IoT DDoS attacks detection using machine learning techniques: A Review

2021 International Conference on Data Analytics for Business and Industry (ICDABI) | 978-1-6654-1656-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICDABI53623.2021.9655789

Amreen Ashraf
Computer Science
University of Bahrain
Sakheer, Bahrain
amreenashraf0@gmail.com

Wael Mohammed Elmedany
Computer Science
University of Bahrain
Sakheer, Bahrain
welmedany@uob.edu.bh

***Abstract*—Internet of things (IoT) is the paradigm that is revolutionizing our daily lives. It has become an important part of our lives due to its ability to transform lives. More and more organizations are using IoT devices as they open new opportunities for healthcare, wearable devices, home appliances and improve sharing and communication of information using the internet. With all these opportunities the challenges related to IoT security are rising. The limited resources and open deployment environment makes it vulnerable to several malicious attacks such as Distributed Denial of service (DDoS). Traditional detection approaches are inadequate for current security requirements. By applying the machine learning techniques these approaches can be improved which in turn will result in improved security. Therefore, this paper is a review of current advancements made in the application of machine learning techniques for the detection of DDoS attacks in IoT. There are many survey papers on intrusion detection in IoT but there is very little which particularly focuses on DDoS attacks. As more and more organizations are using IoT devices, but due to lack of knowledge they sometimes find it hard to understand how to ensure security, this paper will serve as a guide, as its main contribution is to provide an overview of the current state of the art of applications of machine learning techniques in the detection of Distributed Denial of Service attacks in IoT. This paper will also add to the knowledge of researchers who are interested in IoT security.**

I. Introduction

The emergence of the internet of things has improved many aspects of human lives, as it helps to connect billion of things and exchange information easily. Internet of things (IoT) is a technology that is no longer limited to computers only, with continuous advancements it has grown and became a network of devices of all sizes and types like home appliances, vehicles, smartphones, industrial systems, medical equipments, toys, cameras, animals, buildings, and people all connected, sharing information, and communicating using different developed protocols with the aim to achieve smart positioning, tracing, reorganizations, online upgrading, and personal real-time online monitoring and process administration and control [1].

In recent years, the Internet of Things (IoT) seems to have become a trendy topic, as there has been immense increase in the number of devices connected to internet and it is expected to grow further in the coming years. According to reports the number of IoT devices will grow to 5 billion by the end of 2025 which is more than double of number of devices in 2019 (1.3 billion) [2].

However with all the advantages security still is as vital issue as the IoT technology itself. The future application of IoT and the currents are still adopting the host to host communication which has several chronic drawbacks as it was designed several years ago the other reason is its conservative and straightforward design which exposes it to a number of threats. According to statics, in 2016 a security breach was encountered by more than a third of the companies, and IoT systems are in a far worse predicament, as IoT devices have restricted resources.DDoS attacks specifically comprise a determined threat to the Internet of Things (IoT). What's more, DDoS attacks in the IoT have expanded fundamentally [3].

Indeed, developing advanced procedures capable of ensuring sufficient security levels to identify and neutralize cyber-threats whenever they arise, is required to achieve IoT success [2].

A. Background and motivation

In 2016 a French Web host and Brian Krebs a security consultant's website were targeted with the attack traffic of 1.1 Tbps and 620 Gbps respectively. The attack was called Marai which is a Japanese word that means "the future". The attack was conducted using 60,000 infected devices. Marai source code was released to the public, which led to a massive increase in attacks. Dyn was a famous attack with a volume of 1.2 Tbps, hundreds of websites including Reddit, GitHub, Twitter, and Netflix were brought down as a result of this attack. The Internet of Things (IoT) is highly susceptible to DDoS attacks, and it is also being utilized to attack other targets.

The number of DDoS attacks is rising all the time. In 2013 and 2015, it was discovered to be running at a 100 Gbps rate. In 2016 and 2017, the operating volume was increased to 800

978-1-6654-1656-6/21/$31.00 ©2021 IEEE

Gbps and 1.35 Tbps, respectively. Due to poor implementation in IoT devices, attackers are able to hijack these devices easily and then use these devices to attack the intended server. With the growth in the use of IoT devices, the volume of DDoS attacks is also increasing and so is the security spending. According to a survey conducted in the year 2018 at least 20 percent of the organizations experienced DDoS attacks once. Security investment in IoT was 1.5 billion dollars in 2018, and by 2021 it is predicted to reach 3.1 billion dollars [4].

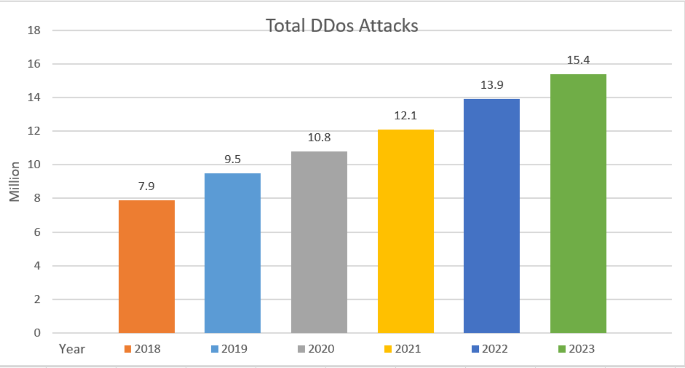

Fig. 1. DDoS Attacks [5]

Fig 1 shows increasing DDoS attacks.These figures are staggering, it is important to pay attention to this issue. Detection accuracy and early detection are some of the most important factors to deal with DDoS attacks effectively. Every detection technique should be able to recognize irregularities with low false negative and false positive ratios and outline normal traffic accurately and intelligently and should be cost-effective in the consumption of resources and per-packet computations [6].

