the IoT malware is categorized into three classes based on the type of vulnerability that is targeted: TELNET, HTTP POST and HTTP GET. The ML classification is performed on IoT access gateway-level traffic rather than device-level traffic. There are two classes of gateway-level traffic: benign and malicious. Furthermore, a testbed consisting of a PC, smartphone and IoT devices connected to an access gateway was used to evaluate the classification performance of EDIMA. During the experimentation phase, a total of 60 traffic sessions were collected for both benign and malicious classes. The records were split into training and test data using a 70:30 split. The ML models utilised in the research are Gaussian Naïve Bayes, k-NN and Random Forest algorithms. The final results showcase that k-NN was the most efficient algorithm amongst the group. It had an accuracy of 94.44%, precision of 0.92 and a F1-score of 0.96. Whereas, Random Forest had an accuracy of 88.8% and Gaussian Naïve Bayes displayed an accuracy of only 77.8%.

V. FALSE DATA INJECTION

False Data Injection attack is a malicious cyber-threat that that target critical infrastructures controlled by Cyber-Physical systems. FDIA was introduced originally in the field of smart grid. It specifically means that an attacker compromises sensor readings in such a complex manner that undetected errors are introduced into the calculations of state variables and values [43]. The strategies pertaining to FDI leverage wireless IoT device communication network vulnerabilities to manipulate and jeopardise the sensory data points. By these manipulations, hackers can mislead the power distribution network and multiple control centres. One such FDI attack that goes by the name of ‘Stuxnet’ was a computer worm that targeted PLC controllers which automated the industrial processes. This worm attacked the Windows OS coupled with Siemens Step 7 real-time data transmission software. This further caused the uranium gas centrifuges to spin out of control and cause threatening damage to the entire power grid. The following studies explores the domain of False Data Injection [44-45].

In this study [31], the use of autoencoders has been established as a machine-learning tool for the detection of FDI attacks. Autoencoders have been successful in recent times and they have been emerging in the machine learning domain. These are neural networks that is capable of learning latent feature representation of the input, generally in a lower dimension. Additionally, another attractive feature of this technique is that they do not require any ‘labelled’ data for the training phase of the model. Autoencoders are able to pick up complex correlation structures in an unsupervised manner. Another methodology has been introduced as denoising autoencoders to clean the compromised data points, by recovering the expected correlation structure. To perform the experiment, the data set was experimentally obtained by utilising a hydraulic test rig. This proposed model used 60% of data for training, and the remaining 40% was used for testing and validating. When comparing with SVM, the SVM model needs to be trained with both types of data (original and false) along with labels. The major advantage of AE-based approaches over SVM’s is that they can detect multiple attacks since they are not trained to classify any ‘specific’ attack. These models detect any attack that causes a significant distortion in the data correlation model. In conclusion, the rate of detection for autoencoders is much higher than the various SVM models. Furthermore, rate of false alarm for the SVM’s is comparatively much greater than the Autoencoder model. On the other hand, the denoising autoencoder demonstrated high ability to recover data to their original state.

HealthGuard is a machine-learning based security framework for Smart Healthcare Systems (SHS) [27]. The usage of implantable medical devices and wearables, SHS can continuously monitor different vital metrics of a patient and automatically detect and prevent critical medical conditions. In this work, a novel ML based security framework is implemented to detect malicious activities. Furthermore, the performance of this technique was evaluated against three different malicious threats. HealthGuard is capable of capturing correlation between different body functions of the patient and further observing the vital signs of different smart medical devices to detect malicious activity. Additionally, HealthGuard was evaluated against three different threats. This research has proved that HealthGuard is capable of detecting multiple threats with high accuracy and precision. There are multiple differences between existing solutions and HealthGuard technology. While existing solution focuses on the sensors located on the wearable devices, this novel methodology detects by considering interconnected body functions. Furthermore, HealthGuard does not include any overhead cost of processing complexity on the sensor node. To evaluate against the malicious attacks, three different threats (tampered device, DOS, False Data Injection) were introduced into the system. Multiple visualizations were developed as a part of attack sequence. It is observed that ANN algorithm performs with highest accuracy (91%) and F1 score of 89%. For DT algorithm, the accuracy decreases to 90% while F1 score increases to 90%. For KNN and RF, both accuracy and F1 score vary from 86% - 87%. While evaluating simultaneous attack scenarios, the accuracy decreases with the number of attacks in the system. The ML technique can effectively work with an accuracy of 91% with three simultaneous attacks. ANN (Deep Learning Technique) has proved to be the best method for detection of these security threats.

