Detection and Mitigation of Known and Unknown DDoS Attacks on Advanced Metering Infrastructure Systems in Nigeria Using Hybrid Machine Learning (Ai) Techniques

Oluwole Solanke* Oludele Awodele
Department of Computer Science, Babcock University, Ilisan-Remo, Ogun, Nigeria.
*E-mail of the corresponding author: solanke0069@pg.babcock.edu.ng

Abstract
The enormous rise of network traffic and its diversity on the Internet have posed new and serious obstacles for detecting network attack activity. Distributed denial of service (DDoS) attack is designed to restrict genuine users from accessing a service for an extended period of time. In this attack, the attacker attempts to compromise a large number of hosts to transmit a large volume of traffic to genuine users. Detecting DDoS attacks is difficult and complicated, primarily different DDoS attacks do not common characteristics through which they can be detected. DDoS attacks are very difficult to fight or trace due to their distributed nature, and automated software tools for conducting DDoS attacks are freely available. This paper proposed a DDoS detection model based on hybrid machine learning technique on AMI systems in the Nigeria Utility Business. It has been discovered that detecting unknown DDoS attacks is difficult to analyze as sometimes the IP packets and header are encrypted. This study proposed a combination of Support Vector Machine (SVM) and Artificial Neural Network (ANN) to detect unknown attacks. The AES 256 algorithm has been employed to decrypt the encrypted IP header.

Keywords: ANN, SVM, DDoS attack, AES algorithm, AMI
DOI: 10.7176/NCS/13-04
Publication date: May 31st 2022

1. Introduction

According to Jia et al. (2017), the enormous rise of network traffic and its diversity on the Internet have posed new and serious obstacles for detecting network attack activity. Traditional detection methods and approaches, particularly for DDoS attacks, have not satisfied the demands of efficient and precise detection for the diversity and complexity of attack traffic in the high-speed network environment as described by Jia et al., 2017. Distributed denial of service has become a major hazard in networking environments as it is designed to restrict genuine users from accessing a service for an extended period of time. In this attack, the attacker attempts to compromise a large number of hosts to transmit a large volume of traffic to genuine users. As a result, the service is unavailable for an extended period of time.

According to Douligeris and Mitrokotsa (2004), DDoS attacks are tricky in nature since they use the same strategies as traditional DoS attacks but on a much bigger scale where the attacker creates a defense force in order to attack in the form of zombies or botnets. A bot refers to a host that is under the attacker's control, while a botnet refers to a network of controlled computers. All these bots are trained to attack the victim and cripple the functionality of the victim system while their legitimate controllers are unaware of that. Najafimehr et al. (2022) described how, in modern attack approaches, attackers forge the target’s IP address and make requests to servers throughout the Internet, resulting in a massive volume of traffic being sent to the target. This type of attack is known as a reflection attack, and the servers are called reflectors. These attacks are usually amplified, called amplification attacks, in which the size of the response traffic is much larger than the requests' size sent by the attacker. Attackers may launch this type of attack by sending a tiny request, querying a list of information. According to the Kaspersky reports, the number of DDoS attacks in 2020 compared to 2019 and 2018 has grown by approximately 88% and 121%, respectively.

The growth of metering devices and infrastructure around the world has been driven by technological advancements, resulting in more accurate and user-friendly equipment with customer contact interfaces. Despite the implementation of various reforms and policies in the Nigerian Utility industry, the evolution of metering technology began with the unbundling of the National Electric Power Authority (NEPA). However, despite the implementation of various reforms and policies in the Nigerian utility industry, it has not progressed smoothly and successfully. The AMI (Advanced Metering Infrastructure) has been adopted by Nigeria utility business to overcome the persisting problems in the business. AMI system provides the utility with real-time data about power consumption and allows customers to make informed choices about energy usage based on the price at the time of use (multi-tariff). It is not a single technology, but rather an integration of many technologies that provides an intelligent connection between consumers and power utilities.

However, distributed denial of service attacks (DDoS) are growing in size and frequency, and it is found that there are an average of 124,000 DDoS events each week and that the peak attack size has grown 73% in the last twelve months, according to research from Arbor Network. Falling prey to a DDoS attack can result in hours of downtime for power utilities, businesses, and consumers as their service and deployed systems are disrupted and made unavailable. Unlike advanced persistent threats, DDoS attacks do not require extreme technical sophistication on the part of cybercriminals. Furthermore, hacker incentives for initiating DDoS assaults are not always, or even typically, monetary, but can be a way for them to demonstrate their skills or achieve credibility among other aspiring criminals.

In light of the aforementioned issues, this paper proposed a DDoS detection model based on hybrid machine learning techniques. The key contribution of this work may be stated as follows:

i. To distinguish between attacks and normal flows in the AMI system, this study proposed a combination of Support Vector Machine (SVM), which is a kind of supervised learning technique to detect known attacks, and the unsupervised learning of an Artificial Neural Network (ANN) to detect unknown attacks based on characteristic patterns.

ii. The study proposes an ensemble feature selection technique to select the most effective features from the given dataset features.

iii. While it has been discovered that detecting unknown DDoS attacks is difficult to analyze as sometimes the IP packets and header are encrypted, the AES 256 algorithm has been employed to decrypt the encrypted IP header.