These attacks can be detected using conventional detection techniques, which are based on anomaly or signature-based approaches. There are many issues associated with the conventional systems, for example, anomaly-based systems raise alert that is not true. The detection accuracy for these systems is also quite low if the baseline data for legitimate users is not adjusted sensibly [7].

The contribution of this paper can be summarized as:

As there is lack of survey papers which particularly focuses on DDoS attacks this paper will close this gap by providing review of current advancements made in the application of machine learning techniques for the detection of DDoS attacks in IoT.

As an ever increasing number of organizations are utilizing IoT devices, however because of absence of information they sometime have difficulty in understanding how to guarantee security, this paper will fill in as an aide.

This study will also help researchers who are interested in IoT security to expand their expertise.

The rest of the paper is organized as follows Section II describes the related work, Section III describes the DDoS attacks and its types, Section IV describes the Machine learning technique, types of machine learning techniques, and Application of Machine Learning techniques, Section V describes Discussion and Section VI gives Conclusion.

II. Related Work

In this paper, the researcher Khadijeh et al [8] provided a comparison of three different approaches of machine learning which can be applied to effectively detect and mitigate the DDoS attacks and provided an overview of the advantages and disadvantages of these approaches. The researcher also provided a performance evaluation of these approaches. The first approach used for comparison is feature extraction and it consists of four phases: Traffic capture, Packet grouping, the third phase is Feature extraction and fourth is binary classification. This approach is designed for the detection of malicious nonlegal traffic, it can be deployed on ISP-controlled switches, network middle-boxes, and gateway routers. Five different ML classifiers like the random forest, Decision tree (DT), KNN, Neural Network (NN), and Support vector machine with the linear kernel (LSVM) were tested. The researcher concludes that simple classification algorithms and low-dimensional features effectively differentiate between traffic from legal IoT devices and DDoS attacks traffic. The second approach reviewed by the researcher is the machine learning approach which uses SDN for the detection of DDoS attacks. The third approach reviewed by the researcher is based on a back propagation artificial neural network integrated with Apache Spark.

The researcher Fatima et al [9] highlights that the resource-constrained nodes in IoT devices make them vulnerable to a number of cyber attacks. In this paper, the researcher surveyed how different machine learning and deep learning techniques can be used to handle the security issues in IoT devices. The researcher provides a review of the current security requirements, current solutions, identify gaps and suggests machine learning solution already present based on the security requirements.

In this paper, the researcher Puja et al [10] focuses on the machine learning techniques for cybersecurity in IoT in healthcare. Wearable sensors in IoT healthcare collect a large amount of data, and the cloud computing environment is excellent for processing and analyzing this data. Many security vulnerabilities affect cloud-based IoT in healthcare. As a result, this paper focuses on providing an overview of the many cloud-based IoT healthcare systems that use various machine learning methodologies utilized in cybersecurity and for data collection in IoT. Following the survey, the researcher proposes a model that integrates IoT healthcare Cloud technologies and machine learning.

III. Distributed denial of Service (DDoS) Attacks

Distributed denial of service attack can be described as “a malicious attempt to make an online service unavailable to users, usually by temporarily interrupting or suspending the services of its hosting server”.

This attack is often carried out through botnets, which are groups of infected devices that are disseminated internation-

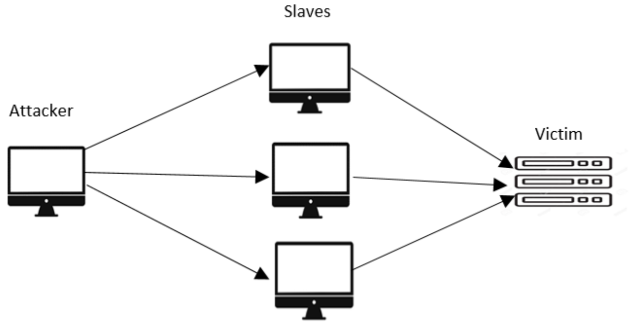

Fig. 2. How DDoS attacks are conducted [11]

ally. It differs from DoS attacks, which are carried out by a single internet-connected device. The fundamental purpose of carrying out this attack is to deprive the legitimate user of the bandwidth resources and network. First, the non-legacy of IoT devices is targeted by the attacker. These devices have constraints like battery capacity, low computation, and power poor built-in security, CCTV cameras, webcams, and baby monitoring gadgets are just a few examples. Due to less security, the attacker easily injects the malware and seizes the control of the device with the help of the tools like the Lizard stressor or Marai Code. A bot is a device in which the attacker injects malware and Botnets are formed when these bots are brought collectively. DDoS attacks on servers are carried out using these botnets, unlike the traditional ones that are used to spam the user or steal the credentials.