VI. CONCLUSIONS

This research paper has a laser-like emphasis and dives into specific security attacks, especially the Distributed Denial of Service (DDoS) and False Data Injection, and how a field is attempting to tackle the underlying concerns. This methodical survey employs and analyses a set of high-quality research papers used in the development of the study. This work builds on an assumption; mitigation of majority attacks will enhance the IoT landscape tremendously. Machine Learning is not only extensively used but is also becoming ubiquitous in protecting IoT-enabled devices and infrastructures. However, the existing ML solutions, especially DDoS attacks, are widespread and abstract. Although after utilisation of various methods, the ML solutions lack precise standards and protocols. Taking inspiration from the field of networking, which is associated with non-IoT infrastructures, undertaking set protocols reduce the vigour of cyber threats as these benchmarks are complete in themselves. Developing proper standards and protocols will result in uncomplicated implementation and alleviation of threats.

This study has given a variety of takeaways and insights after going through the selected research papers. False Data Injection Attacks (FDIA) are majorly found in smart grids, healthcare sensors and associated settings. This happens because in this type of attack, the sensors are hijacked and exploited which are further used as entry points to inject fake data values. The purpose of this attack is to jeopardise the entire operation of the IoT infrastructure either temporarily or permanently. Furthermore, DDoS/DoS Attacks are focused on webservers and databases in order to damage their networking capabilities. The purpose of these attacks is to create a nuisance by temporarily affecting the access to the targeted webserver causing loss of business. Another interesting insight that was generated through this review is that the usage of Deep Learning (DL) has increased rapidly as compared to the traditional Machine Learning (ML). Also, these Deep Learning techniques are highly efficient as compared to Machine Learning. Deep Learning is a broader part of the ML methods particularly based upon the Artificial Neural Networks (ANN). DL has huge data requirements but functions without any human intervention. These techniques will undoubtedly play a role in generating valuable inferences from the massive volumes of data and hence will assist in developing secure IoT infrastructures. The future of this domain lies in standardization of protocols and techniques such that the existing security threats can be eradicated. Testing with real data points and simulated environment will be extremely rewarding. Furthermore, a deep understanding of the existing solutions and practical simulations are required immediately.