2. Literature Review

There has been a lot of research done on applying machine learning and artificial intelligence approaches to identify and prevent DDoS assaults. Shieh et al., (2021), propose a new DDoS detection method based on Bi-Directional Long Short-Term Memory (BI-LSTM), a Gaussian Mixture Model (GMM), and incremental learning. Unknown traffic recorded by the GMM is subjected to traffic engineers' discrimination and tagging, and then sent back into the framework as new training samples. Experiments show that the proposed framework is a viable option for the detection of unknown DDoS attacks. Chen et al., 2021) proposed statistical and machine-learning algorithms to model and analyze the network traffic of DDoS attack aiming at the problem of DDoS attack detection in internet of things (IoT) environment. The authors described that on the internet of things simulation experimental environment, a DDoS attack detection method combined with statistical analysis and machine learning may efficiently identify large-scale DDoS attacks. The IoT network traffic data is collected using a Docker-based virtualization architecture. Different machine learning algorithms are investigated and compared, as well as the relative value of characteristics in various traffic datasets. The KNN method produces the best results in packet-level DDoS attack detection; the accuracy is 92.8 percent and 99.8 percent, respectively, and the voting approach has the highest accuracy (99.8 percent). The RNN method produces the greatest results in second-level discovery, with an accuracy of 97.1 percent.

DDoS attack aims to prevent legitimate users from getting access to a targeted system service by exhausting the resources, bandwidth and so on as described by Mohammed, (2021). Having an automated system that can learn the nature of the attack and instantly detect it is the reason why machine learning is used in this work. Decision tree, KNN and Naïve Bayes are the algorithms used classify a benign traffic from a DDoS attack. The results of the experiment indicate that Decision tree and KNN proved to be the most effective with an accuracy of 100% and 98% respectively. Saied et al., (2016) proposed a model that used an artificial neural network (ANN) technique to identify TCP, UDP, and ICMP DDoS assaults, distinguishing actual traffic from DDoS attacks, and conducting in-depth training on the system using real examples created by common DDoS tools and attack types. In research by (Verma & Kumar, 2021), many machine learning algorithms for detecting DoS/DDoS assaults are reviewed where the author explained that cyberattacks of many types, such as ransomware, phishing, and DOS / DDOS attacks, are becoming more common. Machine Learning is assisting in the training of models to detect attacks and then prevent them from causing maximum damage.

As described by (Khempetch & Wuttidittachotti, 2021), itis anticipated that by 2030, 125 billion gadgets will be in use, resulting in a massive influx of data. It is impossible to add extra security structures to IoT devices due to their restricted resources. The work proposed a deep neural network (DNN) and long short-term memory (LSTM) algorithms to detect DDoS attacks. According to the findings, it can detect over 99.90 percent of all three forms of DDoS attacks. In research from (Sambangi & Gondi, 2020) which aimed to investigate the subject of DDoS attack detection in a Cloud context. The work explained that detecting Distributed Denial of Service (DDos) threats is primarily a machine learning classification challenge. DDoS attacks are most common in the seven-layer OSI model's network, transportation, display, and application layers.

(Pei et al., 2019) presents a machine learning-based DDoS assault detection approach that incorporates two steps: feature extraction and model detection. The feature extraction stage entails analyzing vast volumes of data packages that have been categorized using rules. The recovered features are employed as machine learning input features in the model discovery stage, and the random forest approach is used to train the attack detection model. The results of the experiments reveal that the suggested DDoS attack detection approach has a high detection rate for today's most common DDoS attacks. Denial-of-service (DoS) attacks are continually affecting users and Internet service providers. (Lima Filho et al., 2019) presented a DoS detection system based on machine learning (ML). The proposed method makes inferences based on signatures collected from network traffic samples. The results reveal that attacks are detected online at a rate of more than 96 percent, with high precision (PREC) and low false alarm rates (FAR).

Girma & Wang, (2018) investigates the multivariate correlation between the selected and ranking features in order to meet this need of detecting DDoS attacks and serve the legitimate users with available resources with minimal downtime. The paper covers the research findings and visualizes the experimental data to highlight the degree of parametrical dependency among the selected features and the efficiency of our multivariate correlational technique. Table 1 lists the summary of the few reviewed literature in this study.