The hacked machines are subsequently utilized to launch volumetric DDoS attacks which include the following: ICMP, TCP, and UDP flooding. Open-sourced tools like DDoSim LOIC, GoldenEye, and Pylori's are used to carry the attacks such as HTTP GET/POST.

Some common types of attacks are Zero-day DDoS attacks, NTP Amplification, HTTP Flood, Slow Loris, UDP flood ICMP (Ping) Flood, and SYN Flood [4].

A. Types of DDoS attacks:

There are different types of DDoS attacks:

1) Application Layer Attacks : It includes POST/GET floods, slow and low attacks, attacks that target Windows Apache or Open BSD vulnerabilities. These attacks are composed of innocent and legitimate seeming requests which target the web server and crash it. The magnitude of these types of attacks is measured in Requests per second (Rps).

2) Volume Based Attacks: The purpose of these attacks is to overwhelm the targeted site's bandwidth. It includes the spoofed packet, ICMP, and UDP floods. Bps (bits per second) is the unit of measurement for the severity of these attacks.

3) Protocol Attacks : Fragmented packet attacks, Syn flood, Smurf attacks, ping of death, and many other attacks are included in this type. The impact of these attacks is measured in packets per second, and they employ communication equipment's and numerous resources such as load-balancing firewalls [13].

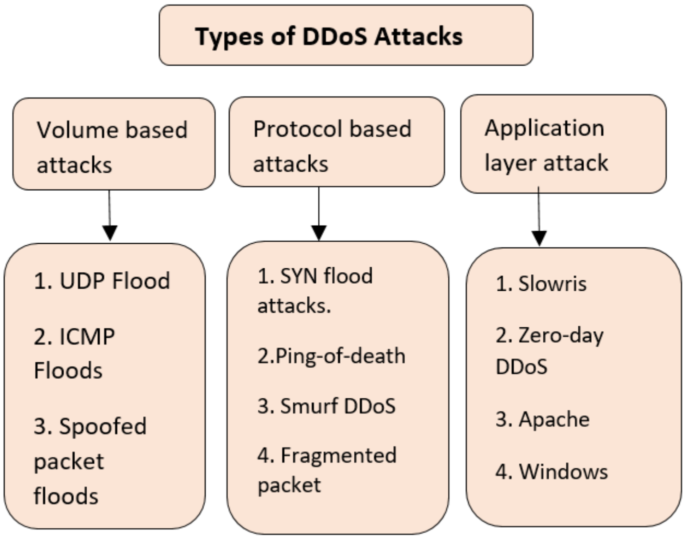

Fig. 3. Types of DDoS attacks [12]

IV. Machine Learning

Machine learning is an artificial intelligence (AI) technology that allows computers to learn and develop without being specifically designed. The field of machine learning deals with the creation and design of computer programs that learn on their own and access data. The learning process starts with data or observations such as instruction or direct experience to make future decisions based on the instances. The main objective is to teach computers to learn on their own, without the need for human intervention [14].Some of the machine learning techniques used in the research papers are described below:

1) J48 Algorithm or C4.5: It is also called the C4.5 descending from the ID3 algorithm. It consists of a root and leaves nodes as it is a tree-based algorithm. Because pruning techniques can help in overfitting the problem. It is built by calculating the entropy value and information gain of all attribute values. It is a widely used algorithm that can handle continuous, uncompleted, and discrete data points and values [15].

2) Naive Bayes : The Bayesian theorem is used to create this classifier and it is used as a decision tree algorithm in a supervised manner. It is very useful when the dimensionality of the input is very high. Although it is built on the simple theory, it has much better-sophisticated classification results. The number of instances do not affect the prediction time in Naive Bayes. It represents learning and using probabilistic knowledge. It provides features learning [15].

3) Artificial neural network : It is employed in the study of vector-valued functions and is built using biological neural networks. The elements used are called the neurons which are interconnected and work in parallel. A subgroup is connected with these elements through links and it is called the layer.

Output input and hidden are the three types of layers in artificial neural networks [15].

4) Support vector machine: It's a supervised machine learning method that can handle regression and classification problems. Primarily it is employed to solve the problems of classification [16].

5) Fuzzy logic : Unlike most of the computers which are based on Boolean logic (0 and 1) or (false or true), it uses the approach ”degrees of truth”. The fuzzy logic operates just like human brain. It aggregates all the data and based on this data facts are generated which are further gathered to generate higher truths. Fuzzy logic can play a crucial role in developing artificial intelligence human-like capabilities [17].

6) Decision tree : It is a technique that is used for solving regression problems and also for classification problems. It consists of classifiers that have a tree structure, leaf node signifies the result, branches represent decision rules, and the internal nodes represent the data set features. Leaf node and decision node are two types of nodes in a decision tree. There are no further branches or nodes. Using the properties of the data set the decisions are made [18].

7) Random Forest : As the name shows, a random forest is comprised of an enormous number of individual decision trees that cooperate as a group. In the random forest, each tree produces a class prediction, the class with the most predictions becomes the prediction of the model [18].