REFERENCES

[1] S. Gupta, S. Vyas and K. P. Sharma, “A Survey on Security for IoT via Machine Learning”, in International Conference on Computer Science, Engineering and Applications (ICCSEA). IEEE, 2020, pp. 1-5.
[2] F. Hussain, R. Hussain, S. A. Hussain and E. Hossain, “Machine Learning in IoT Security: Current Solutions and Future Challenges”, in IEEE Communications Surveys & Tutorials. IEEE, 2020, 22(3), pp. 1686 – 1721.
[3] R. Ahmad and I. Alsmadi, “Machine learning approaches to IoT Security: A systematic literature review”, Internet of Things 14 (2021): 100365
[4] Tahsien, Syeda Manjia, Hadis Karimipour, and Petros Spachos. "Machine learning based solutions for security of Internet of Things (IoT): A survey." Journal of Network and Computer Applications 161 (2020): 102630.
[5] Al-Garadi, M.A., Mohamed, A., Al-Ali, A.K., Du, X., Ali, I. and Guizani, M., “A survey of machine and deep learning methods for internet of things (IoT) security”, in IEEE Communications Surveys & Tutorials, 2020, 22(3), pp.1646-1685.
[6] Cui, L., Yang, S., Chen, F., Ming, Z., Lu, N. and Qin, J., “A survey on application of machine learning for Internet of Things”, in 2018 International Journal of Machine Learning and Cybernetics, 2018, 9(8), pp.1399-1417.
[7] Uprety, A. and Rawat, D.B., “Reinforcement learning for iot security: A comprehensive survey”, in IEEE Internet of Things Journal”, 2020, 8(11), pp.8693-8706.
[8] Haji, S.H. and Ameen, S.Y., “Attack and anomaly detection in iot networks using machine learning techniques: A review”, in Asian journal of research in computer science, 2020, 9(2), pp.30-46.
[9] Alhajri, R., Zagrouba, R. and Al-Haidari, F., “Survey for anomaly detection of IoT botnets using machine learning auto-encoders”, in Int. J. Appl. Eng. Res, 2019, 14(10), pp.2417-2421.
[10] da Costa, K.A., Papa, J.P., Lisboa, C.O., Munoz, R. and de Albuquerque, V.H.C., “Internet of Things: A survey on machine learning-based intrusion detection approaches” in Computer Networks, 2019, 151, pp.147-157.
[11] Hassija, V., Chamola, V., Saxena, V., Jain, D., Goyal, P. and Sikdar, B., “A survey on IoT security: application areas, security threats, and solution architectures” in IEEE Access, 2019, 7, pp.82721-82743.
[12] Al-Hadhrami, Y. and Hussain, F.K., “DDoS attacks in IoT networks: a comprehensive systematic literature review”, in World Wide Web, 2021, 24(3), pp.971-1001.
[13] Munshi, A., Alqarni, N.A. and Almalki, N.A., “Ddos attack on IoT devices” In 2020 3rd International Conference on Computer Applications & Information Security (ICCAIS). IEEE, 2020, (pp. 1-5).
[14] Reda, H.T., Anwar, A. and Mahmood, A., “Comprehensive survey and taxonomies of false data injection attacks in smart grids: attack models, targets, and impacts”, in Renewable and Sustainable Energy Reviews, 2020, 163, p.112423.
[15] Aoufi, S., Derhab, A. and Guerroumi, M., “Survey of false data injection in smart power grid: Attacks, countermeasures and challenges”, in Journal of Information Security and Applications, Elsevier, 2020, 54, p.102518.