Table 1: Summary of Reviewed Literature

S/N	Name of Authors	Name of Journal/Year/Vol/Issue	Issue Addressed	Findings	Gap in the Study
1.	(Mohammed, 2021)	A Machine Learning-Based Intrusion Detection of DDoS Attack on IoT Devices. Vol. 10 Issue (4), July – August 2021, 2792 – 2797.	Decision tree, KNN and Naïve Bayes are the algorithms used classify a benign traffic from a DDoS attack.	The results demonstrate that Decision tree and KNN proved to be the most effective with an accuracy of 100% and 98% respectively.	Naïve Bayes gave a very poor result with an accuracy of 29% as it struggled to forecast DDoS attacks.
2.	(Chen et al., 2021)	DDoS Attack Simulation and Machine Learning-Based Detection Approach in Internet of Things Experimental Environment. 2021, Volume: 15, Issue: 3, ISSN: 1930-1650.	A statistical and machine-learning algorithm was proposed to model and analyze the network traffic of DDoS attack while Docker-based virtualization platform is designed and configured to collect IoT network traffic data.	The results revealed that the KNN algorithm produces the highest results in packet-level DDoS attack detection, with an accuracy of 92.8 percent. The voting method delivers the greatest results in flow-level DDoS attack detection, with an accuracy of 99.8%. The RNN method performs best in second-level DDoS attack detection, with a 97.1 percent accuracy.	It can only effectively detect large-scale DDoS attack on the IoT simulation experimental environment.
3.	(Shieh et al., 2021)	Detection of Unknown DDoS Attacks with Deep Learning and Gaussian Mixture Model; 2021, *Vol.11, Issue* 11, 5213.	New DDoS detection framework featuring Bi-Directional Long Short-Term Memory (BI-LSTM), a Gaussian Mixture Model (GMM), and incremental learning was adopted.	Experiment findings show that the proposed BI-LSTM-GMM can achieve recall, precision, and accuracy of up to 94 percent using the data sets CIC-IDS2017 and CIC-DDoS2019 for training, testing, and assessment. Experiments show that the suggested framework has the potential to be a viable solution for detecting unknown DDoS attacks.	No validation of more datasets and the BI-LSTM and GMM is not auto-configured.

S/N	Name of Authors	Name of Journal/Year/Vol/Issue	Issue Addressed	Findings	Gap in the Study
4.	(Verma & Kumar, 2021)	DOS/DDOS Attack Detection using Machine Learning: A Review	Many machine learning techniques are reviewed to detect DoS/DDoS attacks in this study.	The study reviewed that of all the detection system using machine learning reviewed, Random Forest and CatBoost algorithm has given best accuracy 99.99%.	The review is limited to machine learning techniques only
5.	(Khempetch & Wuttidittachotti, 2021)	DDoS attack detection using deep learning, Vol. 10, No. 2, June 2021, pp. 382~388 ISSN: 2252-8938.	A DDoS attack detection using the deep neural network (DNN) and long short-term memory (LSTM) algorithm was proposed.	A CICDDoS2019 dataset was used as a dataset that has improved the bugs and introducing a new taxonomy for DDoS attacks, including new classification based on flows network. The results show that it can detect more than 99.90% of all three types of DDoS attacks.	The study is limited to the detection of types of Syn Flood, UDP Flood, and UDP-Lag only. Also, the solution has not been tested with other attack types to compare the performance of the model.
6.	(Sambangi & Gondi, 2020)	A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression, 2020, Vol. 63, Issue 1.	Multiple regression analysis for building a machine learning model to predict DDoS and Bot attacks through considering a Friday afternoon traffic logfile was employed.	An ensemble model for the Friday morning dataset achieves prediction accuracy of 97.86 percent and 73.79 percent, respectively, for 16 characteristics obtained via information gain-based feature selection and regression analysis-based ML model. This research paved the way for future research into the value of regression analysis in the construction of machine learning models.	The analysis is limited to one-day log file.

S/N	Name of Authors	Name of Journal/Year/Vol/Issue	Issue Addressed	Findings	Gap in the Study
7.	(Lima Filho et al., 2019)	Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning. Security and Communication Networks, 2019, 1–15.	A machine learning- (ML-) based DoS detection system was proposed in this study	The software uses the Random Forest Tree algorithm to classify network traffic based on samples taken by the sFlow protocol directly from network devices. The software was evaluated based on three intrusion detection benchmark datasets, namely, CIC-DoS, CICIDS2017, and CSE-CIC-IDS2018, and was able to classify various types of DoS/DDoS attacks, such as TCP flood, UDP flood, HTTP flood, and HTTP slow.	The system still needs improvements on a better hit rate among attack classes and an automatic parameter calibration mechanism that maximizes the detection rate of attacks.
8.	(Pei et al., 2019)	A DDoS Attack Detection Method Based on Machine Learning. Volume 1237, Issue 3, 2019.	A Random Forest algorithm model based on machine learning was proposed.	The experimental results reveal that the proposed machine learning-based DDoS attack detection approach has a high detection rate for the most common DDoS attack.	The model has a low detection rate for uncommon DDoS attack.
9.	(Girma & Wang, 2018)	An Efficient Hybrid Model for Detecting Distributed Denial of Service (DDoS) Attacks in Cloud Computing Using Multivariate Correlation and Data Mining Clustering Techniques. Volume 19, Issue 2, pp. 1-12, 2018	Hybrid solution model using DBSCAN and Entropy was proposed.	The experimental results revealed that the hybrid model is effective and efficient in detecting the DDoS attacks.	There is need for improvement on DBSCAN clustering area to improve its capability of handling big data.