8) K-Nearest Neighbor : It is based on a supervised learning technique and it is one of the most basic machine learning algorithms. All the available data is stored and a new type of data is classified by comparison and similarity. It helps easy classification of data based already available data. It can be used for regression and classification problems [18].

A. Application of machine learning techniques in DDoS Attack Detection:

This section describes the advancements made in the application of machine learning techniques for the detection of DDoS attacks in IoT.

The dynamics of security in the internet of things are the topic of study in this paper. The author highlights that the issue of security in IoT is not addressed adequately while designing the security procedures. The paper focuses on the uncertainty and major security issues of IoT. Focusing on these issues the researcher Sayed et al [20] proposed a fuzzy logic and fog-based secure architecture called FLFSIOT which works in real-time. The framework for FLFSIOT consists of six segments. These segments run at three levels. ZAD (Zero-day attack detection), Security threat alarm, Defense evaluation, Security recovery, Attack response, and Active security are the segments. DDoS and collusion attacks are detected by the ZAD segment, which runs on the cluster head fog nodes. The six segments work in sync to detect the zero-day and denial of service attacks. The fuzzy knowledge base helps to detect the attack then raises the alarm and the attack receives the response. FLFSIOT has been made more secure than cloud-supported IoT by eliminating concerns like latency from fog-supported IoT architecture. The proposed framework utilizes the benefits of fog and fuzzy logic to model a solution that can improve security problems. The effectiveness of the proposed solution is tested for the collusion and denial of service attacks and results show that it gives more accurate results. The result of the false-negative rate is 0.75 percent which is lessening further with time. It has a detection accuracy of 95 percent. The greatest FP rate (85 hubs) is 2 percent.

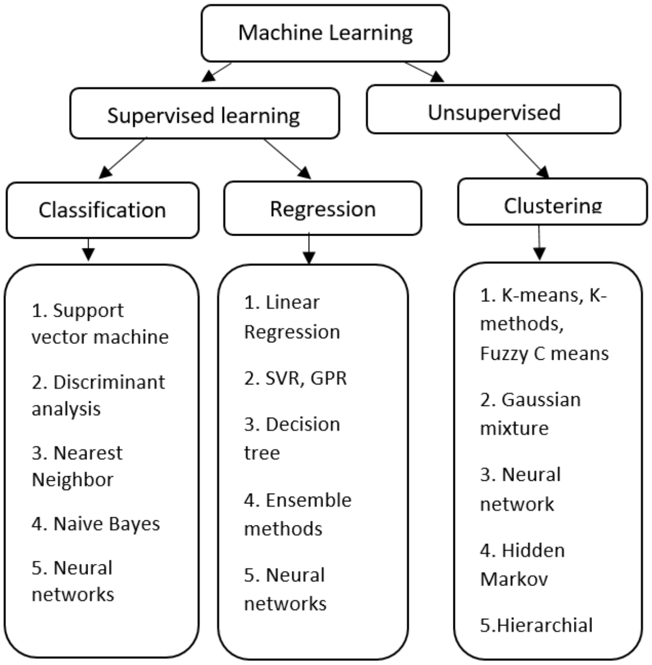

Fig. 4. Types of machine learning techniques [19]

The advancements in machine to machine communication have led to the development of the protocols like MQTT which stands for Message Queuing Telemetry Transport (MQTT). These protocols' accessibility has resulted in a slew of attacks. The researcher Harpriya et al [21] highlight that the issue of security needs to be addressed in MQTT as it is widely used and the most prominent type of attacks in it is DDoS. Focusing on the vulnerabilities in MQTT, a lightweight fuzzy logic-based intrusion detection scheme called Secure-MQTT has been proposed by the researcher for detecting the malicious activities between IoT devices during communications. In the proposed solution fuzzy logic-based system uses the fuzzy rule interpolation technique to identify the node's malicious activities. It avoids the use of a dense rule base with the help of fuzzy rule interpolation as it helps in generating the rules dynamically. The solution is tested through simulation, by making an attacking situation in the organization where the quantity of the noxious hubs is 10-50 percent of the absolute number of hubs conveyed in the network. The results show that the proposed Secure MQTT solution helps in detecting the malicious activities of the nodes. Compared to the traditional solution the proposed solution shows much better results. The

Proposed system has been tested with four different scenarios with 100 nodes it gives the following results False-negative 2 percent, False positive 2 percent, and F Score 0.9090.

In this paper, the researcher Sherazi et al [22] addresses the issue of trustworthiness and privacy in the internet of vehicle(IoV). Due to the involvement of a number of components like vehicles sensors, infrastructures, and humans internet of the vehicle is exposed to a large number of threats like DDoS attacks. To address this issue intrusion detection systems tailored to be added in the internet of vehicles, different machine learning, and artificial intelligence techniques have been investigated which can help in countering the DDoS attacks in IoV environments. The presented approach relies on Q learning and fuzzy logic, and it is simulation tested. Contrasted with the conventional approach, the proposed solution shows much better results and it is helpful in boosting security for heterogeneous networks.