[16] Aboelwafa, M.M., Seddik, K.G., Eldefrawy, M.H., Gadallah, Y. and Gidlund, M., "A machine-learning-based technique for false data injection attacks detection in industrial IoT", in IEEE Internet of Things Journal, 2020, 7(9), pp.8462-8471.
[17] Niu, X., Li, J., Sun, J. and Tomsovic, K., "Dynamic detection of false data injection attack in smart grid using deep learning", in IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), IEEE., 2019, pp. 1-6.
[18] Trevizan, R.D., Ruben, C., Nagaraj, K., Ibukun, L.L., Starke, A.C., Bretas, A.S., McNair, J. and Zare, A., "Data-driven physics-based solution for false data injection diagnosis in smart grids", in IEEE Power & Energy Society General Meeting (PESGM), IEEE, 2019, pp. 1-5.
[19] Bostami, B., Ahmed, M. and Choudhury, S., "False data injection attacks in internet of things. In Performability in internet of things", Springer, 2019, pp. 47-58.
[20] Wang, S., Bi, S. and Zhang, Y.J.A., "Locational detection of the false data injection attack in a smart grid: A multilabel classification approach". IEEE Internet of Things Journal, IEEE, 2020, 7(9), pp.8218-8227.
[21] do Nascimento Alves, H., Bretas, N.G., Bretas, A.S. and Matthews, B.H., "Smart grids false data injection identification: a deep learning approach", in PES Innovative Smart Grid Technologies Europe (ISGT-Europe), IEEE, 2019, pp. 1-5.
[22] Zhang, Y., Wang, J. and Chen, B., "Detecting false data injection attacks in smart grids: A semi-supervised deep learning approach" in IEEE Transactions on Smart Grid, 2020, 12(1), pp.623-634.
[23] Nawaz, R., Akhtar, R., Shahid, M.A., Qureshi, I.M. and Mahmood, M.H., "Machine learning based false data injection in smart grid", in International Journal of Electrical Power & Energy Systems, 2020, 130, p.106819.
[24] Ashrafuzzaman, M., Chakhchoukh, Y., Jillepalli, A.A., Tosic, P.T., de Leon, D.C., Sheldon, F.T. and Johnson, B.K., "Detecting stealthy false data injection attacks in power grids using deep learning, in 14th International Wireless Communications & Mobile Computing Conference (IWCMC), IEEE, 2018, pp. 219-225.
[25] Tiwari, R., Sharma, N., Kaushik, I., Tiwari, A. and Bhushan, B., 2019, October. Evolution of IoT & data analytics using deep learning, in International Conference on Computing, Communication, and Intelligent systems (ICCCIS), IEEE, 2019, pp. 418-423.
[26] Zainab, A., S Refaat, S. and Bouhali, O., "Ensemble-based spam detection in smart home IoT devices time series data using machine learning techniques". Information, 2020, 11(7), p.344.
[27] Newaz, A.I., Sikder, A.K., Rahman, M.A. and Uluagac, A.S., "Healthguard: A machine learning-based security framework for smart healthcare systems", in Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS) IEEE, 2019, pp. 389-396.
[28] Li, M., Sun, Y., Lu, H., Maharjan, S. and Tian, Z., "Deep reinforcement learning for partially observable data poisoning attack in crowdsensing systems", in IEEE Internet of Things Journal, 7(7), 2019, pp.6266-6278.
[29] Mendonça, R.V., Silva, J.C., Rosa, R.L., Saadi, M., Rodriguez, D.Z. and Farouk, A., "A lightweight intelligent intrusion detection system for industrial internet of things using deep learning algorithms", in Expert Systems, 39(5), 2022, p.e12917.
[30] Moudoud, H., Khoukhi, L. and Cherkaoui, S., "Prediction and detection of fdia and ddos attacks in 5g enabled iot", in IEEE Network, 35(2), 2022, pp.194-201.