S/N	Name of Authors	Name of Journal/Year/Vol/Issue	Issue Addressed	Findings	Gap in the Study
10.	(Saied et al., 2016)	Detection of known and unknown DDoS attacks using Artificial Neural Networks, 2016, Volume 172, Issue 1, ISSN 0925-2312.	Based on typical characteristics that distinguish actual traffic from DDoS attacks, the author uses an Artificial Neural Network method to detect TCP, UDP, and ICMP DDoS attacks.	The ANN learning process began with the replication of a network environment that reflects a real-world environment. Then, while normal traffic was flowing through the network, various DDoS attacks were started. In comparison to other ways, the solution was accessed using signature-based and other related academic studies, and the methodology provided higher detection accuracy (98%) than others.	The solution cannot detect DDoS attacks that use encrypted packet headers.

As described by Girma and Wang (2018), cloud computing is made vulnerable and prone to sophisticated distributed intrusion attacks like distributed denial of service (DDoS) attacks due to its distributed nature. To identify network assaults and respond quickly, a reliable security system that can discern anomalies embedded in genuine data is required. Existing systems for monitoring incoming traffic and identifying DDoS assaults generate a high number of false alarms, making them inefficient in detecting high-level flooding attacks and resolving cloud service availability concerns. There have been few studies on the Nigerian utility business; that is, it is worth mentioning that there is little evidence of studies on the detection and mitigation of known and unknown DDoS attacks on AMI Systems in the Nigeria Utility Business using hybrid machine learning (AI) techniques, resulting in a research gap.

3. Related Work

Existing models focus on DDoS attacks and victim attributes, but do not focus on botnet attributes, and botnet becomes the main technology of DDoS organization and management. The key goal of distributed denial of service is to compile multiple systems and form botnets using infected zombies/agents over the Internet. The purpose is to attack a specific target or network with different types of packets. The infected system is controlled remotely by an attacker or a self-installed Trojan.

Saied et al., (2016) proposed a model that used an artificial neural network (ANN) technique to identify TCP, UDP, and ICMP DDoS assaults, distinguishing actual traffic from DDoS attacks, and conducting in-depth training on the system using real examples created by common DDoS tools and attack types. Katkar et al., (2015) offer a signature-based network intrusion detection model for detecting DDoS attacks on HTTP servers. A distributed processing system and a naive Bayesian classifier are included in the model. The efficacy of the proposed model is demonstrated using observational data. The naive Bayesian can only classify sluggish attacks with 97.82 percent precision and regular attacks with 96.46 percent precision.

Khan et al., (2016) propose an artificial neural network-based technique for detection of distributed flooding attacks to things such as sensors or actuators in WMNs called the distributed flood attack detector. In our simulation, sample dataset used to train and test the artificial neural network is generated using NS-2 network simulator. Simulation results and real system implementation proved that the distributed flood attack detector can be used in a real network environment to detect the intermediate and severe distributed flood attacks with low-false positive and false-negative rates.

Mahmoud et al., (2020) discuss how current approaches use Machine Learning (ML) for intrusion detection against DDoS attacks in the SDN network using standard datasets in this study. However, these methods have a number of flaws, including the fact that the datasets employed do not contain the most recent attack trends, resulting in a lack of attack diversity. The author further described that DDoSNet is a deep learning-based intrusion detection system that combines the Recurrent Neural Network (RNN) with an autoencoder. The proposed model was trained and evaluated using the newly released CICDDoS2019 dataset, which contains comprehensive and most recent DDoS types of attacks. When compared to existing well-known traditional machine learning techniques, DDoSNet provides the highest evaluation metrics in terms of recall, precision and accuracy.

Ugwu et al. [11] compared the results of Naive Bayes (NB), Decision Tree (DT), and Support Vector Machine (SVM) machine learning techniques (SVM). The deep learning algorithms LSTM and Singular Value Decomposition (SVD) suggested show a considerable improvement. The network data undergoes pre-processing, which includes data normalization and feature conversion algorithms. The normalizing method necessitates restricting network feature values to a small range. Non-numeric features must be converted to numeric features using the feature conversion method. Kasim [12] classified encoded data as DDoS or normal using dimensional reduction features in the autoencoder (AE) model and a Support Vector Machine (SVM) classifier. The AE-SVM algorithm is capable of distinguishing between regular and DDoS assault traffic. The training vectors for the AE model were constructed using the min-max approach to standardize their data between 0 and 1. The trained model supplied feature learning and feature reduction via the encoding process. The AE-SVM approach worked well in terms of low false-positive DDoS detection rates and quick anomaly identification, according to the findings.

Karan, Narayan, and Hiremath, (2018) proposed two DDoS detection models for SDN networks. To collect network traffic in the first stage, a signature-based snort detection system was employed. SVM and Deep Neural Network (DNN) algorithms are used to classify attacks in the final step. The authors used the KDDCUP99 dataset to train the two detection modules, which were based on 7 out of 41 characteristics. The results of the experiment showed that the DNN outperforms the SVM, with accuracy rates of 92.30 and 74.30 percent, respectively. Mohammed et al. (2018) suggested a new methodology for detecting DDoS attacks on SDN. The NB classifier was trained using the NSL-KDD dataset with 25 chosen features. The authors use a combination of three different selection algorithms (Genetic, Ranker, and Greedy) to choose characteristics from the dataset. Precision, recall, and F1-score had average values of 0.81, 0.77, and 0.77.

(OT, 2019) presented a background information on intrusion detection methodologies as well as a brief description of intrusion detection systems in this study. An Association rule-based method was created for mining known known-patterns in this study. The results of the produced tools are adequate, although they may be better. These tools will go a long way toward resolving data security issues by discovering security flaws in computer systems. (Raman, Ram, Anitha, 2018) basically made three contributions: first, the authors introduced an abstract model for the aforementioned class of attacks, where the botnet emulates normal traffic by continually learning admissible patterns from the environment; second, devise an inference algorithm was devised which was shown to provide a consistent (i.e., converging to the true solution as time elapses) estimate of the botnet possibly hidden in the network; and third, the validity of the proposed inferential strategy on a test bed environment was verified. The test results showed that for several scenarios of implementation, the proposed botnet identification algorithm needs an observation time in the order of (or even less than) one minute to identify correctly almost all bots, without affecting the normal users' activity.

(Myint Oo et al., 2019) present a DDoS attack detection approach based on SDN that will cause the least amount of disruption to normal user activity. With the shortest training and testing times, our detection approach has a detection accuracy of around 97 percent. Using two essential features, namely the volumetric and asymmetric features, our detection technique can cut both training and testing time. (Azeez et al., 2019) described an intrusion detection and prevention system (IDPS) as a device or software program that monitors a network or system. The authors explained that to keep up with the progression of computer-related crimes, it detects vulnerabilities, reports harmful activity, and implements preventive measures. Given that the most current review on the issue was published in 2016, the paper provides an updated review of IDPSs. It will also cover the usage of intrusion detection systems (IDPS) to discover vulnerabilities in various channels via which data is accessed on a network, as well as the prevention techniques used to prevent infiltration.

According to (Awujoola et al., 2021), the problem of network penetration poses a major hazard to all computer users. There is a need for new intrusion detection solutions that are both efficient and effective while also having a low false alarm rate. The research provides a new intrusion detection methodology that combines genetic algorithms with Synthetic Minority Over Sampling (SMOTE) and Resample techniques for class distribution balance. On the KDDCUP99 and NSL-KDD datasets, performance accuracy of 99.9873 percent and 99.8457 percent, respectively, was attained in contrast to other state-of-the-art detection algorithms. In the research from (Irhebhude et al., 2022), an ensembled classifier was used to reduce cyberspace uncertainty. To train the machine learning model, the classification learner in MATLAB was employed. On the merged Comma Separated Value (CSV) UNSW-NB15 dataset and the self-acquired dataset, the studies yielded outstanding classification accuracy of 99.1 percent and 99.4 percent, respectively. When compared to an artificial neural network classifier, ensemble provided a more reliable classification accuracy. (Bandara et.al., 2016) presents a comprehensive survey of preventing DDOS attack recognize by data mining techniques with the use of identifying DDOS attack patterns and analyze patterns by machine learning algorithms. The authors explained that there are some leading machine learning algorithms used to recognize the DDOS attack such as k-Nearest

Neighbors algorithm (KNN), support vector machines (SVM), Random Forest as well as Naïve Base. In addition, the paper highlights open issues, research challenges and possible solutions in this area. The result shows the highest accuracy rate of preventing DDOS attack recognizing by data mining algorithms.

4. Methodology

A. Nature of DDoS Attack

The distributed nature of DDoS attacks makes them extremely difficult to resist or trace. Intruders may also use IP spoofing to conceal their identity, making DDoS assault identification more difficult. The signature-based detection method captures network traffic, which is then compared to well-defined attack patterns, such as packet sequences or bytes. This type of detection strategy is very simple to learn and produce more meaningful findings. The signature-based detection technique, on the other hand, can only detect known attacks with a pre-defined pattern. Using behavioral patterns, the anomaly-based detection scheme is employed to detect the attack and unknown attack can be identified using this detection approach. Existing models concentrate on DDoS assaults and victim characteristics but ignore botnet characteristics, resulting in botnets being the primary technology for DDoS organization and control (Manoj et al., 2020). The main purpose of distributed denial of service is to assemble several computers and build botnets via the Internet using infected zombies/agents. The goal is to use various sorts of packets to attack a specific target or network. An attacker or a self-installed Trojan controls the compromised system remotely.

B. Machine Learning

Machine learning has been used in the realm of security in recent years (Meng, Rice, and Wang, 2018). Machine learning can extract key information from data and combine it with existing knowledge to discriminate and predict new data (Jordan & Mitchell, 2015). As a result, as compared to traditional detection approaches, machine-learning algorithms can be more accurate. The traditional network environment, cloud environment, and software-defined network architecture all entail attack detection for DDoS defensive mechanisms. Artificial neural networks (ANN) and support vector machines (SVM) are two of the most commonly used AI-based techniques to detect distributed denial of service (DDoS) attacks.

C. Hybrid Detection System

Since the beginning of intrusion detection research, the integration of misuse and anomaly intrusion detection has been researched. Misuse intrusion detection can detect attacks precisely with a low false alarm rate, but it can't detect novel attacks; anomaly intrusion detection, on the other hand, can detect unique attacks but has a higher false alert rate. As a result, combining these two strategies to maximize their strengths while minimizing their flaws may yield better outcomes which forms a hybrid approach. In a study conducted by (Ghosh et al., 2014) **a** hybrid multilevel intrusion detection model known as KNN NN was used. The study employs the K-Nearest Neighbor (KNN) method for binary classification, which divides data into 'normal' and 'abnormal' categories. The abnormal class is divided into four major attack kinds using a neural network. It further employs Rough Set Theory and Information Gain for feature selection instead of Principal Component Analysis, which was used in this investigation. The study picked 25 features from the NSL-KDD dataset using these methods, and the experiment was done for 25 and 41 features, while this investigation was done for three categories: 13, 25, and 35 features. Palanisamy et al., (2018) proposed a system that uses machine learning techniques to reduce the number of false positives and false negatives. Existing Intrusion Detection Systems (IDS) data projections that can protect networks from future attacks, with observed development or incursion regularly disclosed ahead of time or gathered midway. It was depicted as if the concept for constructing a secure, intelligent intrusion detection system could protect networks and computers from attack.

D. Data Pre-processing

Data pre-processing is a type of processing that is applied to raw data in order to prepare it for further processing. It is commonly employed as a.preparatory mining procedure since it turns data into a format that can be processed more simply and appropriately according to the user's needs. Association rule mining is used to pre-process data. If/then expressions in a relational database or other archive assist find relationships between seemingly unrelated data. The proposed hybrid detection system pre-process data by which size of data attributes can be reduced and only selected features are extracted for classifiers training and testing.

E. Dimensionality Reduction

Data overload is caused by high-dimensional data, which significantly increases the computing time and storage space needs of data processing (Gui et al., 2017). Dimensionality reduction is used in many data mining and machine learning methods to convert high-dimensional space to low-dimensional space. Text categorization, picture classification, intrusion detection, genome analysis, and other applications are some of the most popular uses of dimensionality reduction techniques. Dimensionality reduction is useful for compressing features by removing redundant and unnecessary data, which enhances the model's efficiency and performance as described by (Khalid, Khalil and Nasreen, 2014).

Feature selection is a technique used to select key data points and explain how they relate to one another. It aids in the simplification of models and the reduction of training and testing time in order to generate unique outcomes.

F. Classification Process

- **Support Vector Machine**

Support vector machine (SVM) is already well-known as the finest binary classification learning technique. SVM, which started out as a pattern classifier based on a statistical learning approach for classification and regression using a range of kernel functions, has now been effectively used to a variety of pattern recognition applications. One of the key benefits of utilizing SVM for IDS is its speed, which is critical since the capacity to identify intrusions in real time is critical. Because it does not rely on traditional neural networks like neural networks, the SVM may choose appropriate parameters.

- **Artificial Neural Network**

Artificial Neural Network (ANN) is a mathematical approach that works similarly to the human brain. The artificial neural network (ANN) is a supervised approach that learns from a collection of training data and then compares the new data to the target sample. If there is an error at the output, the value of the error is transmitted to the hidden layer, which modifies the weight matrix. The general structure of ANN is depicted in figure 1 below. Pandeeswari and Ganesh, (2016), described the working of ANN algorithm as shown in the following process:

{X=X1, X2, X3……………………Xn} are the input presented at the input layer of ANN.

Model while Theta (θ) is the weight function used for modification of the hidden later weight value.

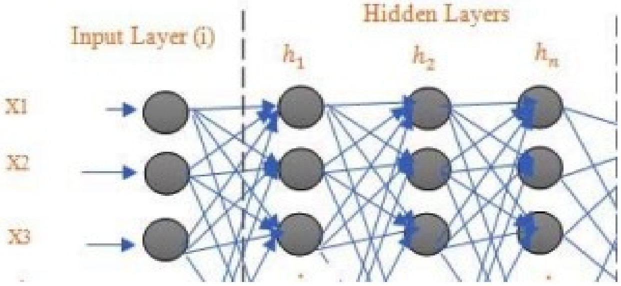

Figure 1: Structure of ANN Algorithm

5. Proposed Model

The proposed technique for network intrusion detection is described in this section. For successful detection, the hybrid strategy SVM with ANN is used in this study. The proposed technique is depicted in Figure 2.

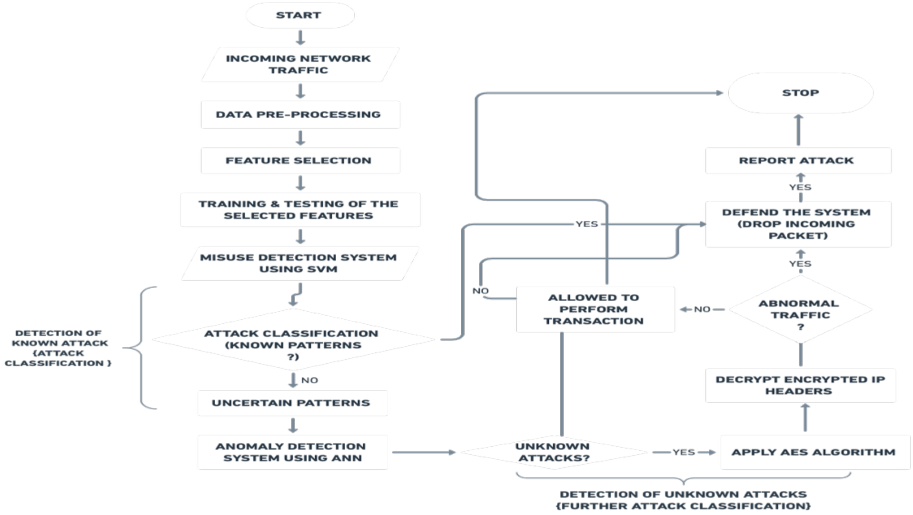
START
INCOMING NETWORK TRAFFIC
DATA PRE-PROCESSING
FEATURE SELECTION
TRAINING & TESTING OF THE SELECTED FEATURES
MISUSE DETECTION SYSTEM USING SVM
DETECTION OF KNOWN ATTACK (ATTACK CLASSIFICATION)
ATTACK CLASSIFICATION (KNOWN PATTERNS ?)
NO
UNCERTAIN PATTERNS
ANOMALY DETECTION SYSTEM USING ANN
UNKNOWN ATTACKS?
YES
APPLY AES ALGORITHM
DECRYPT ENCRYPTED IP HEADERS
ABNORMAL TRAFFIC ?
NO
ALLOWED TO PERFORM TRANSACTION
NO
YES
DEFEND THE SYSTEM (DROP INCOMING PACKET)
YES
YES
REPORT ATTACK
STOP
DETECTION OF UNKNOWN ATTACKS (FURTHER ATTACK CLASSIFICATION)

Figure 2: Proposed Hybrid DDoS Detection Model

To distinguish between attacks and normal flows, this study proposed a combination of Support Vector Machine (SVM), which is a kind of supervised learning technique to detect known attacks, and the unsupervised learning of an Artificial Neural Network (ANN) to detect unknown attacks based on characteristic patterns. The proposed hybrid system begins by accepting incoming network traffic (real-time or batch upload), where the traffic is monitored, datasets are preprocessed, and statistical data is sent for feature selection, which is designed to reduce the number of input variables that are considered most appropriate for the model to predict the target variable. The selected characteristics are then transmitted to be classified, resulting in a dataset for the proposed system to work with. The data is categorised using UDP/ICMP, HTTP/XML, TCP-SYN, and Ping attack patterns in SVM. The attack probability is calculated based on the SVM's output. The packet data is set if there is a chance of an attack and the attack probability is calculated using the SVM result. If an attack is possible, the packet data is configured with termination rules to prevent the packet from being attacked. If the output data does not have a high probability of attack but is suspicious, it is sent to ANN for further classification. When a server receives a Quit/quit message from one of these clients after an attack has been identified and reported, it will close the connection with that client, dropping or deleting that request while normal network traffic is allowed to complete its transaction. In the event of an unknown attack, the AES method is used to decrypt the IP header, and the suspect IP address is banned for a set length of time (e.g. 12Hr). When a server receives many requests for the same service, the AES algorithm is used to monitor the IP address. The MAC IP address of the rogue MAC will be blocked. The AES algorithm can be used to detect encrypted headers.

6. Conclusion

The main contribution of this work is the development of a classification model for the problem that has high intrusion detection accuracy and, more importantly, low false negatives. This was accomplished by combining support vector machines (SVM) and artificial neural networks (ANN) for the detection of known and unknown attacks. The AES 256 technique was also used to decrypt the encrypted IP header since it has been established that detecting unexpected DDoS assaults is difficult to evaluate because IP packets and headers are often encrypted.

Reference

Awujoola, O. J., Ogwueleka, F. N., Irhebhude, M. E., & Misra, S. (2021). Wrapper Based Approach for Network Intrusion Detection Model with Combination of Dual Filtering Technique of Resample and SMOTE. Artificial Intelligence for Cyber Security: Methods, Issues and Possible Horizons or Opportunities, 139–167. https://doi.org/10.1007/978-3-030-72236-4_6

Azeez, N. A., Bada, T. M., Misra, S., Adewumi, A., van der Vyver, C., & Ahuja, R. (2019). Intrusion Detection and Prevention Systems: An Updated Review. Data Management, Analytics and Innovation, 685–696. https://doi.org/10.1007/978-981-32-9949-8_48

Balogun, B. F., Gbolagade, K. A., Arowolo, M. O., & Saheed, Y. K. (2021). A Hybrid Metaheuristic Algorithm for Features Dimensionality Reduction in Network Intrusion Detection System. Computational Science and Its Applications – ICCSA 2021, 101–114. https://doi.org/10.1007/978-3-030-87013-3_8

Chen, H., Meng, C., & Chen, J. (2021). DDoS Attack Simulation and Machine Learning-Based Detection Approach in Internet of Things Experimental Environment. International Journal of Information Security and Privacy, 15(3), 1–18. https://doi.org/10.4018/ijisp.2021070101

Girma A.& Wang P., (2018). An Efficient Hybrid Model For Detecting Distributed Denial Of Service (Ddos) Attacks In Cloud Computing Using Multivariate Correlation And Data Mining Clustering Techniques. Issues In Information Systems. Published. https://doi.org/10.48009/2_iis_2018_1-12

Irhebhude, Martins & Musa, Zahra'u & Kolawole, Adeola. (2022). Uncertainties Classification in Cyberspace using Ensemble Learning Model. 17. 69-76.

Jia, B., Huang, X., Liu, R., & Ma, Y. (2017). A DDoS Attack Detection Method Based on Hybrid Heterogeneous Multiclassifier Ensemble Learning. Journal of Electrical and Computer Engineering, 2017, 1–9. https://doi.org/10.1155/2017/4975343

J. Gui, Z. Sun, S. Ji, S. Member, D. Tao, and T. Tan, "Feature Selection Based on Structured Sparsity : A Comprehensive Study," IEEE Trans. Neural Networks Learn. Syst., vol. 28, no. 7, pp. 1490–1507, 2017

KASIM, M. (2020). An efficient and robust deep learning based network anomaly detection against distributed denial of service attacks. Computer Networks, 180, 107390. https://doi.org/10.1016/j.comnet.2020.107390

Khan, M.A., S. Khan, B. Shams and J. Lloret, 2016. Distributed flood attack detection mechanism using artificial neural network in wireless mesh networks.

Khempetch, T., & Wuttidittachotti, P. (2021). DDoS attack detection using deep learning. IAES International Journal of Artificial Intelligence (IJ-AI), 10(2), 382. https://doi.org/10.11591/ijai.v10.i2.pp382-388

Lima Filho, F. S. D., Silveira, F. A. F., de Medeiros Brito Junior, A., Vargas-Solar, G., & Silveira, L. F. (2019). Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning. Security and Communication Networks, 2019, 1–15. https://doi.org/10.1155/2019/1574749

Mohammed S., (2021). A Machine Learning-Based Intrusion Detection of DDoS Attack on IoT Devices. (International Journal of Advanced Trends in Computer Science and Engineering, 10(4), 2792–2797. https://doi.org/10.30534/ijatcse/2021/221042021

Myint Oo, M., Kamolphiwong, S., Kamolphiwong, T., & Vasupongayya, S. (2019). Advanced Support Vector Machine- (ASVM-) Based Detection for Distributed Denial of Service (DDoS) Attack on Software Defined Networking (SDN). Journal of Computer Networks and Communications, 2019, 1–12. https://doi.org/10.1155/2019/8012568

Najafimehr, M., Zarifzadeh, S., & Mostafavi, S. (2022). A hybrid machine learning approach for detecting unprecedented DDoS attacks. The Journal of Supercomputing, 78(6), 8106–8136. https://doi.org/10.1007/s11227-021-04253-x

Pei, J., Chen, Y., & Ji, W. (2019). A DDoS Attack Detection Method Based on Machine Learning. Journal of Physics: Conference Series, 1237(3), 032040. https://doi.org/10.1088/1742-6596/1237/3/032040

Sambangi, S., & Gondi, L. (2020). A Machine Learning Approach for DDoS (Distributed Denial of Service) Attack Detection Using Multiple Linear Regression. The 14th International Conference on Interdisciplinarity in Engineering—INTER-ENG 2020. https://doi.org/10.3390/proceedings2020063051

Saied, A., Overill, R. E., & Radzik, T. (2016). Detection of known and unknown DDoS attacks using Artificial Neural Networks. Neurocomputing, 172, 385–393. https://doi.org/10.1016/j.neucom.2015.04.101

Shieh, C. S., Lin, W. W., Nguyen, T. T., Chen, C. H., Horng, M. F., & Miu, D. (2021). Detection of Unknown DDoS Attacks with Deep Learning and Gaussian Mixture Model. Applied Sciences, 11(11), 5213. https://doi.org/10.3390/app11115213

S. Khalid, T. Khalil, and S. Nasreen, (2014). “A survey of feature selection and feature extraction techniques in machine learning,” Proc. 2014 Sci. Inf. Conf. SAI 2014, pp. 372–378.

Verma, V., & Kumar, V. (2021). DOS/DDOS Attack Detection using Machine Learning: A Review. SSRN Electronic Journal. Published. https://doi.org/10.2139/ssrn.3833289

Oluwole Solanke (He is a master’s degree holder with an M.Sc. in Computer Science from Babcock University, B.Eng. in Information & Communication Engineering from Covenant University and Certification in Advance Project Management from Unilag Consult, University of Lagos. He is also an active recognized associate member of (COREN) The Council For the regulation of Engineering in Nigeria, an associate member of the Nigerian Institute of management, an associate member of the Nigerian Society of Engineers and the (IEEE) International Institute of Electrical Electronics Engineers.

Oludele Awodele is a Professor of Computer Science and Artificial Intelligence. He had his higher education at the University of Ilorin where he obtained his Bachelor’s degree in 1995 followed by Master’s and Ph.D degrees in Computer Science in Federal university of Agriculture, Abeokuta in 2002 and 2009 respectively. He was the Head of Department Computer Science (2009-2016), Dean, School of Computing and Engineering Sciences (2016-2020), and currently the Director of Academic Planning, Babcock University.