In this paper, researcher Sacranti et al [23] addresses the security issues in software-defined networks (SDN). With the number of benefits, SDN is exposed to a number of attacks. The researcher emphasizes that the traditional approaches are inadequate and proposed an artificial immune system in which fuzzy logic technique has been used for reducing uncertainty, reducing the difficulty of differentiating between malicious and normal traffic, and improving the detection. The proposed system is tested with the public data set having multiple types of DDoS attacks as well as with simulated flooding and port scan attacks. The performance of the proposed system is compared with the Local Outlier factor, k-nearest neighbors, and Naive Bayes. It outperforms all the algorithms in terms of false-positive rate and detection accuracy. The simulated system has a False positive Rate of 0.0014, Accuracy of 0.9996, and F-measure rates of 99.97 percent. On Public data set false positive rate 0.2626, accuracy 0.8903, and F-measure rates 92.28 percent.

In this paper, the researcher khraisat et al [24] emphasizes that it is difficult to ensure the security of IoT device infrastructure owing to the vast number of IoT devices and using the traditional approaches or intrusion detection system. As a result implementation and design of a novel intrusion detection system for the detection of attacks in IoT infrastructure has been presented in this paper. To increase the efficacy of attack detection, the suggested approach is based on data gathered from IoT ecosystems. By combining the one-class support vector machine and c5 classifier the author proposed a hybrid intrusion detection system. It provides the best of both the Anomaly-based Intrusion Detection System (AIDS) and the Signature Intrusion Detection System (SIDS). The aim of creating this system is to increase the detection accuracy of zero-day attacks and well-known intrusions and to decrease the false alarm rate. The system is put to the test using the bot IoT data set, which includes both real IoT traffic and a variety of additional threats. The C5 and one-class SVM are much better in performance compared to the individual technique. When compared to AIDS and SIDS utilized alone, the suggested hybrid IDS has a reduced false positive rate and a greater accuracy rate. The proposed system is tested on a real system and has an accuracy of 99.97 percent, False Positive rate 0, and F measure 0.995.

In this paper, the researcher Hodo et al [25] performed threat analysis of the IoT and proposed a system based on the artificial neural network. Internet packets are used to train a supervised ANN and then its ability to detect and prevent the DDoS attack is tested. To evaluate this, the experimental configuration is done which consists of a five-node sensor IoT network. Out of these five nodes, four act as a client and one as a Server. A network tap is used to capture the traffic, the server perceives and reacts to the information given by sensor hubs relying upon the information which has been received earlier permitting sensor hubs to get acquainted with their behavior. The attack in this experiment is from external intruders. The server node is only targeted as it responsible for logging, analyzing, and responding to sensor nodes. A single host was used for performing the DDoS attack, sending over 10 million packets. The paper is focused on the classification of threat patterns and the classification of normal traffic. The result of the experiment shows that the proposed system accurately detect various kinds of DDoS attack. The proposed system has 99.4 percent overall accuracy.

In this paper, the researcher Om et al [26] highlights the impact of DDoS attacks on businesses. He argues that bandwidth overflow, Internet spoofing can terminate the normal functioning of an organization and can cause huge loss. To address this issue advanced support vector machine is used to analyze the patterns of attacks and protect the users from these attacks. The ASVM, wireshark, and packet instrument are employed to implement the projected system. A new type of data set is generated. The generated testbed consists of 30 switches and three controllers compared to another researcher in SDN security who used one controller. The researcher tested hundreds of scenarios for SYN flooding attacks and UDP flooding attacks. Both malicious and normal traffic was generated. The IoT asymmetric features and IoT traffic from the Open Flow switches are extracted and collected so that the data sets can be created effectively. Cross-validation is used for testing the model. In the proposed ASVM algorithm Linear kernel is also used. ASVM prototype has high detection accuracy of 99 percent.

In this paper, the researcher Shorman et al [27] focuses on DDoS attacks. Based on these a machine learning-based attack detection system is proposed for SDN switches, IoT devices, cloud servers, and IoT gateways. In the IoT system, decision trees are provided for online and offline training for DDoS attacks detection, which sends an alert message to the administrator whenever the attack is detected. For experimenting 8 smart poles were build using different sensors and the sensor data was collected through the wired or wireless network. A heterogeneous gate was implemented for collecting the sensor data and for the authentication of the collected data and the ways the features were extracted based on the types including ICMP flood, UDP flood, and SYN flood for improving attacks detection accuracy. The result of the experiment shows that the

SDN controller can effectively block malicious devices and can effectively detect DDoS attacks.

Focusing on the increased attacks due to the limitation of security in IoT, in this paper the researcher Silveir et al [28] proposed a mechanism to detect and mitigate the DDoS attacks which address the current security challenge in IoT. The proposed system is for IoT controllers which classify the network traffic using machine learning techniques. Testing is done using modern data sets. The proposed system is run by embedding it into the wireless access point, it is based on the signature-based machine learning algorithm. Logistic Regression (LR), Extreme Gradient Boost, and Random Forest (RF) are the three-technique evaluated. Based on samples of network traffic the inferences are made.The open flow switch then forwards the collected data to the detection system. The system was evaluated using recent data sets like CICIDS2017, CIC-DOS tailored to contain DDoS attacks, like TCP flood, HTTP flood, UDP flood, and HTTP slow. Different data sets were used for training and testing. The result of the experiment shows that 20 percent of the sampling data of network traffic has a low false alarm rate, shows a high precision (PR) above 93 percent, and the detection rate above 96 percent.