[31] Cao, J., Wang, D., Qu, Z., Cui, M., Xu, P., Xue, K. and Hu, K., "A novel false data injection attack detection model of the cyber-physical power system", in IEEE Access, 8, 2020, pp.95109-95125.
[32] Xu, G., Li, H., Ren, H., Yang, K. and Deng, R.H., "Data security issues in deep learning: attacks, countermeasures, and opportunities", in IEEE Communications Magazine, 57(11), 2019, pp.116-122.
[33] Vishwakarma, R. and Jain, A.K., "A honeypot with machine learning based detection framework for defending IoT based botnet DDoS attacks", in 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, IEEE, pp. 1019-1024
[34] Hussain, F., Abbas, S.G., Husnain, M., Fayyaz, U.U., Shahzad, F. and Shah, G.A., "IoT DoS and DDoS attack detection using ResNet", in IEEE 23rd International Multitopic Conference (INMIC), 2020, IEEE, pp. 1-6.
[35] Aljuhani, A., 'Machine learning approaches for combating distributed denial of service attacks in modern networking environments", in IEEE Access, 9, 2021, pp.42236-42264.
[36] Islam, U., Muhammad, A., Mansoor, R., Hossain, M.S., Ahmad, I., Eldin, E.T., Khan, J.A., Rehman, A.U. and Shafiq, M., "Detection of Distributed Denial of Service (DDoS) Attacks in IOT Based Monitoring System of Banking Sector Using Machine Learning Models", in Sustainability, 14(14), 2020, p.8374.
[37] Chen, Y.W., Sheu, J.P., Kuo, Y.C. and Van Cuong, N., "Design and implementation of IoT DDoS attacks detection system based on machine learning", in 2020 European Conference on Networks and Communications (EuCNC), 2020, IEEE, pp. 122-127.
[38] Saini, P.S., Behal, S. and Bhatia, S., 'Detection of DDoS attacks using machine learning algorithms", in 7th International Conference on Computing for Sustainable Global Development (INDIACom), 2020, IEEE, pp. 16-21.
[39] Cil, A.E., Yildiz, K. and Buldu, A., "Detection of DDoS attacks with feed forward based deep neural network model", in Expert Systems with Applications, 2021, 169, p.114520.
[40] Gaur, V. and Kumar, R., "Analysis of machine learning classifiers for early detection of DDoS attacks on IoT devices". in Arabian Journal for Science and Engineering, 2022, 47(2), pp.1353-1374.
[41] Roopak, M., Tian, G.Y. and Chambers, J., "An intrusion detection system against ddos attacks in iot networks", in 2020 10th annual computing and communication workshop and conference (CCWC), IEEE, 2020, pp. 0562-0567.
[42] Kamble, P. and Gawade, A., "Digitalization of healthcare with IoT and cryptographic encryption against DOS attacks", in 2019 International Conference on contemporary Computing and Informatics (IC3I), IEEE, 2019, pp. 69-73.
[43] Doriguzzi-Corin, R., Millar, S., Scott-Hayward, S., Martinez-del-Rincon, J. and Siracusa, D., "LUCID: A practical, lightweight deep learning solution for DDoS attack detection", In IEEE Transactions on Network and Service Management, 2020, 17(2), pp.876-889.
[44] Ferrag, M.A., Shu, L., Djallel, H. and Choo, K.K.R., "Deep learning-based intrusion detection for distributed denial of service attack in Agriculture 4.0. Electronics", MDPI, 2021, 10(11), p.1257.
[45] Syed, N.F., Baig, Z., Ibrahim, A. and Valli, C., "Denial of service attack detection through machine learning for the IoT", Journal of Information and Telecommunication, 2020, 4(4), pp.482-503.
[46] Shurman, M.M., Khrais, R.M. and Yateem, A.A., "DoS and DDoS attack detection using deep learning and IDS", in Int. Arab J. Inf. Technol., 2020, 17(4A), pp.655-661.