In this paper, the researcher Cheng et al [29] proposed a DDOS detection system based on machine learning for SDN-enabled IoT networks. The suggested system is a multi-agent-based intrusion detection system that uses the Naive Bayes classification technique. These agents are responsible for the analysis, data collection, and development of inferences. For data classification, Naive Bayes is utilized. The detection framework is deployed in the critical node (such as the controller or the switch) of the SDN network. The proposed system has three phases, the first phase is data capture and pre-processing and the second phase is feature vectors generating and the last is classification. In the first stage data is captured from the network interface and Open flow normalizes the data set for the following step by streaming data from network's raw collected data. Two vector groups (stateful and stateless features) are generated. The generated vectors are divided into testing and training sets. The last phase uses different binary classification algorithms like supports vector machines, random forests, k-nearest neighbor, and random forests, with the testing and training data set for DDoS attack detection. For the experiment, the normal traffic consists of Ping messages, HTTP, MQTT, and HTTPS. The attack traffic consists of TCP retransmission and TCP SYN attacks. The detection of up to 0.78 is achieved.

In this paper, the researcher Aysa et al [30] proposed a solution keeping in mind the security challenges due to node density, power limitations, and processing power in IoT devices. Many data mining and machine learning algorithms like decision trees, LSVM, and neural networks are used in the proposed solution to detect the abnormal features or activities in DDoS attacks. Three main types of WiFi-connected IoT camera devices that are used to collect data are Equipment's (security camera), Philips (camera), and Simple Home (safety camera). Two kinds of IoT botnet attacks have been analyzed for the evaluation of the proposed system the primary is BASHLITE, it is a type of attack that taints Linux framework and the second is mirai. The proposed framework has four stages, first is the data set preparation stage, second is data preprocessing, the third step is to train and learn on both normal and typical data sets, and fourth is testing and evaluations. The approaches utilized are useful in detecting the attack, but when decision trees and random forests are combined, excellent accuracy in detecting attacks is attained.

In this paper, the researcher Soe et al [31] based on machine learning, proposed an attack detection system with sequential architecture for IoT. It focuses on many security weaknesses such as lacking computational assets for solid security systems or having inadequate memory and computational assets for successful safety measures. The proposed system is based on different machine learning algorithms, including decision tree, J48, artificial neural network (ANN), and naive bayes. An efficient feature selection approach is applied to attain high performance in implementing a lightweight detection system. The results indicate an accuracy of 0.99.

In this paper, researcher Doshi et al [32] focuses on the threats faced by insecure consumer IoT devices. One of the major types of attacks that make IoT devices insecure is the botnet attacks such as mirai. Based on IoT-specific network behaviors such as, regular interval between a limited number of endpoints and packets the researcher tends to prove that feature selection can result in higher accuracies when detecting the DDoS attacks with several machine learning algorithms like decision tree using Gini impurity scores (DT), support vector machine with the linear kernel (LSVM), K-nearest neighbor "KDTree" algorithm (KN), random forest using gini and neural network. To experiment a network of consumer IoT devices was put up to capture malicious and benign communications. Raspberry Pi v3 was set up as a WiFi access point to act as a middle box. To collect the non-DDoS traffic three IoT devices interacted for 10 minutes. Neural networks, were developed using the keras python library, whereas the scikit-learn python package was used to develop other machine learning models. The accuracy for the naïve algorithm is 0.93. For four other classifiers, accuracy ranges between 0.91 to 0.99.

Karthik et al [33] concentrate on the primary attacks in IoT in this study. The author focuses on preventing DDoS attacks by exploiting SDN (software-defined network) properties such as multidimensionality and flexibility. In this paper, an intrusion detection system by combining the support vector classification algorithm has been proposed which is based on SDN environment. The average accuracy achieved by the system is 96.23 percent.

In this paper, the researcher Amouri et al [34] proposed an intrusion detection system using machine learning techniques. Considering the vulnerabilities resulting from usage of low resources nodes in wireless sensor networks (WSNs) and Mobile ad-hoc networks (MANETs) and the internet of things, an intrusion detection system has been proposed in this study which consist of two-stages.The first stage has sniffers which are used for the collection of data from the network and mac layer. The collected data can be fed into the forest classifier which generates the correctly classified instances (CCIs). The second phase then starts where the CCIs which have been generated are fed into the supernode. The researchers used DDoS and black hole attacks for testing the system. High power/node velocity situations have detection rates of over 98 percent, whereas low power/node velocity situations have detection rates of approximately 90 percent.

In this paper, the researcher Nimbalkar et al [35] highlights that due to a large amount of network traffic the machine learning might take some time to detect attacks hence proposed an intrusion detection system with feature selection using the Gain Ratio (GR) with top-ranking 50 percent features and Information Gain (IG). The system was validated using KDD Cup 1999 data set and it has been tested using the BoT-IoT dataset. The proposed system has much better results.

V. Discussion

Different Researchers have spent a lot of time looking into how to secure IoT devices. These researchers covered a wide range of topics related to IoT device security and tested multiple machine learning techniques to address the issue most of these have proven to be very effective. Researcher Sayed et al [20], Harpriya et al [21], Sherazi et al [22], Sacranti et al [23] focused on the application of fuzzy logic combined with different approaches. When applied individually the detection accuracy of 95 percent is achieved whereas when this technique is combined with another approach it gives an accuracy of 99.96 percent. Most of the research is done on simulated systems that have proven to be very successful but if a different environment is chosen, the performance may be affected.

Support vector machine in paper [24] combined with C5 classifier has an accuracy of 99.97 percent which is tested on the real system and when support vector machine used alone it has detection accuracy of 99 percent in the study [26]. Whereas artificial neural network in the study [25] has detection accuracy of 99.6 percent.

In the paper [28] Extreme Gradient Boost (XGB), Logistic Regression (LR), Random Forest (RF), supports vector machines, and k-nearest neighbor are tested they achieve a high detection rate of over 96 percent and high precision (PR) of over 93 percent. Similarly, multiple approaches have been tested in [29] they have a detection rate of 0.78. Artificial neural networks (ANN), Naïve Bayes, J48, and decision tree are combined and achieve a detection accuracy of 99 percent. Support vector machine, K-nearest neighbors algorithm (KN), Decision tree using Gini impurity scores (DT), Random Forest were tested using the Scikit-learn Python library and Keras library. The author provided an evaluation of different approach in terms of the accuracy the result showed for naïve algorithm accuracy is 0.93 and the rest of the classifiers have detection accuracies from 0.91 to 0.99. [32]

In all the approaches that have been used by different researchers, the support vector machine proves to have higher detection accuracy whether applied individually or combined with another approach. In most of experiments it's evident that when more than one approach is combined it increases the detection accuracy as proved by Aysa et al [30] where the researcher tested the approaches like Decision tree, Neural networks, Machine learning, and data mining algorithms such as LSVM .The result of the experiment has shown that when two approaches are combined they have better results.

There is a lack of research that tests the approaches under a heterogeneous environment, most of the studies are limited to testing the approaches under one condition. It's very important to test the proposed approaches under a heterogeneous environment to understand how the detection rate varies and what factors impact the accuracy.

There is also limited research which tests each of the machine learning technique under similar platform. For example in this paper, [30] the researcher provided a very good comparison of the techniques like Decision tree, Neural Network, machine learning, and data mining algorithms such as LSVM, the only drawback of this study is there are other approaches like support vector machine and naive Bayes techniques which are not covered. As the use machine of learning techniques for security is increasing day by day in the field of computer science, using a common platform to assess the performance of different techniques would be a great achievement as it will help researchers to choose the best technique.

VI. Conclusion

Botnet attacks on the Internet of Things (IoT) have become much more common in recent months, because of the expansion of IoT devices that can be promptly hacked. Botnets are a regular risk that may dispatch a progression of Distributed Denial of Service (DDoS) attacks by abusing the absence of fundamental security systems in IoT devices. More danger is faced due to the inaccuracy of detection techniques. The application of machine learning techniques can help in the detection of these attacks as it is evident from the researches done previously where a detection rate of up to 99.97 percent is achieved. For future, machine learning strategies should be evaluated on a single platform in the future to have a better knowledge of how each methodology works on the same platform. More study should be done into combining two or more techniques since it has been proven in certain studies that this can enhance the accuracy outcome as well as the detection rate.

References

[1] K. K. Patel, S. M. Patel *et al.*, "Internet of things-iot: definition, characteristics, architecture, enabling technologies, application & future challenges," *International journal of engineering science and computing*, vol. 6, no. 5, 2016.

[2] J. Galeano-Brajones, J. Carmona-Murillo, J. F. Valenzuela-Valdés, and F. Luna-Valero, "Detection and mitigation of dos and ddos attacks in iot-based stateful sdn: An experimental approach," *Sensors*, vol. 20, no. 3, p. 816, 2020.

[3] G. Liu, W. Quan, N. Cheng, H. Zhang, and S. Yu, "Efficient ddos attacks mitigation for stateful forwarding in internet of things," *Journal of Network and Computer Applications*, vol. 130, pp. 1–13, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804519300062

[4] M. M. Salim, S. Rathore, and J. H. Park, “Distributed denial of service attacks and its defenses in iot: a survey,” *The Journal of Supercomputing*, pp. 1–44, 2019.

[5] Five most famous ddos attacks and then some. [Online]. Available: https://www.a10networks.com/blog/5-most-famous-ddos-attacks/

[6] P. Kaur, M. Kumar, and A. Bhandari, “A review of detection approaches for distributed denial of service attacks,” *Systems Science & Control Engineering*, vol. 5, no. 1, pp. 301–320, 2017.

[7] S. Apale, R. Kamble, M. Ghodekar, H. Nemade, and R. Waghmode, “Defense mechanism for ddos attack through machine learning,” *International Journal of Research in Engineering and Technology*, vol. 3, no. 10, pp. 291–294, 2014.

[8] K. Wehbi, L. Hong, T. Al-salah, and A. A. Bhutta, “A survey on machine learning based detection on ddos attacks for iot systems,” in *2019 SoutheastCon*. IEEE, 2019, pp. 1–6.

[9] F. Hussain, R. Hussain, S. A. Hassan, and E. Hossain, “Machine learning in iot security: Current solutions and future challenges,” *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1686–1721, 2020.

[10] P. Ghosal, D. Das, and I. Das, “Extensive survey on cloud-based iot-healthcare and security using machine learning,” in *2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. IEEE, 2018, pp. 1–5.

[11] Ddos. [Online]. Available: https://www.keycdn.com/support/ddos-attack

[12] N. N. Tuan, P. H. Hung, N. D. Nghia, N. V. Tho, T. V. Phan, and N. H. Thanh, “A ddos attack mitigation scheme in isp networks using machine learning based on sdn,” *Electronics*, vol. 9, no. 3, p. 413, 2020.

[13] Ddos attacks. [Online]. Available: https://www.imperva.com/learn/ddos/ddos-attacks/

[14] What is machine learning? a definition. [Online]. Available: https://www.expert.ai/blog/machine-learning-definition/

[15] Y. N. Soe, Y. Feng, P. I. Santosa, R. Hartanto, and K. Sakurai, “Machine learning-based iot-botnet attack detection with sequential architecture,” *Sensors*, vol. 20, no. 16, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/16/4372

[16] Understanding support vector machine(svm) algorithm from examples (along with code). [Online]. Available: https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/

[17] What is fuzzy logic. [Online]. Available: https://searchenterpriseai.techtarget.com/definition/fuzzy-logic

[18] Decision tree classification algorithm. [Online]. Available: https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm

[19] Machine learning in matlab. [Online]. Available: https://se.mathworks.com/help/stats/machine-learning-in-matlab.html

[20] S. R. Zahra and M. A. Chishti, “Fuzzy logic and fog based secure architecture for internet of things (flfsiot),” *Journal of ambient intelligence and humanized computing*, pp. 1–25, 2020.

[21] A. Haripriya and K. Kulothungan, “Secure-mqtt: an efficient fuzzy logic-based approach to detect dos attack in mqtt protocol for internet of things,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, p. 90, 2019.

[22] H. H. R. Sherazi, R. Iqbal, F. Ahmad, Z. A. Khan, and M. H. Chaudary, “Ddos attack detection: A key enabler for sustainable communication in internet of vehicles,” *Sustainable Computing: Informatics and Systems*, vol. 23, pp. 13–20, 2019.

[23] G. F. Scaranti, L. F. Carvalho, S. Barbon, and M. L. Proença, “Artificial immune systems and fuzzy logic to detect flooding attacks in software-defined networks,” *IEEE Access*, vol. 8, pp. 100 172–100 184, 2020.

[24] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, and A. Alazab, “A novel ensemble of hybrid intrusion detection system for detecting internet of things attacks,” *Electronics*, vol. 8, no. 11, p. 1210, 2019.

[25] E. Hodo, X. Bellekens, A. Hamilton, P.-L. Dubouilh, E. Iorkyase, C. Tachtatzis, and R. Atkinson, “Threat analysis of iot networks using artificial neural network intrusion detection system,” in *2016 International Symposium on Networks, Computers and Communications (ISNCC)*. IEEE, 2016, pp. 1–6.

[26] O. R and O. Yadav, “Detecting ddos attack using advanced support vector machines (asvm) in iot network,” vol. 7, pp. 5–19, 08 2020.

[27] A. Al Shorman, H. Faris, and I. Aljarah, “Unsupervised intelligent system based on one class support vector machine and grey wolf optimization for iot botnet detection,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 7, pp. 2809–2825, 2020.

[28] F. A. F. Silveira, F. Lima-Filho, F. S. D. Silva, A. d. M. B. Junior, and L. F. Silveira, “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.

[29] H. Cheng, J. Liu, T. Xu, B. Ren, J. Mao, and W. Zhang, “Machine learning based low-rate ddos attack detection for sdn enabled iot networks,” *International Journal of Sensor Networks*, vol. 34, no. 1, pp. 56–69, 2020.

[30] M. H. Aysa, A. A. Ibrahim, and A. H. Mohammed, “Iot ddos attack detection using machine learning,” in *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. IEEE, 2020, pp. 1–7.

[31] Y. N. Soe, Y. Feng, P. I. Santosa, R. Hartanto, and K. Sakurai, “Machine learning-based iot-botnet attack detection with sequential architecture,” *Sensors*, vol. 20, no. 16, p. 4372, 2020.

[32] R. Doshi, N. Apthorpe, and N. Feamster, “Machine learning ddos detection for consumer internet of things devices,” in *2018 IEEE Security and Privacy Workshops (SPW)*. IEEE, 2018, pp. 29–35.

[33] A. Mubarakali, K. Srinivasan, R. Mukhalid, S. C. Jaganathan, and N. Marina, “Security challenges in internet of things: Distributed denial of service attack detection using support vector machine-based expert systems,” *Computational Intelligence*, vol. 36, no. 4, pp. 1580–1592, 2020.

[34] A. Amouri, V. T. Alaparthy, and S. D. Morgera, “A machine learning based intrusion detection system for mobile internet of things,” *Sensors*, vol. 20, no. 2, p. 461, 2020.

[35] P. Nimbalkar and D. Kshirsagar, “Feature selection for intrusion detection system in internet-of-things (iot),” *ICT Express*, 2021.