[47] Sriram, S., Vinayakumar, R., Alazab, M. and Soman, K.P., "Network flow based IoT botnet attack detection using deep learning", in IEEE INFOCOM 2020-IEEE conference on computer communications workshops (INFOCOM WKSHPS), IEEE, 2020, pp. 189-194.

[48] Alhajri, R., Zagrouba, R. and Al-Haidari, F., "Survey for anomaly detection of IoT botnets using machine learning auto-encoders" in Int. J. Appl. Eng. Res, 14(10), 2019, pp.2417-2421.

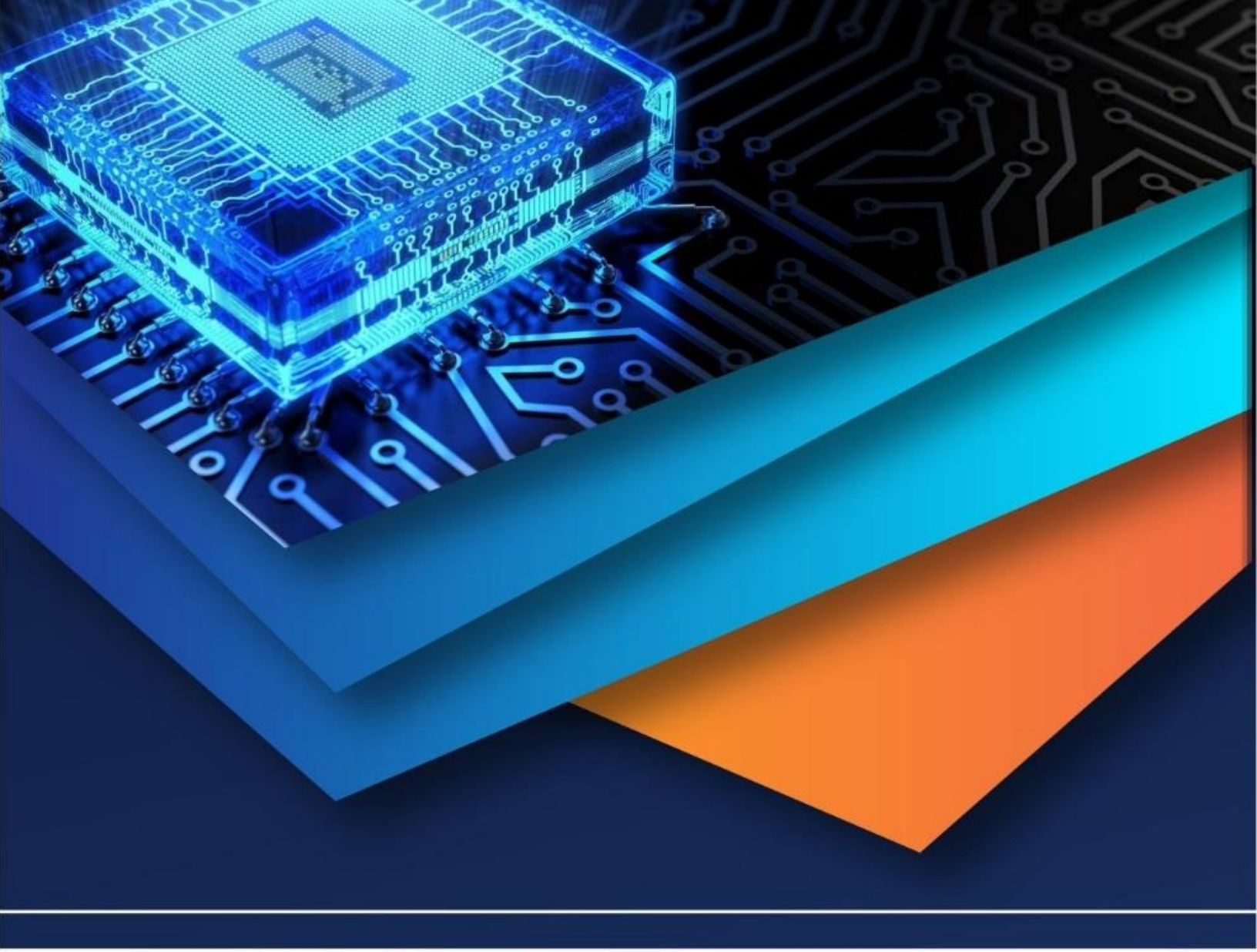
[49] Bhatia, R., Benno, S., Esteban, J., Lakshman, T.V. and Grogan, J., "Unsupervised machine learning for network-centric anomaly detection in IoT", in Proceedings of the 3rd acm conext workshop on big data, machine learning and artificial intelligence for data communication networks, 2019, pp. 42-48.

[50] Gurulakshmi, K. and Nesarani, A., "Analysis of IoT bots against DDOS attack using machine learning algorithm" in 2018 2nd International conference on trends in electronics and informatics (ICOEI), IEEE, 2018, pp. 1052-1057.

[51] Kumar, A. and Lim, T.J., "EDIMA: Early detection of IoT malware network activity using machine learning techniques". in 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), IEEE, 2019, pp. 289-294.

[52] Silveira, F.A.F., Lima-Filho, F., Silva, F.S.D., Junior, A.D.M.B. and Silveira, L.F., "Smart detection-IoT: A DDoS sensor system for Internet of Things", In 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), IEEE, 2020, pp. 343-348.

[53] Ge, M., Fu, X., Syed, N., Baig, Z., Teo, G. and Robles-Kelly, A., "Deep learning-based intrusion detection for IoT networks" in 2019 IEEE 24th pacific rim international symposium on dependable computing (PRDC), IEEE, 2019, pp. 256-25609.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY