REPUBLIC OF TURKEY
ALTINBAŞ UNIVERSITY
Institute of Graduate Studies
Electrical and Computer Engineering

NEW SMART GRID FRAMEWORK BASED OPTIMIZED AI METHOD: CASE STUDY ATTACK DETECTION

Abdulmunem Yaseen Khudhair KHUDHAIR

Master's Thesis

Supervisor

Prof. Dr. Osman Nuri UÇAN

Istanbul, 2022

NEW SMART GRID FRAMEWORK BASED OPTIMIZED AL METHOD: CASE STUDY ATTACK DETECTION

Abdulmunem KHUDHAIR

Electrical and Computer Engineering

Master's Thesis

ALTINBAŞ UNIVERSITY

2022

The thesis titled NEW SMART GRID FRAMEWORK BASED OPTIMIZED AI NEW SMART GRID FRAMEWORK BASED OPTIMIZED AI METHOD: CASE STUDY ATTACK DETECTION **prepared** by ABDUL MUNEM YASEEN KHUDHAIR KHUDHAIR and submitted on (20/05/2022) has been **accepted unanimously** for the degree of Master of Science in Electrical and Computer Engineering.

_____

Prof. Dr. Osman Nuri UÇAN

the Supervisor

Thesis Défense Committee Members:

Prof. Dr. Osman Nuri UÇAN	Faculty of Engineering and Natural Sciences, Altinbaş University	_____
Asst. Prof. Abdullahi Abdu IBRAHIM	Faculty of Engineering and Natural Sciences, Altinbaş University	_____
Dr. Mohammed Hasan ALWAN	Faculty of Engineering Bilad AlRafidain University	_____

I hereby declare that this thesis meets all format and submission requirements as Master's thesis.

Submission date of the thesis to the Graduate Education Institute: ___/___/___

I hereby declare that all information/data presented in this graduation project has been obtained in full accordance with academic rules and ethical conduct. I also declare all unoriginal materials and conclusions have been cited in the text and all references mentioned in the Reference List have been cited in the text, and vice versa as required by the abovementioned rules and conduct.

Abdulmunem Yaseen Khudhaır KHUDHAIR

Signature

DEDICATION

To the owner of a wonderful biography and an enlightened thought, the one who played the biggest role in my education, my wonderful father may God have mercy on him.

To the one who set me on the path of life and taught me the values of it, my lovely mother.
To my life's partner, the one who stood by my side and was there for me, my glamorous wife.

To my brilliant teachers, those who helped me when I needed it.

I'd like to dedicate my project to these mentioned above.

PREFACE

I would like to thank my supervisor Prof, Dr. Osman Nuri Uçan and my all family.

ABSTRACT

NEW SMART GRID FRAMEWORK BASED OPTIMIZED AI METHOD: CASE STUDY ATTACK DETECTION

Khudhaır, Abdulmunem

M.Sc., Electrical and Computer Engineering, Altınbaş University,

Supervisor: Osman Nuri UÇAN

Date: 20/05/2022

Pages: 68

A smart grid is an electrical network based on digital technology that is used to deliver electricity to consumers through two-way digital communications. The system provides supply chain monitoring, analysis, control and reporting to increase efficiency, reduce energy consumption and costs, and ensure maximum supply chain transparency and energy reliability. The smart grid was introduced to fill the gaps in the existing power grids with smart electricity meters.

In this paper, we propose the use of an automatic TSA-based deep encoder to solve the IDS problem in the Internet of Things. It is the first automatic encoder to be used to reduce the size of input data. In the second step, the output of the first autoencoder is connected to the second encoder and the size of the input data is also reduced. Then the output of the second autoencoder is connected to SoftMax. SoftMax categorizes the attributes into three attacks (normal and abnormal). The proposed system showed results with an accuracy of 98% compared to some studies.

Keywords: IDS, Data Mining, Computer Security, Smart Grid.

TABLE OF CONTENTS

Pages

LIST OF TABLES

Pages

LIST OF FIGURES

Pages

ABBREVIATIONS

SVM	:	Support vector machine
NN	:	Neural Network
RBF	:	Radial Basis Function
ESD	:	Energy Spectral Density
IoTs	:	Internet of Things
IDS	:	Intrusion Detection Systems
TSA	:	Tree seed algorithm
ANN	:	Artificial-Neural-Network

1. INTRODUCTION

Unlike other products such as electricity like oil and food, electricity cannot be stored for the future. To be given to the consumer immediately after manufacture. In short, the power grid is a network of power plants, which are the power lines that supply electricity to consumers. Rapid population growth and the proliferation of mining companies make it difficult to meet the electricity needs of households and industry. An increase in the demand for electricity at certain times of the day can lead to various problems, including short circuits and transformer failures. To solve the transmission problems of traditional networks, it is necessary to anticipate consumer behavior to ensure efficient transmission. In this context, the concept of Smart Grid (SG) was introduced. SGs can intelligently forecast electricity demand and supply power based on forecast demand. SG, B. Aggregation and Intelligent Forecasting Demand forecasting can solve a variety of traditional network problems. B. Reduce energy consumption by reducing the risk of short circuits and preventing energy loss, personal injury and property damage [1], 5th Generation (5G), other wireless networks, large data analytics and deep learning (DL) have unlocked the true potential of SG . SG has multiple stakeholders including smart cars, smart buildings, smart power plants and smart cities.

Developments that enable smart cities and focus on sustainability are also driving the growth of renewable energy. But just as adapting smart cities is a challenge, so is adapting renewable energy. Before the advent of renewable energies, networks consisted of a small number of producers, a network, and a large number of consumers with well-defined roles. There was a link between these two groups where the current only flowed in one direction. However, with the development of electricity production methods that can be installed in homes such as solar energy, consumers have become both consumers and producers, that is, "producing consumers". You can not only use your own power, but also supply it to the grid. This change requires a unidirectional transmission line to be bidirectional. This greatly complicates the generation, consumption and flow of energy [2].

Also, in the old network, a high inertia element working as a generator was replaced with a low inertia element such as a solar power plant. Therefore, more complex systems should be more reliable than before [3].

Searching for ways to address the requirements of new scenarios has long been the subject of academic research. Particular attention is paid to the sustainability of SG, which is one of the applications of the smart city concept [4].

The smart grid collects information about consumer demand, compares it with current supply, charges for electricity based on this balance between supply and demand, and sends it so users can decide whether to use it. This process is time sensitive, so it is important to find the SG stability dynamically. In addition to the technical part of the system, other economic effects on energy prices need to be considered in order to analyze fluctuations in energy supply and demand.

Smart grids try to solve this problem by using frequency, which is one of the main features of networks. During periods of over generation, the grid frequency increases and during periods of underproduction, the grid frequency decreases. Basically, this means that all the information required for dynamic energy pricing can be found at the grid frequency [5]. Measuring the frequency of each user provides the network administrator with all the necessary information. So the network administrator can update the energy forecast and inform the user.

A mathematical model based on smart grid differential equations. This is called a study [6] and evaluated as a study [7]. The goal is to determine the jitter of a reference star architecture network with 4 nodes consisting of a power source (intermediate generation node) that powers 3 consumer nodes. Model: Total energy balance (nominal energy produced or consumed by each node in the network), response time inputs, and inputs related to energy price elasticity, which adjust consumption and production in response to price changes. Consider the gender of the participant. This leaves a mathematical model predicting the stability of the network. If so, it provides a robustness score and solves the binary classification problem, but this model relies on significant simplifications.

Through low-cost calculating, the cloud, big data, analytics, and mobile phone knowledges, physical things can part and gather data with minimal human interference. In this extremely related world, digital schemes can greatest, screen and switch every communication among connected entities. The physical world encounters and cooperates in the digital world [8].

There are numerous limits to investigating the transferred data. Regular home-grown internet operator does not income Time, they do not have the practical knowledge to screen

their network movement, and smooth so, if the data Encrypted, it will not be intelligent to recited the brought data. Normal house you can use additional than ten home smart devices that contact the Internet at one time and this is a growing number [9].

Companies are fair start to learn and take benefit of the chances if by the IoTs. The capability to imprisonment and analyze information from dispersed linked devices delivers the aptitude to enhance processes, make new income watercourses, and recover client service. Though, the IoTs also disclosures governments to new security hovels presented by augmented network connectivity and leaky plans by project. Progressive assailants have established the capability to pivot in other systems by captivating benefit of faintness in IoT devices.

In this context, DOS attack happens when the package would Work is usually not available. There can be many Factors for lack of availability, but it generally indicates Infrastructure that cannot be handled due to capacity Overload. Distributed Deprivation of Service (DDoS) attack, a large number of spiteful systems attacking one target. This is often done by robots since many devices are programmed to enter service at the same time.

Developments in computer and internet technologies cause security vulnerabilities of companies and governments. In recent years, social, mobile, big data, internet of things etc. While information and communication technologies have made great progress in fields, network principles have almost never changed. It has become very difficult to meet the current market requirements with the traditional network structure. Traditional

It is seen that the network architecture has remained unchanged in recent years and is quite cumbersome [10].

The traditional network causes the limitation of hardware and software parts. This causes high costs in purchasing new hardware and purchasing new licenses when additional cases are needed. After all, it is not useful for new business situations. Internet technologies for mobile devices, cloud technologies, social networks, big data, multimedia, and digital world make management and configuration very complex and demanding.

In addition, access to high performance, scalability and dynamic management is essential today. the Internet and mobile networks mature and increased availability and dynamic manageability and the need for higher bandwidth networks have become critical issues.

A framework called SDN has been proposed to solve the problems and limitations of the existing network. This architecture separates network management from routing mechanisms, allowing direct programming and control

ICTs complicate and manage existing networks due to their high bandwidth, availability, high-speed connectivity, and dynamic management requirements. To avoid these problems, governments and businesses are starting to use software-defined networking (SDN). SDN allows you to manage multiple devices by physically separating the network control plane from the data and control planes. Secure SDN has replaced the traditional network as it can realize a secure and reliable network system. SDN provides the ability to centrally monitor, modify and manage your network from a single console. SDN is a computer system that separates the control plane from the data plane, leading to a new understanding of networking that leads to traditional network architectures. SDN can be enhanced during installation and later as needed

The IoTs is an innovative uprising for the Internet. it resolves Make themselves recognized and gets smart Make or allow setting decisions thanks to the fact they can connect data about themselves. They data composed by others can be edited Things, or mechanisms can be complex facilities. This is the change escorts the arrival of a cloud Computing and Internet broadcast abilities Toward IPv6 with nearly limitless processing ability. The objective of the IoTs is to allow things Connected anywhere, anytime, with everything and perfect person applying any network / route and any facility. See Figure 1.1

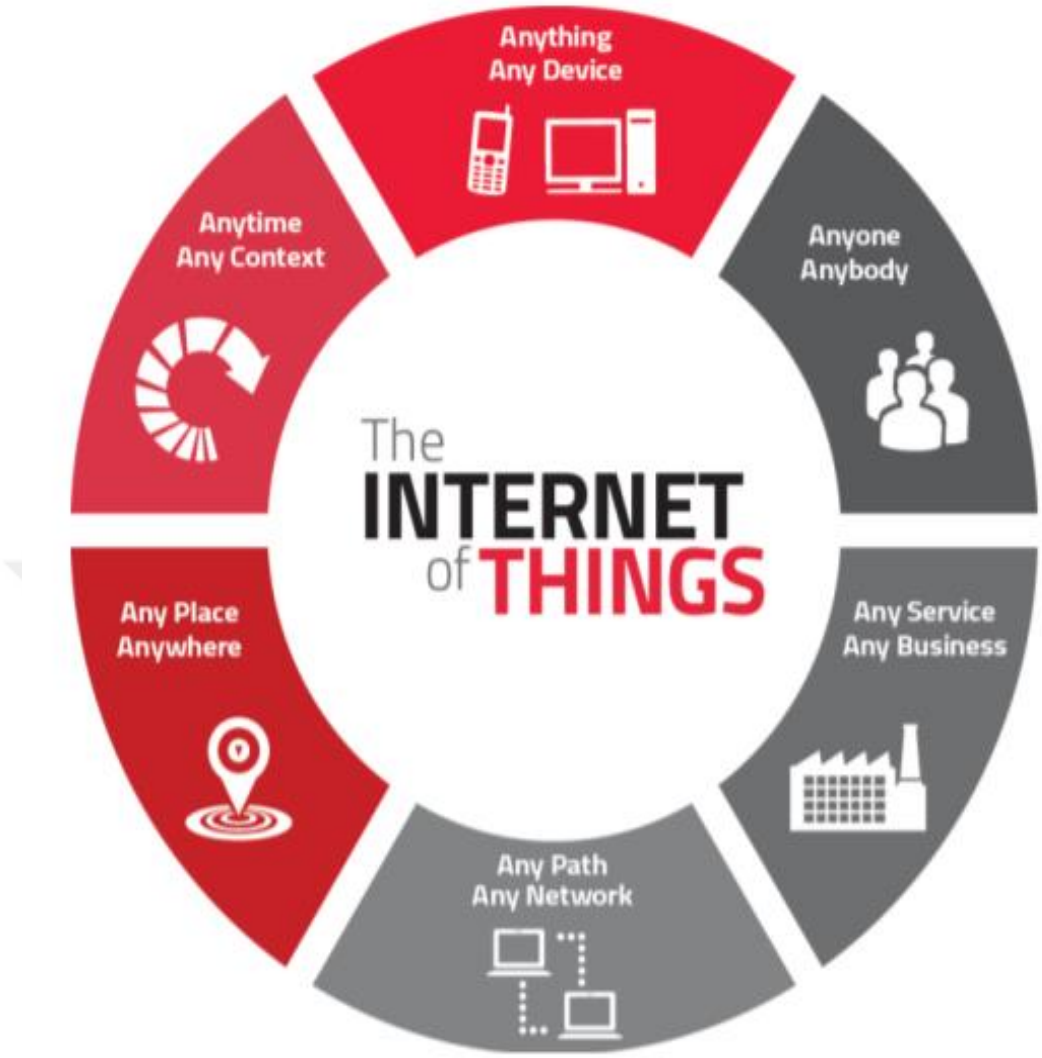

Figure 1.1: Internet of Things [11].

Intrusion Detection Systems (IDS) are essentially system-level security mechanisms for the Internet of Things. The published identifier must be applied to the IoT system.

Identities, designed for intelligent IoT-based environments, must operate in complex environments such as low processing power, fast response times, and big data throughput. Therefore, intrusion detection systems may not be very common.[12].

The most obvious difference between deep learning and traditional networks is that deep learning is software-based while traditional networks are usually hardware-based. Since deep learning is software-based, it provides more flexibility, control, and convenience for users to manage their resources. Large organizations such as Google, Yahoo, HP, and

Cisco are investing heavily in protecting their systems from distractions. Existing networks require new hardware to increase network capacity. This requires companies to use deep learning technology to reduce costs. At the same time, deep learning provides high performance for modern applications [13].

Tech researchers and developers around the world still have to agree on a single standard architecture for IoT systems. Three levels make up the IoT infrastructure [14]. Other researchers [15], [16] suggest structures of four or five levels. Of course, there are security and privacy issues as there is no standard architecture. In practice, the smart environment consists of different IoT systems such as sensors from different technology companies and hardware or software applications that do not use a common standard language [17].

Solving this problem requires effective discovery as well as the development and implementation of reliable and realistic security measures, including network attack detection, malware detection, and network discovery [18]. Therefore, we need a stable and powerful AI algorithm that can detect and respond to network attacks effectively. The development of these algorithms requires a representative and well-organized dataset to train and validate the reliability of the system [19]. Much research has been done in this area, but the vastness of cyberspace and the random and unexpected attacks in cyberspace have made the Internet of Things a subject that needs more research and development. Additionally, there are significant gaps in research, and there are gaps in methods, practices, and data sets. In order to fill these gaps, it is very important to conduct a detailed study of artificial intelligence algorithms, real datasets and anomaly detection [20]. Therefore, knowledge extraction, subtle strategies, old fashioned brain tissue, genetic computation, neural genetic computation, local molecular knowledge, overload, understanding learning, discovery systems of heavy traditional knowledge in light of traditional AI calculations are an IoT organization [21].

Depending on the type of algorithm, each machine model has different dataset properties. For example, the attributes of a linear trend have a large impact.

A large set of astronomical data. This allows you to analyze and select them and predict which characteristics of the dataset (eg columns) are most important.

Network anomalies are malicious traffic that compromises the security of your network. New types of malicious attacks and data breaches appear every day. The detection of anomalies is particularly important to prevent them. In its simplest form, anomaly

detection is a method of detecting abnormal locations or conditions in a given population. These contingencies have structures or patterns that don't really match the expected behavior of the data. Anomaly detection aims to detect normal behavior caused by defective organs. Therefore, they are very important in computer applications. The biggest challenge in anomaly detection is managing large datasets.

Network error detection is presented as an unusual situation in network traffic. Credit card fraud, inaccessibility of internet services and transfer of sensitive information from one computer to another are just some of the anomalies [22].

Network Intrusion Detection Systems (NIDS) are essential to protect against malicious and suspicious cybersecurity activities. NIDS plays an important role in protecting against threats and attacks on network systems. NIDS faces many challenges and challenges in detecting anomalies in network traffic. As the types of attacks continue to increase and traditional methods are used, a large number of false positives are occurring.

Analyzing traffic behavior through monitoring, preventing the collection of important information, and preventing unauthorized access to computer systems are some of the main activities of malware.

What sets this study apart from other studies is the detection of anomalies in smart grids and IoT data using real-world data. The dataset used consisted of data retrieved from a realistic network environment.

1.1 OBJECTIVE OF STUDY

This article attempts to develop a new framework as an optimized data mining technology for automatic detection of IDS in smart grids. The proposed structure is validated against a common dataset in the are and the outcome are compared to the better method obtained in the previous researches.

1.2 MOTIVATION

To date, data mining studies in previous study areas have given satisfactory results. This includes image recognition, disease detection, and data classification. Data mining technology optimized to detect IDS attacks on the smart grid. Data mining methods have

low precision for complex tasks. However, through optimized data mining, it was possible to improve the accuracy of the technology.

1.3 RESEARCH QUESTIONS

A. Which dataset represented IDS attack very well in smart grids?

B. How can we improve the accuracy of more precisely defined functions and classifiers?

C. How optimize data mining technique in best form?

1.4 LIMITATION OF STUDY

The first weakness is that when looking at a new topic like data mining, the time it takes to complete a thesis is short, so there are no other research in this area. In addition, other methods involve different computer hardware such as CUDA, but it is expensive and scarce. This problem can also be solved by using CUDA in a function as it can work very well for such tasks.

2. OVERVIEW

2.1 RELATED WORK

[23] Heriberto Valentia. Propose a threat model to identify potential weaknesses in low voltage distribution networks. Then calculate the probability of exploitation based on a realistic attack scenario. Finally, apply a formal check to check the probability model against the attack characteristics. The results obtained provide information about potential threats and possible successful attacks. Explain the impact of security breaches on the security and privacy of energy consumers. Next, we will discuss future considerations for improving the security of critical energy sectors [24].

Silva et al [25] Announcement of centralized, decentralized, and specification-based intrusion detection mechanisms. In this mechanism, a node called a watch node finds a node that violates the rules. Each node has an error counter. When a node violates a rule, the counter is incremented. This node generates an alert when the counter value exceeds a certain limit "th" during the period "t". The author uses a C++ simulator that he developed to test his work. The results show that the proposed method is energetically efficient and 100% of the final attacks detected by black holes, selective redirection and wormholes.

The Romans and others [26] introduced the so-called "weather watcher" method of exploring the region. As a result, an Intrusion Detection System (STS) agent is installed on each sensor node to monitor data from the nodes and neighboring nodes within wireless range. An alert is generated when a node behaves abnormally.

Lu et al [27] proposed a fixed-width clustering algorithm. According to the algorithm, an STS agent is installed on each node and each node acts as a monitoring node. Then B. A total of 12 characteristics are defined, including the number of packets received, sent, and combined. These properties are used in general messages to determine the means or standard deviation of each neighbor. These values are calculated and summarized by simulating various attack scenarios. After analyzing these groups, we found damaged nodes. The disadvantage of this model is the definition of characteristics used to identify specific attacks. According to the authors, the proposed methodology allows to effectively detect simulated redirect attacks with low levels of false positives.

Drozda and Schherbeka [28] demonstrate a detection mechanism based on an Artificial Immune System (AIS). This mechanism does not require many computations and provides better detection performance. This method first checks for normal behavior and then analyzes the message for attacks. The authors describe the principles for developing the proposed method and demonstrate the effectiveness of the approach for performing simulation experiments in Network Simulator 2 (NS-2).

Gupta et al [29] proposed a basic anomaly detection mechanism (ANDES) for intrusion detection. ANDES creates a route to each node and checks for its existence. ANDES detects data and data anomalies collected by the control aircraft.

Kroniris and his colleagues [30] proposed a common specification-based local control mechanism for selective targeting and detection of black hole attacks. Following his approach, an STS agent is installed on each sensor node and ensures that the node's messages conform to the rules. When specifications are violated, a warning is sent to the link discovery engine. This component then communicates with other nodes to determine the status of these nodes. On most nodes, an alert is generated when a node is detected to be down.

Bogkowice et al [31] Let's take a quick look at some of the attacks that break the overall working of the WSN. According to them, "STS is a deprecated service that helps in scenarios where a node may be attacked and captured by an adversary." .studied the mechanism of abnormality detection. His work focuses only on unusual talents.

jerks, etc. [32] principles an algorithm to detect compromised nodes. Faulty nodes can produce false positives compared to healthy nodes, so we developed a jointly distributed STS to detect these nodes. The verification is carried out in two stages. First report the damaged node. If no information is available, randomly select n neighbors based on threat level and send a handshake request packet. The decision phase begins when a node receives an acknowledgment request packet. because depending on the response of the randomly selected nodes, there are three stable systems. OK (node is not healthy), disagree (not sure), ignore (node shipping alert is compromised). The proposed model not only helps to detect damaged nodes.

Ahmed et al. [33] proposed another abnormal method for detecting lymph nodes. They use signature and anomaly based methods to identify bad nodes. Using this method, the sensor network is split into pairs. Each sensitive node controls the behavior of that node. This approach has two problems: coupling and abnormal activity detection mechanisms.

Li et al. [34] Implements a group search mechanism. This mechanism divides the network of sensors into n groups. Starting from the fact that all the nodes of a given group must do the same thing, the author provides information on one or more properties of the environment. However, the information perceived must differ from the specific threshold. First, the sensor nodes are grouped based on the similarity of the collected data. This information can be used to detect abnormal activity on a specific host in an attack scenario. If a particular node fails, it is removed from the routing table. The author reports that he found a low level of false positives after applying the proposed methodology.

Chang et al. [35] presents a graph-based approach for efficient detection of damaged signals. The IDS agent must be installed on the beacon host. The address node provides information about the location of the sensor node. A corrupt Beacon node sends incorrect information to other nodes, which affects the performance of routing protocols. The base station receives these alerts using the secure transmission protocol. After collecting the generated dataset, apply the proposed graph theory-based detection mechanism to ensure that the information comes from a trusted source. The authors of this approach propose a practice-oriented structure. Its purpose is to identify reliable or dangerous sources of information. It is expensive to install the Global Positioning System (GPS) on all sensor nodes. The beacon node concept makes efficient use of network resources using location-based routing.

The IDS mechanism proposed in Tiwari et al. [36] follows a specification-based approach that defines rules that define behavior as normal or abnormal. The number of messages a node leaves behind is used to detect black holes and further redirection attacks. The cluster leader is responsible for determining the legitimacy of the node where the cluster is sending and receiving packets.

Fried et al. [37] used Bayesian classifiers and decision trees for data mining techniques to detect various network attacks. The proposed algorithm also solves some of the following

data mining problems: B. Feature processing, handling missing feature values, and noise reduction of training data.

Mamoun et al [38] proposed a new IDS by comparing different IDS approaches. With this architecture, the total surface of the sensor network is divided into several zones. Sensor nodes in each zone are controlled by cluster nodes. Two or more cluster nodes are controlled by a local node. Instead, local nodes are monitored and controlled by base stations.

Chitrakar and Huang [39] proposed a hybrid technique for detecting abnormal behavior. The authors attempted to apply a hybrid learning approach by combining the k-medoid-based clustering method with the simple Bayesian classification method.

Copolino et al. [40] propose an improved IDS model for WSN using data mining techniques. In the proposed IDS, two types of agents were categorized as central agents and local agents. The central proxy is the base station that makes the final decision on intrusion detection, which is on the server, and the local proxy is at the node.

Sajjad et al. [41] propose a method for intruder detection based on neighbor trust calculation. In the proposed IDS, each node monitors the trust level of its neighbors. Based on this confidence score; Neighbors can declare themselves to be trustworthy, dangerous, or harmful. The proposed scheme successfully detects Hello Flood attacks, interference, and selective routing by analyzing network statistics and the behavior of malicious hosts. The simulation results show that the network performance is better when anomaly detection method based on neighbor confidence is applied.

Balakrishnan and Reno [42] attempted to detect and detect intrusions using a rule-based anomaly detection system. Anomalies are detected under specific rules. The survey identifies the heads of different groups from different nodes in the network and collects the remaining nodes. The cluster leader receives intrusion detection data from the cluster members. The cluster leader also forms the basis of the routing architecture, so that attacks against other cluster leaders can be detected. But behind the party tree stand the party members. If a block header is detected to be defective, the block header monitor shuts it down. After the malicious block header is identified and removed, it retrieves another

block header from the block member. The performance of the algorithm was evaluated by the authors using a simulation tool specially developed in C++.

Momani et al. [43] developed a proprietary WSN dataset to detect and classify four types of denial of service attacks (black hole, gray hole, flood, and programming). The dataset collected is referred to as the Wireless Sensor Network Dataset (WSN-DS). Artificial Neural Networks (ANNs) are trained on data sets to detect and classify various Denial of Service (DoS) attacks. The results showed that WSN-DS improves the ability of STS to achieve higher classification accuracy

2.2 SMART CITIES

Cities are constantly in design and have complex dynamics. There is an event full of events happening around the world. Those who will live on an Earth now will be at 70% of that rate by 2050 [44]. It can be pure system infrastructure or the same plan information and social infrastructure for cities to progress. For this, the concept of “Smart cities” put forward brings all modern production factors under a common roof. For this purpose, research is carried out in the literature on smart cities and applications .

In the study of Caragiulo et al. [45], it is presented that the main focus of smart cities is ICT. The European Union (EU) and other international institutions draw attention to the importance of ICT-based city development. However, the quality of ICT is not the only definition of the smart city; The role of human capital and education in urban development is also important.

carries. Cities with a high skilled labor force in the past are able to attract more skilled workers. For EU policies that place emphasis on spatial homogeneity, this poses a problem. The study examines the characteristics of the smart city; the use of a networked infrastructure to strengthen social/cultural/urban life with economic efficiency

We define smart cities as “government-run cities where natural resources encourage smart governance”. For the wisdom of these cities, consider e-government, gross domestic product per capita (GDP-PPP), employment rate in the entertainment industry, multimodal access, time spent in public transport common and human capital derived from purchasing power. Parity. There is a strong relationship between multimodal access, transit time,

accessibility, and GDP. The long-term impact of the entertainment industry on urban development is a subject of controversy in the literature, but ideas for solutions for the relationship between work and the creative classes (people working in the professional community, people specialized in combinations of new ideas and new knowledge). Percentage of people involved in the problem or value creation) and creative activities and GDP per capita. finds a slight correlation. There is also a positive correlation between e-government and GDP per capita, but the study also points out that rich cities may place more emphasis on e-government. According to neoclassical theories, human capital is a strong predictor for economic performance, but the study shows a complex relationship (a quadratic curve) between human capital and GDP per capita.

In 2008, IBM chief executive officer (CEO) Samuel Palmisano gave a speech to the Council on Foreign Relations, promoting the concept of the 'Smart Planet'. In 2009, IBM CEO received positive feedback from Obama on the same speech. In the same year, IBM's CEO in China introduced the 'smart planet' concept as the final strategy. The then Chinese prime minister, Wen Jiabao, opened the IoT institute in Wuxi.

During his visit, he emphasized the importance of the development of the sensor network. According to the study of Kehua Su et al; IBM's smart planet concept aims to combine the information technology of the future with structures such as networks, transportation systems, hospitals, and to provide the combination of IoT and ICT.

Thus, the smart city is the combination of digital city and IoT. For the smart city; installation of application systems, which is especially important among the installation of collective infrastructure, platform installation and installation of application systems; the establishment of wireless cities, the establishment of smart homes by combining sensor devices with IoT, the establishment of smart transportation systems, the establishment of smart social services and management, the establishment of network control mechanisms through systems such as e-government with the help of wireless technologies for city management, the establishment of smart health services, the establishment of tourism system summarizes it as smartening [46].

In the study of Oliveira and Campo largo [47], the rise of digital society brought; The tension between social cohesion and building sustainable economies is underlined. It is

argued that cities will become 'Human Smart Cities' when they take full advantage of human capital. New management models that come with innovative projects such as Peripheria, CitySDK and My Neighborhood enable city residents to contribute more to city service management. The study shows that the concept of smart city was created by the ICT industry and the industry failed to engage citizens in the first place and failed to take ownership of the authorities.

argues that it continues. They aim to offer models for the joint participation of society and authorities and want to integrate the paradigm they have presented with the name Human Smart City into cities. In the study, the My Neighborhood project implemented in Lisbon, Milan, Aalborg, and Birmingham is given as an example; project in the neighborhood aims to recreate and strengthen social ties. ICT is used to retrieve user data at the local level; It is aimed that the cooperative societies bring innovation from the bottom up. Solution: It is based on three phases: rebuilding neighborhoods/communities, empowering them and scaling their values. Big data has been used by governments to support the sustainability and development of smart cities. What drives cities to this is to gain the economic, environmental, and social benefits that smart cities bring. According to the study of Nuaimi et al. [48], some advantages of smart cities are efficient use of resources, better services, more efficient business models, better quality of life with less waste, and sharing data and resources. As for the problems encountered during the use of big data in smart cities; Most of the big data obtained is unstructured, ensuring data sharing between different parts of the city and providing a fine line between collecting enough data and protecting the privacy rights of citizens, low quality of data from many sources, and since some of the collected data will be confidential, this data It is given as protection against unauthorized use and attacks, to be able to fight problems such as traffic that will be brought by the increase in the size of the city.

2.3 INTERNET OF THINGS (IOTS)

IoT, smartphones connected systems and schemes influence data calm by embedded devices and actuators on machineries and additional physical sensors complaints. IoT is predictable to feast fast in the pending years, and this meeting will disclose a new element of facility that upsurges customers' excellence of life and the output of businesses, GSMA "Related Life".

There are several technologies complicated in applying IoT. Simple structure of IoT represented below in Figure 2.1 and some of them represented below:

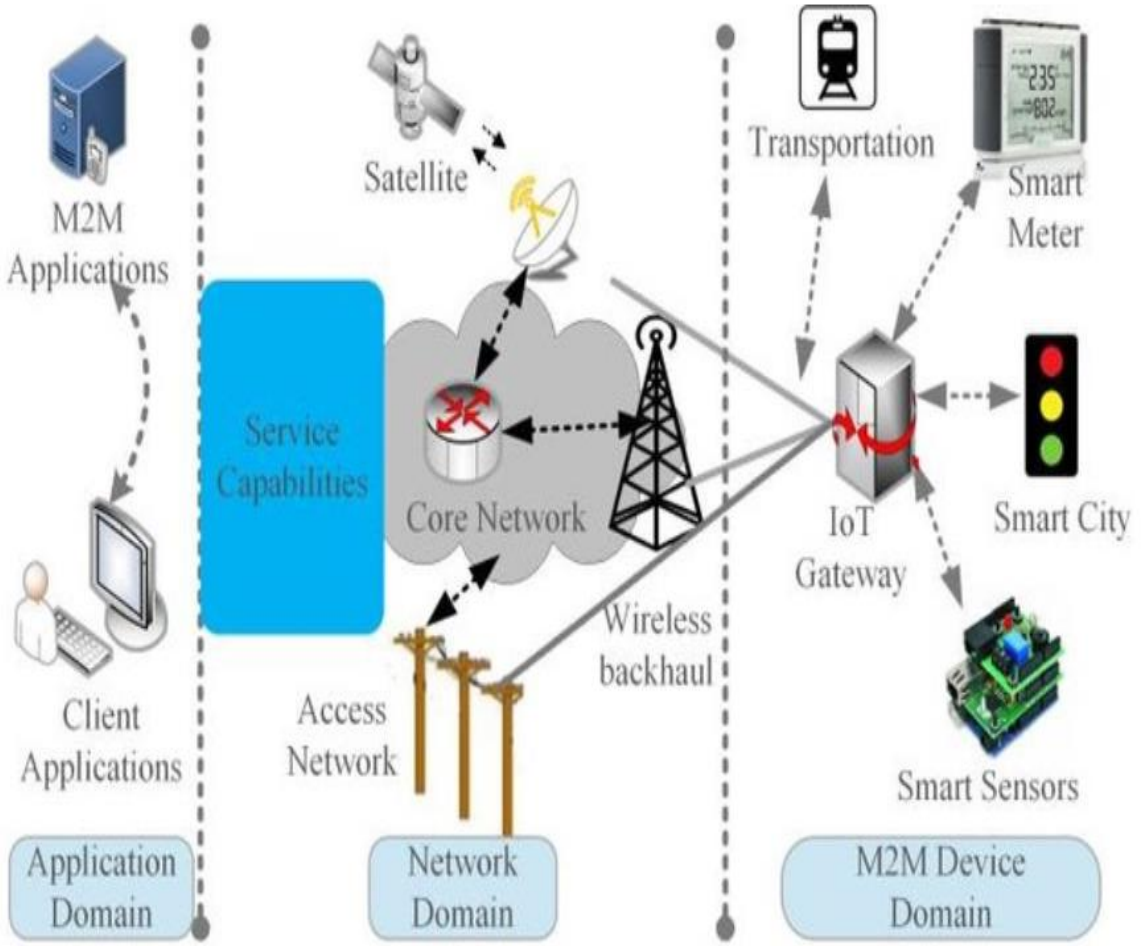

Figure 2.1: IoT Architecture [49].

2.3.1 Six Layered Architecture

As shown in the in figure 2.2. A six-tier organization depend on the network hierarchical organization:

Figure 2.2: Six (6) layer IoT organization [50].

i. Coding Layer: The coding layer is the foundation for IoT, which permits the identification of items of attention. In this layer, individually item is allocated an exclusive ID that types it informal to classify items.

ii. Detection Layer: Also identified as a sensor layer or physical layer. It represents as the five organs of IoT. Recognizes items to gather data. For this determination, Radio Frequency Identification (RFID) tags, barcodes, Bluetooth, wireless sensors, LTE and so on. Dissimilar kinds of sensors are further to the items. A sensor is determined conferring to the needs of the operators and the items to which they are joined. It sends the composed data to the spectator layer to check the validation of these sensors and equipment.

iii. Network Layer: It achieves the subsequent purposes; gateway - routing and dispensation - network competences - conveyance competences - Error recognition and alteration. It also receipts maintenance of communication advancing, publishing and contribution. With obligatory request for High-speed dispensation facilities, context-sensitive presentations, and so on. Deliver a wide variety of IoT facilities and presentations, such as several networks with dissimilar knowledges and admission etiquettes are obligatory to work heterogeneously with apiece extra Preparation. These networks can be tradition, general, or mixture models and are intended to be reinforced. Joining supplies for dormancy, bandwidth, or safety.

iv. Middleware Layer: Folds data sent from a transport layer. Procedures the composed data. It is grateful to remove any extra data that types no sense and to excerpt beneficial data. Though, it also removes the big information problem in IoT. Big data covers a lot of data that can move the enactment of IoT.

v. Application layer: Covers the IoT presentation. This layer is at the upper of organization and Accountable for providing several applications to several operators in IoT. Applications can be dissimilar areas like logistics, industrial, trade, public security, environment, and medicine etc. Many applications with augmented adulthood of RFID knowledge. They are emerging below the top of IoT.

Business layer: The business layer denotes to the envisioned behavior of an request and performances as the manager of the whole scheme. IoT is answerable for handling and regulatory its requests, business and income replicas. The user's confidentiality is also achieved by this layer. It too has the aptitude to control how data can be shaped, stowed, and adapted. The weakness in this layer permits assailants to misuse an request and evade

business reason. Greatest security subjects are weaknesses produced by tainted or lost security payments in a submission [50].

Any IoT execution needs considerable quantity of skills to be able to provision the obtainable facilities in footings of software, hardware, and network as well. There are numerous experiments essential to be studied when we are talking about security of an IoT situation.

a) The main issue in IoT world is the items themselves, also physical or effective. By countryside, IoT is dependent on the growth of items as greatly as likely and the extra items we have the extra possible difficulties may have. Tens of ages back we had only to defend our PC to admission the internet, now we need to upkeep about PCs, Smartphones, smart plans, car, wearables, whatever almost that is coupled completed the net [51].
b) Communication and data transfer. Irrespective of the amount of IoT equipment, the data existence taken must someway be advanced to the following equal for additional dispensation. Verification and discretion define must be a key component of IoT security when encryption wants to be applied
c) Unofficial admission to smart equipment.

2.3.2 Future of IoTs

Rendering to [52] by 2020 there will be 20 billion associated IoT equipment. Though, the amount of linked IoT equipment in 2020 may be level upper owed to the creations of novel IoT equipment. Before, the development tendency was as tall as 31% year-on-year. In

2017, the customer section now accounts for 63% of the total amount of IoT equipment. Besides, the key issue will be on the customer lateral and low-cost equipment play an significant part in capacities.

When talking about the future of IoT, marketplace rate also dramas a significant character. Though customer's acquisition a large amount of IoT equipment, the business lateral is capitalizing additional in IoT equipment. In 2017, 57% of the spending on the IoT marketplace was enclosed by the business. The development trend comprises IoT facilities; it is immobile a minor but fast rising subdivision. Figure 2.3 below demonstrations the marketplace values for the IoTs market projected among 2016 and 2020 [53].

Category	2016	2017	2018	2020
Consumer	532,515	725,696	985,348	1,494,466
Business: Cross-Industry	212,069	280,059	372,989	567,659
Business: Vertical-Specific	634,921	683,817	736,543	863,662
Grand Total	**1,379,505**	**1,689,572**	**2,094,881**	**2,925,787**

Figure 2.3: IoT market value, 2016 – 2020 [53].

2.3.3 Advantages of IoTs

Effective use of resources: If we distinguish the function and the method in which every expedient works, then we certainly upsurge the effective use of resources in addition to monitoring natural resources.

a) Reducing human effort: Since IoT devices interrelate and connect with all additional and do many tasks for us, they reduce social exertion.
b) Save time: Since it decreases human exertion, it certainly protects period. Time is the chief issue that can be saved through the Internet of Things platform.
c) Improving data collection:

d) Improving security: Now, if we have a system that is interconnected of all of these things, we can make the system safer and more efficient.

2.4 DETECTION METHODS

There are 2 chief approaches of recognition for IDS: “misuse-based detection” & “anomaly-based detection”. Individually technique has its benefits and disadvantages and that is the aim why a new group has been deliberate intensively in the last insufficient years entitled hybrid finding. The main determination of hybrid recognition is to enhance the recognition precision by merging the powers of the two chief approaches before declared: the correctness of the misuse-based recognition on beforehand recognized attacks and the simplification rule of the anomaly-based recognition to beforehand hidden attacks [55].

2.4.1 Misuse-Based Detection

Misuse-based recognition, typically, an IDS procedures log archives established after the OS for an exact era of period in instruction to have a whole usual of operator action. Afterward that, the IDS does examination of the present action, using an instruction base system, indicators, or a consistent empirical, in instruction to decide the conceivable incidence of irregularity or interruption. For the current, the misuse appliance usages a predefined matrix of interruption forms so the scheme distinguishes in loan the arrival of misuse or misuse. The better this bench, the additional interplanetary and calculation period is consumed in the examination. Preceding exertion by a GA castoff a bench of scope 28x48, wherever 28 is the total of events to be measured—an action might be an operator logging, interpretation of an exact folder, performing a package, fourth and 48 is the amount of intrusion forms. Current works uses hundreds and extra than one thousand interruptions, authorizes earlier effort, and proposes an iterative scheme as additional alternate. An intrusion container be stated by a collection of actions to plaid, wherever respectively admission stipulates the numeral of actions of an exact kind that must happen in instruction toward take an intrusion. Similarly, the penalties of the operator annals calm can be realized as a collection, anywhere apiece admission stipulates the entire amount of actions of that exact kind achieved by an operator. Figure 2.4 illustrations an instance of arrays P of intrusion forms and an array UA of operator action [56,57].

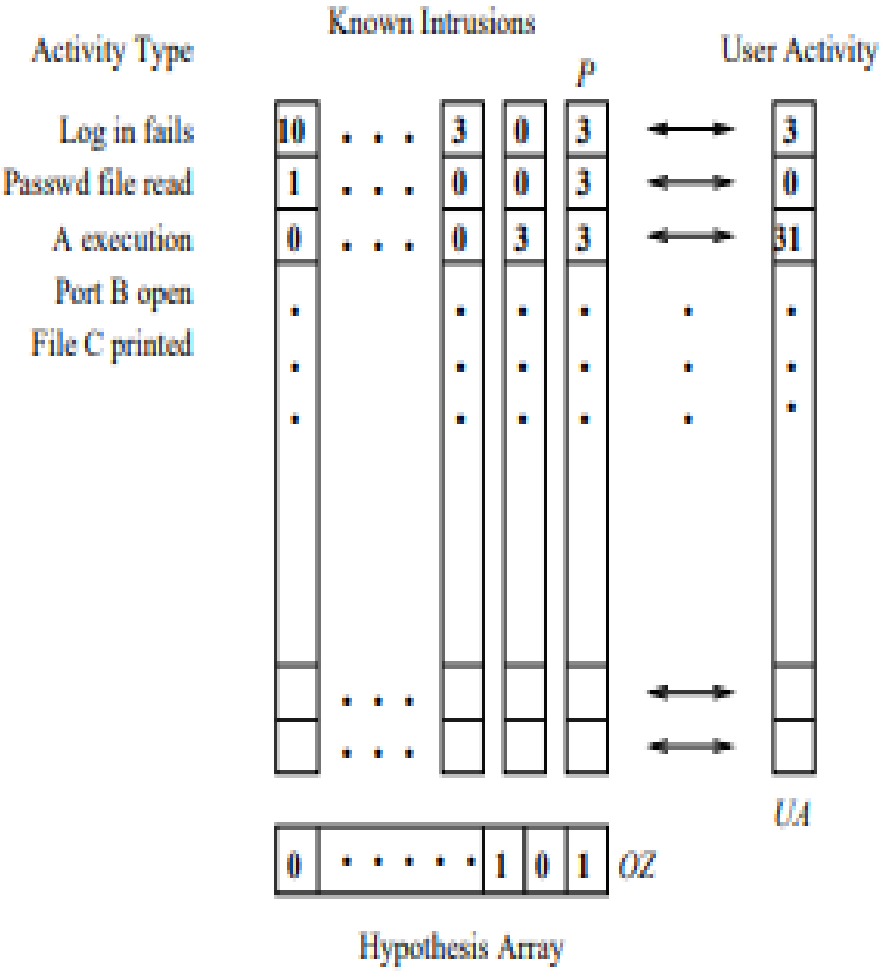

Figure 2.4: Arrays of intrusions P, array of user activity UA, and hypothesis array OZ [58].

2.4.2 Anomaly-Based Detection

IDS are security methods that have ability gather and examine data from numerous kinds of method and net bases to identify action that could be an attack or unlawful admission on the method. IDS mention to recognition of hateful actions like break-ins, diffusions, and additional procedures of machine misuse in a host or net. Interference is altered after the normal conduct of the model, and henceforth irregularity recognition methods are appropriate in intrusion recognition area. Contingent on the info foundation careful, an IDS

may be also host or network-based. A host-based IDS examines proceedings like procedure identifiers and system appeals It is associated with operating system data [59].

Interrupts take the form of incorrect subsequences of sentences. Irregular sequences interpret hate programs, errors and rule violations. Capturing a continuous scene of information requires a violation detection technique that is useful for detecting host-dependent failures. Each method requires changing the order information or calculating the similarity of orders. There are two ways to identify anomalies using system call suggestions. The order of short commands depends on their frequency. Instead, the network-based IDS monitors network-related activities. B. Intrusions from network IDS such as traffic bandwidth, IP address, service port, protocol, standard, etc. They are often called irregular and continuous surveillance, and they exhibit normal behavior in EU networks. Information is usually a large dimension with a combination of defined and persistent attributes. Thus, anomaly detection techniques are needed to improve computational efficiency by collecting these critical inputs. In addition, information is often flexible in style and requires online searches. Brand information consistent with normal behavior is usually available, but intrusion tags are not available. Thus, partial, and unattended anomaly detection methods are preferred in this area. The above network connection detection methods used by IDS can be divided into four main categories: grouping, arithmetic, data theory, and aggregation. Due to the high volume of input and active network configuration, Ottum Rul is a problem.

To solve this problem, you can get used to the simplistic behavior and manage your actual networking environment. This entry is recommended for self-learning systems [60]. Internet technologies for mobile devices, cloud technologies, social networks, big data, multimedia, and digital world make management and configuration very complex and demanding. In addition, access to high performance, scalability and dynamic management is essential today. The traditional standard method is very boring. traditional networks: It is a hardware-centric network with significant disadvantages in terms of innovation, reliability, scalability, flexibility, and manageability.

Managing large numbers of network devices in such situations is difficult and error-prone. As a result, Software Defined Networking (SDN) is a platform designed to simplify and improve network management and increase flexibility by separating the control and data layers. Although ITA is seen as an emerging research topic, it has attracted the attention of

many researchers in industry and academia. Academies such as Google, Yahoo, HP and Cisco are starting to use YTA technology.

As the Internet and mobile networks mature and increased availability and dynamic manageability and the need for higher bandwidth networks have become critical issues. .

A framework called SDN has been proposed to solve the problems and limitations of the existing network. This architecture separates network management from routing mechanisms, allowing direct programming and control. SDN uses many simple packet forwarders (SDN switches) and controllers that are logically centralized, and are managed and configured through an open, stream-like interface.

An SDN switch consists of one or more routing tables managed by a central controller. That is, it is controlled and planned in the control plane. This mechanism allows software developers to easily manage network resources. Packets are also processed in the routing table. This means that SDN switches can be used in routers, switches, firewalls, etc., depending on the policies implemented by the central controller in the routing table. means it can be used. You can work the same way. SDN simplifies modern network management and provides opportunities for innovation. This allows researchers to test ideas, explore and evaluate results. SDN attracts researchers as it plays an important role in modern Internet technologies.

Control Plane**:** NOX, POX, Beacon, Maestro, and Floodlight are a few of the SDN controller platforms. POX is an open source platform for SDN control applications. While NOX and POX platforms are based on Python; Beacon, Maestro and Floodlight platforms are based on Java. Since POX, which enables rapid development and prototyping, is used more widely than NOX, the POX platform will be used in our thesis.

Data Plane: Consists of communication devices. It includes instances of data sources, classifiers, meters. This plane is programmed and managed by the controller.

Application Plane: It is shared with applications and services that define the behavior of the network. Applications that directly (or primarily) support routing programming operations (for example, control plane routing operations) are not considered part of application programming.

SDN is a logic network device that exposes data plane layer view. In the SDN control plane, SDN makes decisions such as routing, sending and route selection. Provides

information about the status of devices. SDN applications in the SDN application layer are the programs that provide the connection between the needs of the network and the behavior of the network explicitly and directly.

2.5 ARCHITECTURE

In direction to construct effective IDS, one must reflect round the dissimilar mechanisms of their manner. Usually, IDS take at smallest 2 mechanisms: an information group section and an investigation section [61].

i. The information gathering unit is in custody of gathering smithereens of data castoff as indication of an imposition and bringing it to the break of the system to choose if there is an anomaly. This module gathers review tracks such as employer logs, net tracks, organization calls, etc. The information group unit looks like a net of devices watching the review tracks in real time.

ii. The investigation unit examines efforts conventional by the information group unit and chooses if there is an attack or not. It is the essential of the IDS and can differ contingent on the detection technique. Its drive is to deliver the system manager with a strong swift of the scheme's state. This unit can be actual reserve wearing. That is why it is often joint with additional unit which procedures the information primary to lesser the computational complexity. Additional complex IDSs have extra features like an information preprocessing module, a reply to module and a signature producer module on highest of the two chief modules [61].

iii. The information pre-processing unit is in custody of plummeting the size of the information in command to enhance the computational speed of the IDS. NIDSs have to analyze the traffic in actual time deprived of presenting too much stays in the network. For this purpose, this module is crucial to brand the IDS feasible in a real setting. Additional particulars will be given about noise decrease and feature selection [63].

iv. The reply to module can be also active or energetic. Greatest of the present IDSs are active. This income that they usual a terror when an intrusion receipts home, but do not stab to

pawn the attack. Lively IDSs (intrusion prevention systems) distinguish the attack and effort to stop it. Dissimilar kinds of responses are likely. The IDS can near the present TCP connection, reconfigure mechanically the firewall to avoid additional intrusion, etc.

v. The signature producer module makes new attack signatures rendering to the info see Figure 2.5.

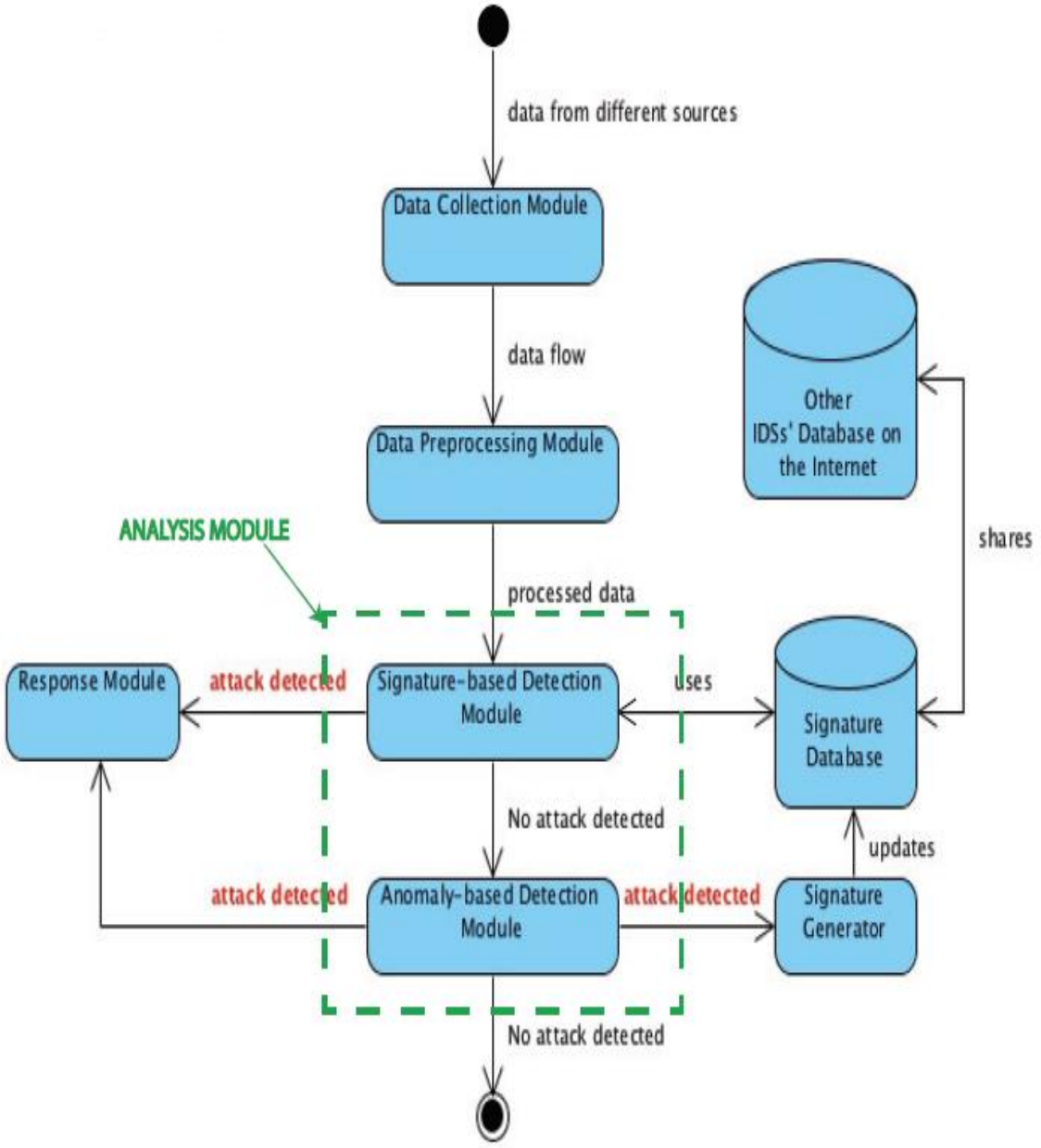

Figure 2.5: Possible IDS architecture [64].

2.6 CHIEF CLASSES OF ATTACKS

A decent catalogue brands it likely to categorize separate attacks into clusters distribution mutual possessions [65]. One extensively used classification division's attacks into 4 lessons: Investigations, “Denial of Service (DoS), User to Root (U2R) and Remote to Local (R2L)”. The lesson Usual that will be castoff advanced in this effort denotes system traffic which is careful attack-free. Numerous main ideas must be unspoken beforehand starting to dig deep into the dissimilar lessons of attacks [66].

2.6.1 Probe

This subdivision presents the manners, status, and kinds of net probes. System probes are actual central kind of net attacks. They are glared as the primary stage to takeoff greatest of net attacks. Some network attacks consist of three steps that typically happen in order: system probe, attacks action, and footmark permission. Deprived of network probes nobody of greatest attacks can be hurled. Network probes are recognized as investigation attacks since their chief impartial is to gather data around hosts and net facilities consecutively in the network. A. Host Curve Attacks Net attacks typically twitch by execution net detection, which comprises host curve and harbor examination attacks. Host bend attack controls the living hosts in the net; the attacker stabs to discover an entry to twitch through or arrive from. Consequently, interlopers will be penetrating for feeble opinions to cooperation. Deprived of this attack, the interloper will not be talented to unveiling greatest of the attacks in contradiction of net capitals. There are three mutual kinds of host stroke scan methods: TCP ECHO, UDP ECHO, and ICMP Sweep [67]. The Probe attacks listed in Table 2.1.

Table 2.1: Probe attacks [68]

Name	Service	Vulnerable Platforms	Mechanism	Time to Implement	Effect
Ipsweep	icmp	All	Abuse of feature	Short	Find active machines
Mscan	many	All	Abuse of feature	Short	Looks for known vulnerabilities
Nmap	many	All	Abuse of feature	Short	Find active ports on a machine
Saint	many	All	Abuse of feature	Short	Looks for known vulnerabilities
Satan	many	All	Abuse of feature	Short	Looks for known vulnerabilities

User to Root (U2R):

The fast growth of computer networks and typically of the Internet has shaped numerous safety injuries. Throughout New Year's, a numeral of attacks on network have melodramatically augmented. So, safeguarding of network capitals are actual vital particularly U2R attacks. A single U2R attack might reason a large damage of a business. Consequently, defensive company capitals are actual grave from these attacks. In these sections, we temporarily present related work, U2R attack and its kinds [69].

User to root attack includes illegal access to native wonderful operator freedoms by a local privileged operator. Instances of such attacks are “Load module, Perl” and “Buffer overflow. Load module”: This attack feat a fault in in what way SUNOS animatedly load units. This fault brands it likely for any operator of the method to grow root freedoms. Perl: Exploits a bug in some PERL applications on approximately previous systems. This germ contains in these PERL applications indecorously treatment their root freedoms. This principal to a state where any operator can get root freedoms. Buffer overflow: Contains in teeming input buffers acceptable to overwrite memory sites covering security pertinent data.

Table 2.2: User to Root attacks [70]

Name	Service	Vulnerable Platforms	Mechanism	Time to Implement	Effect
Eject	Any user session	Solaris	Buffer Overflow	Medium	Root Shell
Ffbconfig	Any user session	Solaris	Buffer Overflow	Medium	Root Shell
Fdformat	Any user session	Solaris	Buffer Overflow	Medium	Root Shell
Loadmodule	Any user session	SunOS	Poor Environment Sanitation	Short	Root Shell
Perl	Any user session	Linux	Poor Environment Sanitation	Short	Root Shell
Ps	Any user session	Solaris	Poor Temp File Management	Short	Root Shell
Xterm	Any user session	Linux	Buffer Overflow	Short	Root Shell

U2R attacks might consequence in important damage of period and money for several governments. Then, the recognition of such attacks is very vital to safeguard security in computer and network systems [71].

2.6.2 Remote to Local (R2L)

R2L attack includes illegal access from a remote instrument.

a) Imap: It reasons a buffer overflow by abusing a germ in the verification process of the imap server on about forms of LINUX. The attacker becomes root freedoms and can perform a random arrangement of instructions [72].

b) Ftp writes: This attack feat a misconfiguration moving inscribe freedoms of nameless books on an FTP server. This lets any ftp operator to enhance random records to the FTP server [73].

c) Ph.: This is an instance of seriously printed COl writings that is dispersed through the Apache server. Abusing this fault lets the attacker to perform codes with the http freedoms [74].

d) Warez master: This attack is likely in a state where inscribe consents are indecorously allocated.

e) FTP Seiter: When this is the circumstance, the attacker can upload reproductions of prohibited software that can then be transferred by additional operators.

f) Warez client: This attack contains in transferring unlawful software before uploads throughout a warez master attack. There are several approaches for their deterrence, but they all have approximately limits like variable countryside of attacks [75]. An example of numerous attacks examples is presented in Table 2.2.

Table 2.3: Remote to Local attacks [76]

Type	Signatures
Imap	5,tcp,imap4,SF,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2,3,0.00,0.00,0.50,0.67,0.50,1.00,1.00,6,58,0.17,0.50,0.17,0.05,0.00,0.02,0.17,0.97.
Ftp write	280,tcp,ftp_data,SF,283618,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,2,0.00,0.00,0.00,0.00,1.00,0.00,1.00,1,1,1.00,0.00,1.00,0.00,0.00,0.00,0.00,0.00.
Normal	0,tcp,http,SF,313,3545,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,11,11,0.00,0.00,0.00,0.00,1.00,0.00,0.00,11,255,1.00,0.00,0.09,0.02,0.00,0.00,0.00,0.00.
Warezmaster	0,udp,private,SF,28,0,0,3,0,0,0,0,0,0,0,0,0,0,0,0,0,0,69,69,0.00,0.00,0.00,0.00,1.00,0.00,0.00,255,69,0.27,0.02,0.27,0.00,0.00,0.00,0.00,0.00.
Warezclient	0,udp,private,SF,28,0,0,3,0,0,0,0,0,0,0,0,0,0,0,0,0,0,69,69,0.00,0.00,0.00,0.00,1.00,0.00,0.00,255,69,0.27,0.02,0.27,0.00,0.00,0.00,0.00,0.00.
Phf	0,icmp,ecr_i,SF,1480,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0.00,0.00,0.00,0.00,1.00,0.00,0.00,255,1,0.00,0.90,0.00,0.00,0.00,0.00,0.89,0.00.

2.6.3 Dos

DoS attacks nowadays are portion of each Internet user's life. They are trendy altogether the time, and all the Internet operators, as a public, have approximately share in making them, sorrow from them or smooth loosing time and cash since of them. DoS attacks do not have whatever to prepare with contravention into computers, captivating switch over remote hosts on the Internet or theft advantaged data like praise card records. Using the Internet way of language DoS is neither a Hack nor a Crack. It is an entire fresh and dissimilar topic [77].

A DoS attack is a class of attacks by individuals or groups of individuals who use the functions of Internet Scissors to prevent other employees from actually gaining access to models and data. Previously, DoS attacks were linked to SMURF attacks on routers. Hosts that were previously completely removed from the router are effectively blocked if the attacker allows the router to stop sending packets. In recent years, however, new attack models have been applied to attacks on web servers, mail servers and other objects [78].

Low-bandwidth networks can be impacted by increased coverage, which quickly increases bandwidth. The response speed depends on the collaboration between the service providers, as if the filters were distributed over the upstream routers [79].

Or DDoS punishes the target host for additional target parts in a combination of DoS attacks approved by other hosts in theaters and performances. This term is used to refer to multiple DDoS attack databases, rather than a single database. DDoS is often rejected by many opinions set by the administrator and cannot be removed simply by filtering the default IP address. Notable DDoS attack tools include Mstream, Torino, TFN2K (Tribe Flood Network), Barbed Wire, and Shaft. DDoS attacks are an example of attacks on bandwidth [80].

Table 2.4 which was modified from, illustrations dissimilar kinds of DoS with approximately belongings for a specific kind like the facility that the attack usages, the stages susceptible to this caring of "DoS", the kind of susceptibility (instrument) that the attack receipts benefit of, the period obligatory to tool it and the outcome produced by the attack. This tilt of DoS attacks is by no earnings thorough but stretches an impression of the diversity originate in this session of attacks [81].

Table 2.4: Denial of Service

Name	Service	Vulnerable Platforms	Mechanism	Time to Implement	Effect
Apache2	http	Any Apache	Abuse	Short	Crash httpd
Back	http type	Any Apache	Abuse/Bug	Short	Slow server response
Land	N/A	SunOS	Bug	Short	Freeze machine
Mailbomb	smtp	All	Abuse	Short	Annoyance
SYN Flood (Neptune)	Any TCP	All	Abuse	Short	Deny service on one or more ports for minutes
Ping of Death	icmp	None(*)	Bug	Short	Crash the targeted computer
Process Table	Any TCP	All	Abuse	Moderate	Deny new processes
Smurf	icmp	All	Abuse	Moderate or Long	Network Slowdown
Syslogd	syslog	Solaris	Bug	Short	Kill Syslogd
Teardrop	N/A	Linux	Bug	Short	Reboot machine
Udpstorm	echo/chargen	All	Abuse	Short	Network Slowdown
ARP Poisoning	ARP	All	Abuse	Short	Deny access to one or more machines on the network

3. MATERIAL AND METHODS

3.1 CYBER ATTACK METHODS

Introduced by IDS Anderson in 1980, a cyberattack attempt or threat, accessing information, changing information, or making a system unreliable or unusable. Cyber-attacks are generally divided into two according to their purpose. In the first type of attack, the system resources may be attempted. This is called an active attack. In the second type, an attack is made to learn information from the system and use this information in a way that does not affect system resources (e.g., wiretapping). Such attacks are called passive attacks. In addition, cyber-attacks are divided into two according to the place where they are made. Attacks by an entity within the security perimeter are defined as internal attacks. Attacks made by an unauthorized or illegitimate user of the system are called external attacks. Cyber-attacks are carried out using various methods. The following sections describe methods for international cyberattacks. intrusion Detection Systems (IDS) are security mechanisms that operate primarily at the level of the Internet of Things system [82].

The most obvious difference between deep learning and legacy networks is that deep learning is software-based, while legacy networks are typically hardware-based. Since deep learning is software-based, it is flexible and provides control and convenience for users to manage their resources. Big companies like Google, Yahoo, HP, and Cisco spend a lot of money to protect their systems and protect them from distractions. Existing networks require new hardware to increase network capacity. This allows organizations to use deep learning technologies to reduce costs. At the same time, deep learning provides high bandwidth to today's applications.

Over the past decade, the Internet of Things has grown significantly, reaching 4G and more than 20 billion devices, as well as smart applications such as computers, various smart devices, old or smart sensors, medicine, education, energy, and transportation. 5G Internet. he won. Access to Devices Due to the rapid growth and spread of IoT devices, this number is expected to exceed 50 billion by 2030. Many devices connected in this way are constantly exchanging and transmitting large amounts of data (Big Data) [83]. This saves valuable time and resources by identifying and predicting the most important features of

your dataset (such as columns), allowing you to analyze and focus on the most effective metrics [84]. Network error detection is displayed as an anomaly in network traffic. Credit card fraud, inability to access Internet services, and the inability to transfer sensitive information from one computer to another are some anomalies [85].

Network intrusion detection systems (NIDS) are essential to protect against malicious and suspicious cybersecurity activities. NIDS plays an important role in protecting against threats and attacks on network systems. NIDS faces many challenges and difficulties in detecting anomalies in network traffic. The increased variety of attacks and the use of traditional methods lead to high quality false positives.

Analyzing traffic behavior through monitoring, preventing the collection of sensitive information, and preventing access to computer systems without access rights is one of the main tasks of malware [86].

The differences between this study and other studies are as follows: This is anomaly detection using real data from the smart grid and data from the Internet of Things. The dataset used is the data taken from the actual network environment. The purpose of this study is to find the most appropriate measure by observing the effects of various properties on the results. Our model has been tweaked to provide exceptional performance and precision.

Moreover, meeting the needs of social users requires multimedia analytics, advanced processing, continuous data collection, high bandwidth and less complex computer encryption techniques [87]. However, it is very difficult for existing networks to meet these requirements. The trend is to integrate next-generation network technologies to provide quality of service (QoE) to target users. However, because social networks are expensive, complex, and open to investment and maintenance, it is important to address the security and privacy issues inherent in these networks before introducing new technologies. Therefore, it is necessary to establish a communication network to solve the problems created by multimedia social networks. One way to manage and control this network was to use software defined networks [88]. The most common anomaly detection algorithms are:

Fuzzy logic is a computational method based on the precision of modern computers. Intrusion detection activities also include many numeric attributes in the collected data and various derived statistics. Building models based directly on numerical data often leads to many false positives [89]. Similarly, Garbia et al. Detect malicious use in cyber services using sequential hypothesis testing algorithms in end-to-end path computation. Detection accuracy was very good, but false positives were high. Another drawback is that real-time data is not used [90]. Nguyen and Rogan monitored data from various ISPs and used hidden Markov models to detect anomalies. This works well for small data but quite poorly for large data [91]. There are exceptions for real and simulated network traffic using principal component analysis algorithms, but the calculation is very complex [92]. To avoid this complexity, Ray used a support vector machine in his studio [93]. SVM.

It is a machine learning algorithm with a teacher analyzing the given labeled data to categorize it. Draw a hyperplane to classify the data into categories. By definition, hashtags are used to detect cyberattacks or network hackers. SVM imposes an upper bound on arguments between different classes to reduce generalization errors. Logistic regression is basically a statistical model that uses logistic functions to model binary dependent variables. In regression analysis, logistic regression estimates the coefficients of a logic model [94].

Existing algorithms based on conventional heavyweight methods, fuzzy methods, classical neural networks, genetic algorithms, neural algorithms, statistical learning, and data mining go beyond traditional detection methods. However, these very traditional methods do not work as expected because they are inefficient, have no selection function, or use all properties of the dataset. Traditional IDS methods usually form all record properties or use the wrong record properties (columns) [95].These conditions increase the number of calculations in the packet processing time and consume more energy. However, since IoT systems mostly suffer from power efficiency issues, IoT systems need low power consumption IDS systems. Therefore, we need a lightweight (low power) IDS instead of heavy identifiers based on the characteristics of a complete or imprecise dataset. Consequently, the accuracy and performance of these methods are greatly reduced [96]. As most research on malware or intrusion detection shows, the traditional IDS approach faces positive challenges, and the high rate of false alarms reduces the quality of service (QoS)

of these networked systems. If a user pack is accidentally deleted, the user will receive a payment error message and the user pack will be delayed. Error-based identity recognition systems solve the problem of illegal scanning methods that threaten[97].

3.1.1 Security Principles in Wireless Sensor Networks

The ultimate goal of security services in wireless sensor networks is to protect detected data and network resources against attacks and misbehavior [98]. A secured sensor network should be able to securely collect data from the environment in which it is deployed, ensure that the content of the obtained data is not changed, and deliver the data securely to the target recipients. In order to ensure these principles, security principles such as data privacy, authentication, data integrity and sustainability of the network must be fulfilled.

3.1.2 Data Privacy

Data privacy is the mechanism that ensures that the data obtained by the sensor nodes is delivered only to the desired recipients and is prevented from being accessed by unauthorized persons, including neighboring nodes. Data privacy can be critical in some situations. The standard approach used to keep sensitive data private in sensor networks is to encrypt the data with a secret key that only the intended recipients have [99].

3.1.3 Data Integrity

Although mechanisms to ensure data confidentiality prevent information from being leaked, they alone are not sufficient to ensure data security. Because a malicious node can add fake packets to the network or change the data in the packet. Moreover, data loss/damage may occur even without an attack due to difficulties arising from the communication environment. Data integrity, on the other hand, ensures that the data is not changed during the transmission from the sending node to the receiving node. To ensure data integrity, message authentication codes (Message Authentication Code, MAC) or circular codes (cyclic codes) can be used [100].

3.1.4 Availability

Availability ensures that a sensor network can function uninterruptedly. The sustainability of the network requires that all the resources of the network, including the sensor and cluster head nodes and the base station, are ready for communication. The availability of the network can be disrupted either by external interventions such as denial of service attacks or by internal causes such as exhaustion of resources. This can cause serious losses in critical areas such as healthcare, disaster monitoring or military applications. E.g.; In disaster monitoring applications, large losses may occur due to the inability to detect a possible disaster, or a backdoor may be opened for an enemy invasion in a military application [101].

For the availability of the network, a centralized access control system has been proposed for additional communication between nodes or for successful forwarding of messages [102]. However, these approaches require additional calculation, thus additional energy consumption.

3.1.5 Data Freshness

While ensuring the confidentiality and integrity of the data, it is essential to safeguard the clarity of the data. Data freshness indicates that the information is new and guarantees that older communications are not retransmitted. This requirement becomes particularly significant when using keys. Ideally, a key generation process should ensure that every key (session key) shared with participants is new . Otherwise, a replay attack can be launched using the old key. A nonce or other time-related counter can be added to the packet to ensure data freshness [103].

3.1.6 Self-Organization

In smart citifies architecture, unlike traditional networks, there is no permanent transportation for network management. Sensor nodes should be flexible enough to organize themselves according to different situations for communication and fault tolerance purposes. Otherwise, in case of any attack or error on the network, devastating results are likely to occur. This feature, which is critical for the sustainability of the sensor network, also brings serious difficulties in terms of network security [104].

3.1.7 Secure Positioning

Sensor nodes must be correctly located and automatically discoverable. Because many WSN applications need accurate location information of nodes. E.g., In disaster monitoring applications, accurate location information will be needed to pinpoint the location of a potential disaster. However, attackers can manipulate location information by reporting false signal strength or repeating signal. Researchers have conducted several studies on secure positioning. Verifiable Multilateration (VM) mechanism based on range and distance limitations and Secure Range-Independent Localization (SERLOC) mechanism using location markers are some of these studies [105].

The security principles described for securing sensor network traffic are of primary and secondary importance. Principles that can be evaluated within the scope of the primary objective, principles of data privacy, authentication.

3.1.8 DoS-DDoS Attacks

DoS attacks are the most popular type of attack in network security with the development of the internet and the network. In a DoS attack, the system cannot respond to the service request by exceeding its resources. An infected host, controlled by an attacker, initiates a DoS attack. In this type of cyber-attack, the machine or network resources are rendered unusable for the intended user by disrupting the service of the host connected to the internet. Attacking from a single device with a single internet connection is called DoS and attacking from different parts of the internet from more than one connected device at the same time is called DDoS (Distributed Denial of Service) attack.

DDoS attacks are generally divided into 3 categories: volume-based, protocol and application layer attacks. In volume-based attacks, the goal is to saturate the bandwidth of the attacked site. Protocol attacks are attacks with the aim of consuming so much resources that the server cannot respond to valid requests. Application layer attacks, also called seventh layer attacks, cause overload by sending a large number of requests and requiring intensive resource use and processing on the server.

3.1.9 Attacks on privacy and authentication:

The sensor network's use of wireless communication infrastructure allows for the injection of fraudulent or malicious packets into the network and the surreptitious monitoring of sensitive data obtained by the sensors. Data privacy ensures that the data obtained by the sensor nodes is delivered only to the desired recipients and is prevented from being accessed by unauthorized persons, including neighboring nodes. The basic approach to ensure data privacy is the use of standard encryption techniques. The authentication process ensures that the transmitted packets come from the correct source. However, a compromised node has the ability to authenticate itself on the network because it has the secret keys of a legitimate node. Intrusion detection systems can be used to detect compromised nodes and prevent possible attacks.

3.1.10 Attacks on network availability:

Availability allows a sensor network to function uninterruptedly. Denial of service attacks (DoS) against the sensor network can disrupt the availability of the network. Network interruption can cause serious losses, especially in critical applications. For the availability of the network, a centralized access control system has been proposed for additional communication between nodes or for successful forwarding of messages [106]. However, these approaches require additional calculation, thus additional energy consumption.

3.1.11 Attacks on service integrity:

Injecting fake packets into the sensor network or changing the data in the existing packet causes the network to produce erroneous results. To ensure data integrity, MAC or circular codes can be used [107].

Attacks on wireless sensor networks can be classified into passive attacks and active attacks according to their structure. Because sensor networks are connected wirelessly, passive attackers can easily listen to the network's radio frequency bands and steal personal or sensitive information. In fact, a lot of information can be obtained from sensor networks through direct field observations. However, passive attacks do not require a physical presence in the field for monitoring and surveillance activities and can collect information anonymously with little risk.

Basically, in a passive privacy attack, an attacker monitors unencrypted traffic for sensitive information that can be used in other types of attacks. Passive attacks include monitoring and analyzing network traffic, decrypting weakly encrypted traffic, and gathering credentials. Attacks collect user information or data files without the user's consent or knowledge. [108].

3.1.12 Attacks on service integrity

Injecting fake packets into the sensor network or changing the data in the existing packet causes the network to produce erroneous results. To ensure data integrity, MAC or circular codes can be used [109].

Attacks against wireless sensor networks can be classified as passive or active according to their nature. While the sensor network communicates over a wireless channel, a passive attacker can easily eavesdrop on the network's radio frequency range in order to steal private or sensitive information [110]. In fact, a lot of information from the sensor network could possibly be gathered through direct field surveillance. However, in passive attacks, there is no need to be physically present in the field for monitoring and surveillance activities, and information can be collected anonymously with low risk [111].

3.2 DATA MINING TECHNIQUES

3.2.1 K-means Clustering

k-means clustering is a technique of vector quantization, initially from signal processing, that goals to divider n comments into k clusters in which individually comment fits to the cluster with the nearest mean [112]. The steps of algorithm listed below:

A. Primary we prepare k facts, named means, arbitrarily.

B. We classify individually element to its neighboring nasty, and we inform the averages organizes, which are the means of the substances branded in that nasty so distant.

C. We recurrence the procedure for an assumed amount of iterations and at the end, we have our clusters.

3.2.2 Support Vector Machine (SVM)

SVM is a technique which can be applied for both regression or classification tests. Though, it is typically applied in organization difficulties. In the SVM technique, we conspiracy apiece feature item as an opinion in n-dimensional interplanetary with the worth of apiece feature existence the charge of a specific organize. Before, we achieve organization by discovery the hyper-plane that distinguishes the two sessions actual well. The Figure 3.1 presented the SVM procedure [113].

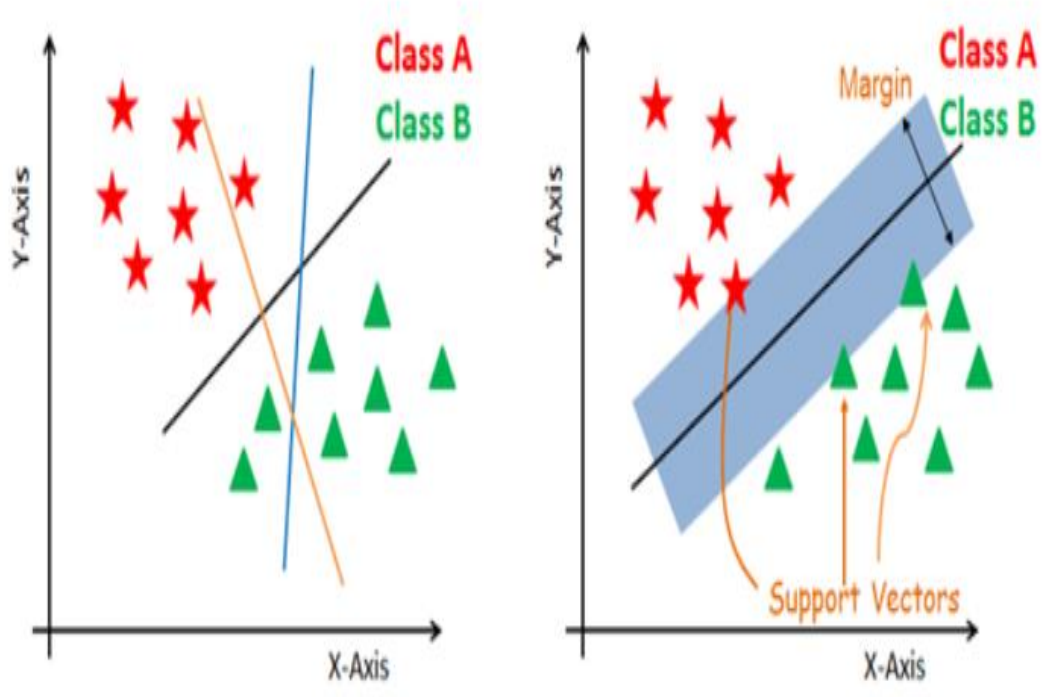

Figure 3.1: SVM Process [114]

3.2.3 Decision Tree (DT)

DTs are a non-parametric supervised learning technique applied for regression and classification. DT learns from data to estimate a sine curve with a set of if-then-else verdict rules. DT constructs regression or classification representations in the form of a tree construction. It disruptions depressed a feature set into lesser and minor subsections although at the similar time a related DT is incrementally industrialized. The concluding outcome is a tree with decision bulges and foliage bulges. A decision bulge has two or

extra twigs. Leaf node signifies a decision or classification. The highest decision bulge in a tree which agrees to the greatest forecaster named root bulge. DTs can grip both definite and arithmetical feature [115].

3.2.4 Neural Network

It's a method for office block a computer software that acquires from feature. It is depending very insecurely on how human reason the human brain mechanism [116]. First, a group of software "neurons" are produced and associated composed, letting them to show communications to each supplementary. Following, the network is requested to resolve a problem, which it efforts to do over and over, separately period consolidation the influences that principal to achievement and lessening those that principal to disappointment. Figure 3.2 presented the structure of MLP.

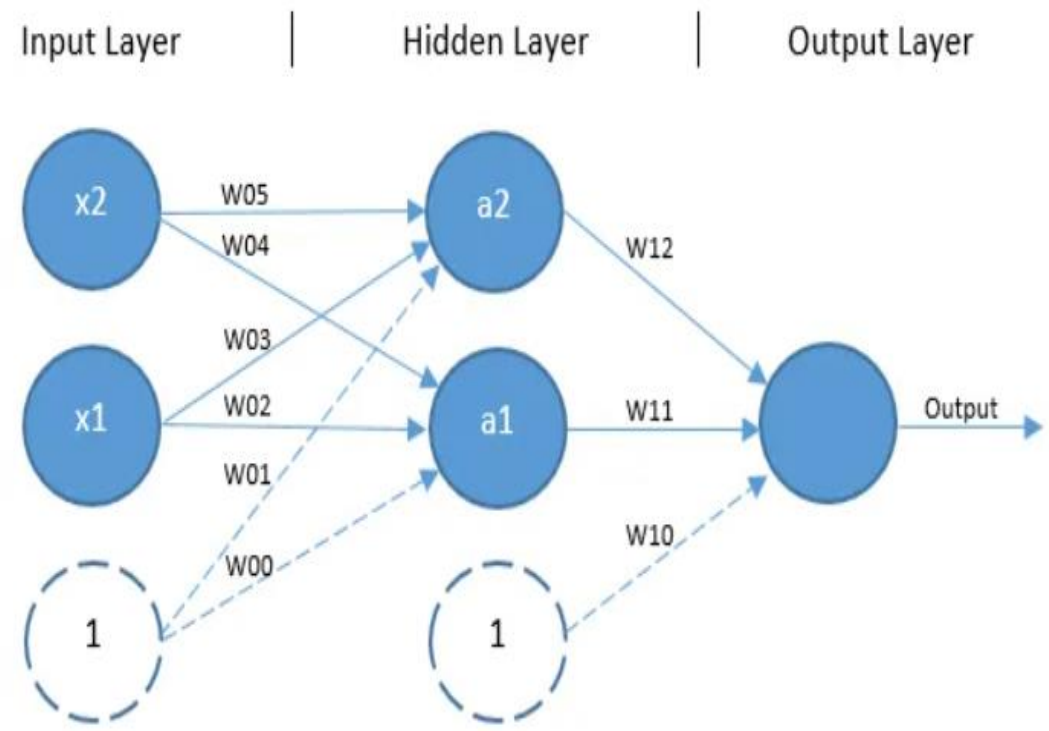

Figure 3.2: Neural Network Structure [117]

3.3 DEEP LEARNING (DL)

DL has changed hand-in-hand with the numerical age, which has transported about an bang of feature in very procedures and from each area of the domain. This data, recognized just as big feature, is haggard from bases similar communal media, internet search trains, e-commerce stages, and online cinemas, among others [118]. This huge quantity of feature is willingly nearby and can be communal finished fintech software's such as cloud computing.

Name	Service	Vulnerable Platforms	Mechanism	Time to Implement	Effect
Apache2	http	Any Apache	Abuse	Short	Crash httpd
Back	http type	Any Apache	Abuse/Bug	Short	Slow server response
Land	N/A	SunOS	Bug	Short	Freeze machine
Mailbomb	smtp	All	Abuse	Short	Annoyance
SYN Flood (Neptune)	Any TCP	All	Abuse	Short	Deny service on one or more ports for minutes
Ping of Death	icmp	None(*)	Bug	Short	Crash the targeted computer
Process Table	Any TCP	All	Abuse	Moderate	Deny new processes
Smurf	icmp	All	Abuse	Moderate or Long	Network Slowdown
Syslogd	syslog	Solaris	Bug	Short	Kill Syslogd
Teardrop	N/A	Linux	Bug	Short	Reboot machine
Udpstorm	echo/chargen	All	Abuse	Short	Network Slowdown
ARP Poisoning	ARP	All	Abuse	Short	Deny access to one or more machines on the network

3.4 DATASET

The KDD99 dataset and the extended NSL KDD type are two of the well-known intrusion detection datasets, but they are relatively old datasets and most of the traditional designs and attack designs have changed from the previous period. To address this problem, in 2015 we completed a comprehensive system-specific data set that replicates modern system traffic scenarios from various interventions with low impact and complex organized system traffic data. The IDS UNSW-NB15 is a pool of approx. 100 GB of system traffic, which consists of the Cyber Range Lab in the Cyber Security Center (ACCS) in Australia. That dataset now contains the characteristics of today's normal and man-made attacks. This dataset contains 2,540,044 examples, including 48 data and reports

on typical anal attacks and 9 attacks. The research team created subsections of exercise routines and test suites with 175,341 and 82,332 records, respectively. In this study, the accuracy of the model was assessed using a 20% subsection of the UNSW-NB15 dataset [119].

3.5 PROPOSED METHOD

This study presented new study based deep sparse autoencoders and TSA. The two autoencoders applied to reduce the size of input data. The aim of applying Autoencoders to extracted high level features gradually and then wired the features to the SoftMax. The output of this step wired to the second autoencoder, which used to classify the extracted features by the first autoencoder. The deep sparse autoencoders are good at classifying, processing, and forecasting based on time series data because delays of unknown duration can occur between major events in a time series. Sparse autoencoders are designed to address the loss gradient problem that can occur in traditional autoencoders formation. Relative spatial length insensitivity is an advantage of sparse autoencoders in many applications over autoencoders,

$$f_t = \left(W_f X_t + U_f c_t + b_f\right) \tag{3.2}$$

Then, the TSA applied for w and b estimation to obtain best accuracy for the sparse autoencoders classification. The TSA was proposed by Kiran in 2015 based on the basic idea of the relationship between trees and seeds to solve continuous optimization problems. At TSA, trees and seeds represent possible solutions to optimization problems. Initially, the trees are randomly generated in the search area. The number of trees in the TSA is called "planting size", it is investigated in, and it is recommended between 10% (minimum number of seeds) and 25% (maximum number of seeds). The seeds are generated by equations. (3.3) or (3.4) for each tree in each iteration.

$$S(k) = T(i) + \propto \left(B - T(r)\right) \tag{3.3}$$

$$S(k) = T(i) + \propto \left(T(i) - T(r)\right) \quad (3.4)$$

where $T(i)$ is ith tree, $S(k)$ is kth seed of $T(i)$, $\propto$ is a consistently dispersed chance amount in the assortment of [-1, 1], B is the finest tree gotten so far, $T(r)$ is a chance tree which is dissimilar from the $T(i)$. The most important kind is what equation is defined to create a new starting location. This selection (either equation (1) or equation (2)) is controlled by a method control parameter called search trend (ST) in the range [0, 1]. Higher ST values provide reliable local searches and faster convergence. A lower ST value results in slow convergence but strong global search. In other words, the TSA detection and usage functions are controlled by the ST parameter. In short, TSA detection and usage functions are controlled by the ST parameter.

At the beginning of the search with TSA, the initial tree positions are created using the equation that represents possible solutions to the optimization problem. (3.5).

$$T_{i,j} = L_{j,min} + r_{i,j}\left(H_{j,max} - L_{j,min}\right) \quad (3.5)$$

where $L_{j,min}$ is the lower bound of the search space, $H_{j,max}$ is the higher bound of the search space and $r_{i,j}$ is a random number produced for each dimension and location, in range of [0, 1]. For minimization, the best solution is selected from the population using Eq. (3.6).

$$B = \min\{f(\vec{T}_i)\}\ i = 1, 2, \ldots., N \quad (3.6)$$

where N is the number of trees in the population. When creating new seed locations for a tree, the number of seeds can be greater than one, and this number depends on the size of the population. When analysing the effect of the control parameters on the efficiency of

CST, 10% of the population size is the minimum number of seeds produced for the tree and 25% of the population size is the maximum number of seeds produced. made of wood. The number of semen production in TSA is completely random

4. RESULTS

First, the data was evaluated with the first automated feature extraction encoder. The goal of this phase is to clear away unnecessary data and minimize execution time. This happened so because autoencoder is a basic mathematical model that extracts useful and crucial characteristics.

The second section of the table is indeed the second part of the automated encoder, which is an abbreviation of a function. The first auto-encoder trained, as well as the training procedure shown in the Figure 4.1.

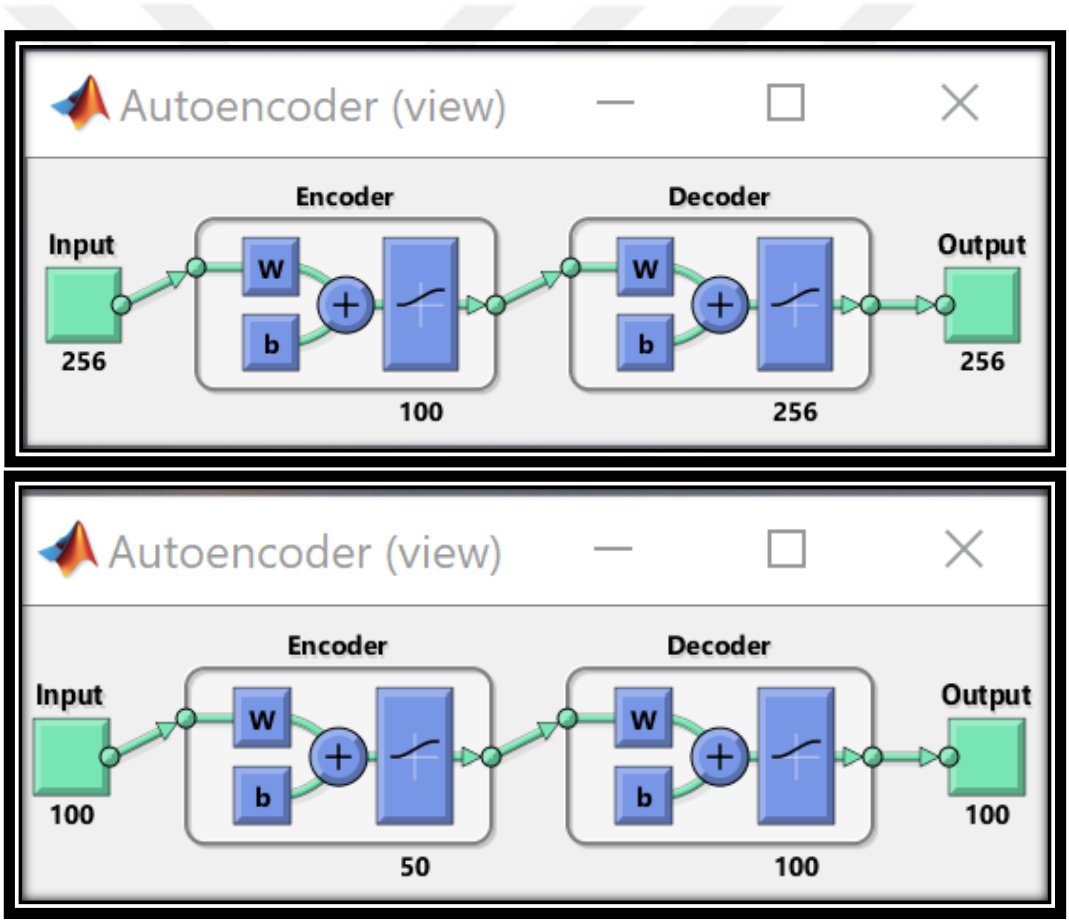

Figure 4.1: First Auto-encoder.

The last layer of the deep learning technique in this study, is the SoftMax which trained in supervised fashion. This lead to classify the extracted features by the auto-encoders to the labels normal and abnormal. The Figure 4.2 presented the training step of SoftMax.

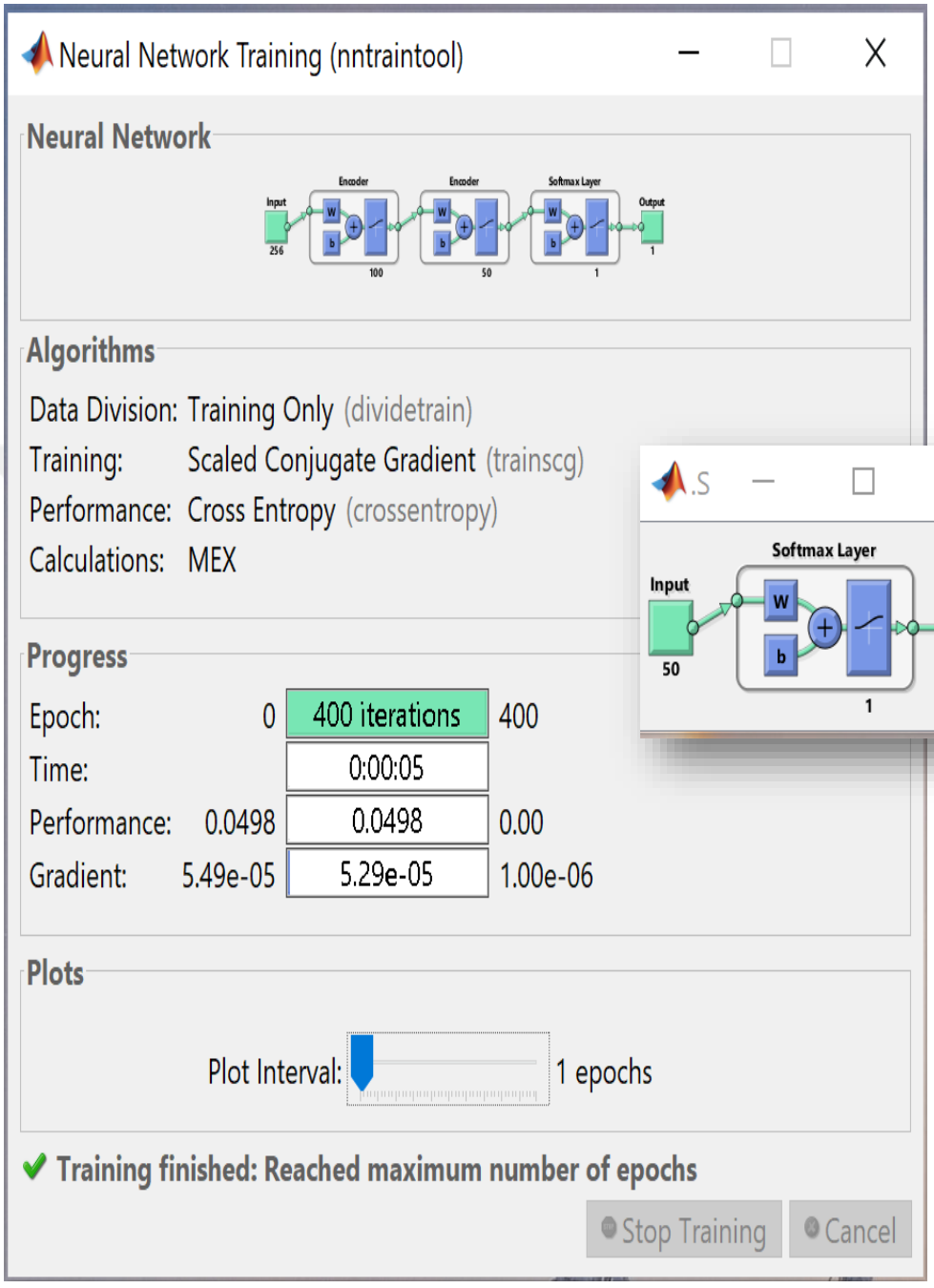

Figure 4.2: SoftMax Training Process.

The TSA applied to select best weight and basis for the autoencoders which mean that the TSA used as training technique.

Then, several parameters such as accuracy, F1 Score, and Matthews Correlation Coefficient which these results are presented in the Table 3.1. furthermore, the confusion matrix presented in 4.3.

Parameters	**Values (%)**
Accuracy	98
F1 Score	96.10
Matthews Correlation Coefficient	95.90

Table 4.1: Results

We obtained the dataset from the KAGGLE machine learning website

Calculate the average power of each train from the row data obtained from the UCI website which leads to reducing the size of the data and extracting, If the length of the vector is not a multiple of matrix columns the operations are repeated for the test data, and the data that are calculated by the average of power function are wired to the deep sparse auto-encoder

the first auto-encoder reduces the size of the input data from 256 to 100 which is the number of the first hidden layer. the data that are extracted by the first auto-encoder are wired to the second sparse auto-encoder.

the second auto-encoder reduces the size of the input data from 100 to 50 which is the number of the first hidden layer, SoftMax was applied to train the extracted features by the autoencoders to the labels, was stacked the autoencoders and SoftMax and trained in the supervised technique, calculated the confusion matrix for the model to evaluate the results

The confusion matrix used to evaluate the proposed metho results as shown, in figure 4.3

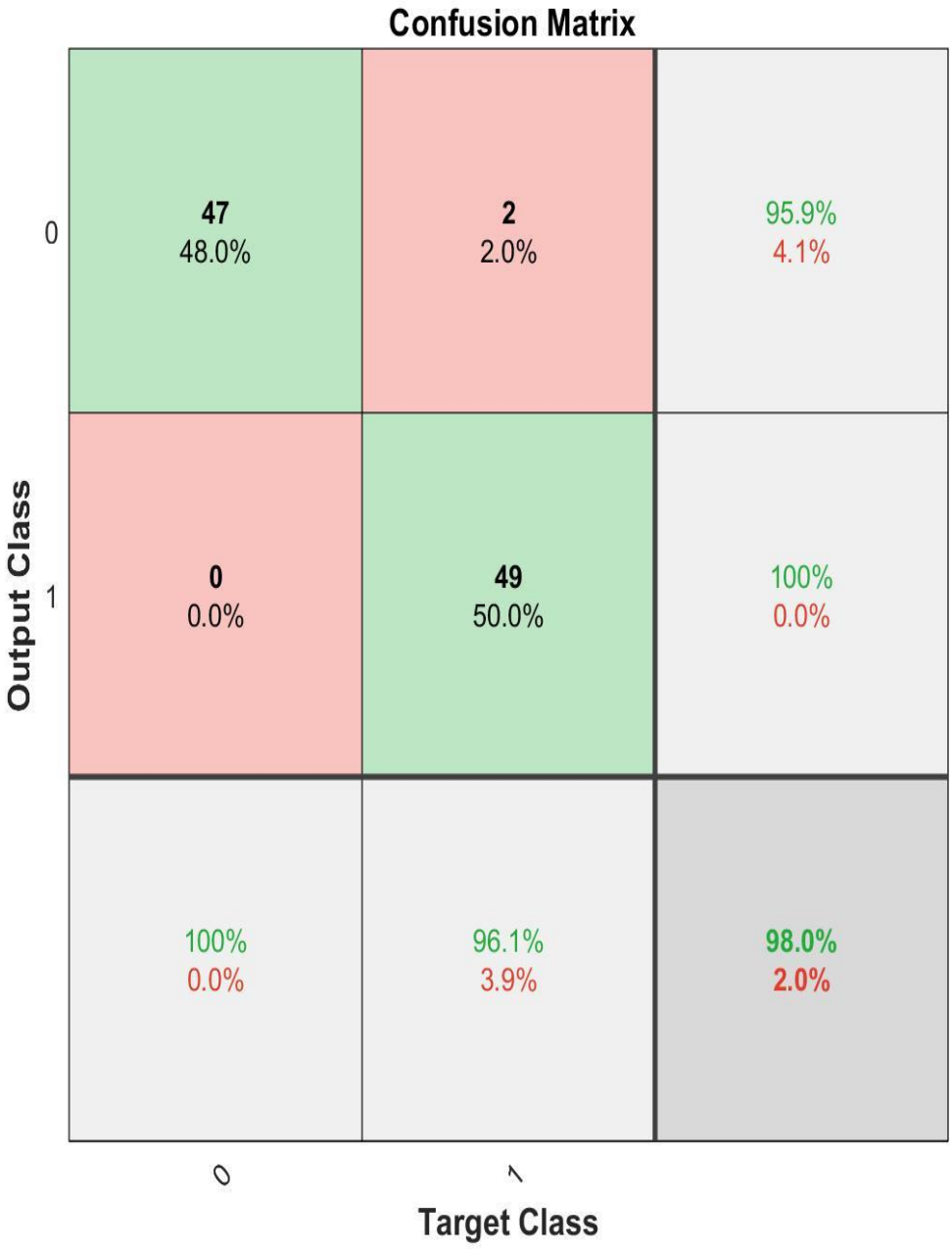

Figure 4.3: Confusion Matrix.

The presented results would be compared with several state of art studies to evaluate the presented method. The previous studies applied COS-P-Poly3 in the [121] and presented 88.84% accuracy. In the [122] Logistic Regression classifier and DT applied to detect malware attack and show suitable results which presented 86.6 and 96.7 % respectively of accuracy.

Then, our method presented 98 which is remarkable when compared with previous studies. The main issue in our method we combine TSA with deep sparse auto-encoder to obtain best accuracy.

Table 4.2: Comparison our method with previous studies

Methods	**Methods**	**ACC (%)**
[121]	COS-P-Poly3	88.84
[122]	Logistic Regression	86.6
[122]	Decision Tree	96.7
[120]	Linear SVM	93.1
[122]	Ensemble learning	95.1
[123]	fairly adequate+ classifier	89
[124]	Markov chain + Random Forest	92
Our Method	TSA based deep sparse auto-encoders	98.00

In [120] Alzaylaee et L. The authors proposed a second investigation in which actual phones were utilised for automated feature extraction in order to overcome the constraints of dynamic analysis of particular Android viruses. It is automatically determined which dynamic properties come from sample apps that are being executed on Android phones. In addition, some of the testing included benchmarking malware detection algorithms utilising simulators and devices, which was done as part of the research. Alzaylaee and colleagues conducted research in which they found This is particularly essential because a simulator is used in the majority of research based on dynamic and behavioural analysis. We employed 222 malware samples from the MalGenome project as well as 1222 benign

samples from McAfee Labs, which is affiliated with Intel Security, to conduct our research and analysis. Another recent study assisted by artificial intelligence assessed the source codes of 200 malicious Android applications as well as the source codes of 200 non-malicious Android applications, yielding an F1 score of 95.1 percent [123].

Another study [124] employed real-time machine learning techniques on the device to extract behavioural attributes from Android malware. Researchers used reports from the Cuckoo Sandbox file analysis platform to extract behavioural attributes from Android malware.

It was obtained through a malware sharing portal known as Virus Share, then they ran a machine learning (ML)algorithm on the device they built on a total of 2000 malware samples belonging to 18 different malware families, all of which were obtained from of the malware sharing platform. The accuracy rate of 89 percent was achieved by the usage of tenfold cross-validation. It was possible to get an accuracy rate of 89 percent by employing this strategy (10-fold cross-validation).

Recently published study by Onwuzurike et al[125]. suggests a framework that integrates static, dynamic, as well as hybrid analytic methodologies, using a model based similar to MAMADROID. Researchers found that combining static and dynamic analysis approaches produced the best results, with just an F1 score of 92 percent. They also found that static analysis techniques were least as successful at dynamic analysis techniques, if not more so [126]

.

5. CONCLUSIONS

Developments in Internet technologies and rapidly changing traffic demands require new network security tools and support. Therefore, in this study, the IDS method was used instead of the traditional network method. Since the traditional network infrastructure used today cannot meet the ever-increasing requirements (big data transmission, speed, cloud computing, virtualization) in the world of informatics, the IDS method was used based on the OpenFlow protocol.

Furthermore, by utilising the bot-Iot dataset, we hope to overcome some of the drawbacks of existing datasets in the future. This dataset was built with real data, and I used the most recent version of the software (real data for 2018 was used on my lab). It was at the UNSW Canberra Cyber Center that the BoT-IoT dataset was generated, which involved establishing a realistic network environment.

Over the past decade, deep learning has been shown to provide higher accuracy in anomaly detection than traditional methods. Developers have developed various methods to detect and prevent attacks based on network anomalies. Traditionally, detecting different attack scenarios and protecting networks has been a daunting task. Intrusion detection systems are designed to analyze and detect different types of network attacks to provide a more secure environment for communication.

The proposed framework will be used in future computer network security, speech recognition, video classification, etc.

REFERENCES

[1] S. Stolfo, W. Fan, W. Lee, A. Prodromidis, and P. Chan, “Cost-based modeling for fraud and intrusion detection: results from the jam project, in: DARPA Information Survivability Conference and Exposition,” DISCEX’00, Proc., vol. 2, no. IEEE, 2000, pp. 130–144, 2000.

[2] M. Alkasassbeh, A. B. A. Hassanat, and G. Al-naymat, “Detecting Distributed Denial of Service Attacks Using Data Mining Techniques,” vol. 7, no. 1, pp. 436–445, 2016.

[3] A. Abraham, C. Grosan, and C. Martin-Vide, “Evolutionary design of intrusion detection programs,” Int. J. Netw. Secur., vol. 4, no. 3, pp. 328–339, 2007.

[4] A. M. Brues, “Genetic effects of the atom bomb,” J. Hered., vol. 38, no. 5, pp. 137–137, 1947.

[5] H. Liu, Y. Sun, and M. S. Kim, “Fine-grained DDoS detection scheme based on bidirectional count sketch,” Proc. - Int. Conf. Comput. Commun. Networks, ICCCN, 2011.

[6] C. Khammassi and S. Krichen, “A GA-LR wrapper approach for feature selection in network intrusion detection,” Comput. Secur., vol. 70, pp. 255–277, 2017.

[7] A. Abraham and J. Thomas, “Distributed intrusion detection systems: a computational intelligence approach,” Idea Gr. Inc. Publ. Usa, Chapter, vol. 5, pp. 1–28, 2005.

[8] N. Sharma and S. Mukherjee, “A Novel Multi-Classifier Layered Approach to Improve Minority Attack Detection in IDS,” Procedia Technol., vol. 6, pp. 913–921, 2012.

[9] B. Škrbić and N. Durišić-Mladenović, “Principal component analysis for soil contamination with organochlorine compounds,” Chemosphere, vol. 68, no. 11, pp. 2144–2152, 2007.

[10] N. Patani and R. Patel, “A Mechanism for Prevention of Flooding based DDoS Attack,” vol. 13, no. 1, pp. 101–111, 2017.

[11] L. K. Xr et al., “8VLQJ 6WDFNHG ’ HQRLVLQJ $ XWRHQFRGHU IRU WKH,” pp. 483–488, 2017.

[12] Y. Feng, R. Guo, D. Wang, and B. Zhang, “Research on the active DDoS filtering algorithm based on IP flow,” 5th Int. Conf. Nat. Comput. ICNC 2009, vol. 4, no. October, pp. 628–632, 2009.

[13] A. Meek, “DDoS attacks are getting much more powerful and the Pentagon is scrambling for solutions,” 2015.

[14] J. Mirkovic and P. Reiher, “A taxonomy of DDoS attack and DDoS defense mechanisms,” ACM SIGCOMM Comput. Commun. Rev., vol. 34, no. 2, p. 39, 2004.

[15] S. Taghavi Zargar, J. Joshi, D. Tipper, and S. Member, “A Survey of Defense Mechanisms Against Distributed Denial of Service (DDoS) Flooding Attacks,” pp. 1–24, 2013.

[16] T. T. Oo and T. Phyu, “Analysis of DDoS Detection System based on Anomaly Detection System,” 2014.

[17] S. Sinha and M. Sharma, “Simulation and Analysis of DDoS Attacks by Specialized Simulator using Virtualization,” vol. 3, no. 2, pp. 2–4, 2014.

[19] B. Hang, R. Hu, and W. Shi, “An enhanced SYN cookie defence method for TCP DDoS attack,” J. Networks, vol. 6, no. 8, pp. 1206–1213, 2011.

[20] D. Wang, Z. Yufu, and J. Jie, “A multi-core based DDoS detection method,” Proc. - 2010 3rd IEEE Int. Conf. Comput. Sci. Inf. Technol. ICCSIT 2010, vol. 4, pp. 115–118, 2010.

[21] N. Hoque, H. Kashyap, and D. K. Bhattacharyya, “Real-time DDoS attack detection using FPGA,” Comput. Commun., vol. 110, pp. 48–58, 2017.

[22] J. Singh, M. Sachdeva, and K. Kumar, “Detection of DDoS Attacks Using Source IP Based Entropy,” vol. 3, no. 1, pp. 201–210, 2013.

[23] S. M. Lee, D. S. Kim, J. H. Lee, and J. S. Park, “Detection of DDoS attacks using optimized traffic matrix,” Comput. Math. with Appl., vol. 63, no. 2, pp. 501–510, 2012.

[24] H. Rahmani, N. Sahli, and F. Kamoun, “DDoS flooding attack detection scheme based on F-divergence,” Comput. Commun., vol. 35, no. 11, pp. 1380–1391, 2012.

[25] Senirkentli, Güler B., Fatih Ekinci, Erkan Bostanci, Mehmet S. Güzel, Özlem Dağli, Ahmad M. Karim, and Alok Mishra. 2021. "Proton Therapy for Mandibula Plate Phantom" Healthcare 9, no. 2: 167. https://doi.org/10.3390/healthcare9020167.

[26] V. P. Nigam and D. Graupe, “A neural-network-based detection of epilepsy,” Neurol. Res., vol. 26, no. 1, pp. 55–60, 2004.

[27] A. Abraham, U. States, and C. Grosan, “Genetic Systems Programming,” vol. 13, no. May, 2006.

[28] L. Deng and D. Yu, “Deep Learning: Methods and Applications,” Found. Trends® Signal Process., vol. 7, no. 3–4, pp. 197–387, 2014.

[29] K. Ahmed, K. Nia, S. A. Khan, and A. Shaukat, “Identifying Be est Feature Subset for Ca ardiac Arrhy ythmia Classification,” pp. 494–499, 2015.

[30] A. Shenfield, D. Day, and A. Ayesh, “Intelligent intrusion detection systems using artificial neural networks,” ICT Express, vol. 4, no. 2, pp. 95–99, 2018.

[31] A. Kaushik, H. Gupta, and D. S. Latwal, “Impact of Feature Selection and Engineering in the Classification of Handwritten Text,” 2016 Int. Conf. Comput. Sustain. Glob. Dev., pp. 2598–2601, 2016.

[32] A. Gupta, A. T. Müller, B. J. H. Huisman, J. A. Fuchs, P. Schneider, and G. Schneider, “Generative Recurrent Networks for De Novo Drug Design,” Mol. Inform., vol. 37, no. 1, 2018.

[33] S. Ibrahim, R. Djemal, and A. Alsuwailem, “Electroencephalography (EEG) signal processing for epilepsy and autism spectrum disorder diagnosis,” Biocybern. Biomed. Eng., vol. 38, no. 1, pp. 16–26, 2018.

[34] N. Kohli, N. K. Verma, and A. Roy, “SVM based methods for arrhythmia classification in ECG,” 2010 Int. Conf. Comput. Commun. Technol. ICCCT-2010, pp. 486–490, 2010.

[35] V. Sze, Y.-H. Chen, T.-J. Yang, and J. Emer, “Efficient Processing of Deep Neural Networks: A Tutorial and Survey,” vol. 105, no. 12, pp. 2295–2329, 2017.

[36] D. Petkovic et al., “SETAP: Software engineering teamwork assessment and prediction using machine learning,” 2014 IEEE Front. Educ. Conf. Proc., pp. 1–8, 2014.

[37] C. Sobie, C. Freitas, and M. Nicolai, “Simulation-driven machine learning: Bearing fault classification,” Mech. Syst. Signal Process., vol. 99, pp. 403–419, 2018.

[38] J. H. Lee, J. Shin, and M. J. Realff, “Machine learning: Overview of the recent progresses and implications for the process systems engineering field,” Comput. Chem. Eng., 2017.

[39] T. M. Mitchell, (Mcgraw-Hill International Edit) Thomas Mitchell-Machine learning-McGraw Hill Higher Education (1997). .

[40] S. RAY, “Understanding Support Vector Machine algorithm from examples,” 2017.

[41] A. M. Karim, Ö. Karal, and F. V Çelebi, “A New Automatic Epilepsy Serious Detection Method by Using Deep Learning Based on Discrete Wavelet Transform,” no. 4, pp. 15–18, 2018.

[42] Ahmad M. Karim, Mehmet S. Güzel, Mehmet R. Tolun, Hilal Kaya, Fatih V. Çelebi, A new framework using deep auto-encoder and energy spectral density for medical waveform data classification and processing, Biocybernetics and Biomedical Engineering, Volume 39, Issue 1,2019,Pages 148-159, ISSN 0208-5216, https://doi.org/10.1016/j.bbe.2018.11.004.

[43] S. M. Hosseini Bamakan, H. Wang, and Y. Shi, “Ramp loss K-Support Vector Classification-Regression; a robust and sparse multi-class approach to the intrusion detection problem,” Knowledge-Based Syst., vol. 126, pp. 113–126, 2017.

[44] A. M. Karim, F. V. Çelebi, and A. S. Mohammed, "Software Development for Blood Disease Expert System," Lecture Notes on Empirical Software Engineering, vol. 4, no. 3, pp. 179–183, 2016.

[45] A. M. Karim, M. S. Güzel, M. R. Tolun, H. Kaya, and F. V Çelebi, "A New Generalized Deep Learning Framework Combining Sparse Auto-encoder and Taguchi Method for Novel Data Classification and Processing," pp. 1–22.

[46] L. Wang, C. Wang, W. Du et al., "Parameter optimization of a four-legged robot to improve motion trajectory accuracy using signal-to-noise ratio theory," Robotics and Computer-Integrated Manufacturing, vol. 51, pp. 85–96, 2018..

[47] R. Huang, C. Liu, G. Li, and J. Zhou, "Adaptive Deep Supervised Autoencoder Based Image Reconstruction for Face Recognition," Mathematical Problems in Engineering, vol. 2016, 2016.

[48] Alex Sherstinsky, Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network, Physica D: Nonlinear Phenomena, Volume 404, 2020, 132306, ISSN 0167-2789.

[49] Pritika Bahad, Preeti Saxena, Raj Kamal, Fake News Detection using Bi-directional LSTM-Recurrent Neural Network, Procedia Computer Science, Volume 165, 2019, Pages 74-82, ISSN 1877-0509.

[50] Zhiyong Cui, Ruimin Ke, Ziyuan Pu, Yinhai Wang, Stacked bidirectional and unidirectional LSTM recurrent neural network for forecasting network-wide traffic state with missing values, Transportation Research Part C: Emerging Technologies, Volume 118, 2020, 102674, ISSN 0968-090X.

[51] Ying Chen, Voltages prediction algorithm based on LSTM recurrent neural network, Optik, Volume 220, 2020, 164869, ISSN 0030-4026.

[52] Neculai Andrei, A Dai–Yuan conjugate gradient algorithm with sufficient descent and conjugacy conditions for unconstrained optimization, Applied Mathematics Letters, Volume 21, Issue 2, 2008, Pages 165-171, ISSN 0893-9659.

[53] Muhammed Maruf Öztürk, İbrahim Arda Cankaya, Deniz İpekçi, Optimizing echo state network through a novel fisher maximization based stochastic gradient descent, Neurocomputing, Volume 415, 2020, Pages 215-224, ISSN 0925-2312.

[54] Andreea Koreanschi, Oliviu Sugar Gabor, Joran Acotto, Guillaume Brianchon, Gregoire Portier, Ruxandra Mihaela Botez, Mahmoud Mamou, Youssef Mebarki, Optimization and design of an aircraft's morphing wing-tip demonstrator for drag reduction at low speed, Part I – Aerodynamic optimization using genetic, bee colony and gradient descent algorithms, Chinese Journal of Aeronautics, Volume 30, Issue 1, 2017, Pages 149-163, ISSN 1000-9361.

[55] Arnulf Jentzen, Philippe von Wurstemberger, Lower error bounds for the stochastic gradient descent optimization algorithm: Sharp convergence rates for slowly and fast decaying learning rates, Journal of Complexity, Volume 57, 2020, 101438, ISSN 0885-064X, https://doi.org/10.1016/j.jco.2019.101438..

[56] Rui Ye, Qun Dai, Implementing transfer learning across different datasets for time series forecasting, Pattern Recognition, Volume 109, 2021, 107617, ISSN 0031-3203,.

[57] Yu Liu, Ao Li, Xing-Ming Zhao, Minghui Wang, DeepTL-Ubi: A novel deep transfer learning method for effectively predicting ubiquitination sites of multiple species, Methods, 2020, ISSN 1046-2023,..

[58] Jingyao Wu, Zhibin Zhao, Chuang Sun, Ruqiang Yan, Xuefeng Chen, Few-shot transfer learning for intelligent fault diagnosis of machine, Measurement, Volume 166, 2020, 108202, ISSN 0263-2241,..

[59] Shanshan Wang, Lei Zhang, Jingru Fu, Adversarial transfer learning for cross-domain visual recognition, Knowledge-Based Systems, Volume 204, 2020, 106258, ISSN 0950-7051,..

[60] Xiyun Yang, Yanfeng Zhang, Wei Lv, Dong Wang, Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier, Renewable Energy, 2020, ISSN 0960-1481,..

[61] Chuan Li, Shaohui Zhang, Yi Qin, Edgar Estupinan, A systematic review of deep transfer learning for machinery fault diagnosis, Neurocomputing, Volume 407, 2020, Pages 121-135, ISSN 0925-2312,..

[62] Zhengjun Qiu, Shutao Zhao, Xuping Feng, Yong He, Transfer learning method for plastic pollution evaluation in soil using NIR sensor, Science of The Total Environment, Volume 740, 2020, 140118, ISSN 0048-9697,..

[63] Santi Kumari Behera, Amiya Kumar Rath, Prabira Kumar Sethy, Maturity status classification of papaya fruits based on machine learning and transfer learning approach, Information Processing in Agriculture, 2020, ISSN 2214-3173,..

[64] Seunghyeon Kim, Yung-Kyun Noh, Frank C. Park, Efficient neural network compression via transfer learning for machine vision inspection, Neurocomputing, Volume 413, 2020, Pages 294-304, ISSN 0925-2312,..

[65] Xiang Li, Wei Zhang, Hui Ma, Zhong Luo, Xu Li, Partial transfer learning in machinery cross-domain fault diagnostics using class-weighted adversarial networks, Neural Networks, Volume 129, 2020, Pages 313-322, ISSN 0893-6080,..

[66] Piyush Kant, Shahedul Haque Laskar, Jupitara Hazarika, Rupesh Mahamune, CWT Based Transfer Learning for Motor Imagery Classification for Brain computer Interfaces, Journal of Neuroscience Methods, Volume 345, 2020, 108886, ISSN 0165-0270,..

[67] Xinghua Li, Zhongyuan Hu, Mengfan Xu, Yunwei Wang, Jianfeng Ma, Transfer learning based intrusion detection scheme for Internet of vehicles, Information Sciences, Volume 547, 2021, Pages 119-135, ISSN 0020-0255,..

[68] Sheikh Saud, Basharat Jamil, Yogesh Upadhyay, Kashif Irshad, Performance improvement of empirical models for estimation of global solar radiation in India: A k-fold cross-validation approach, Sustainable Energy Technologies and Assessments, Volume 40, 2020, 100768, ISSN 2213-1388,..

[69] Jie Wei, Hui Chen, Determining the number of factors in approximate factor models by twice K-fold cross validation, Economics Letters, Volume 191, 2020, 109149, ISSN 0165-1765,.

[70] Tzu-Tsung Wong, Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation, Pattern Recognition, Volume 48, Issue 9, 2015, Pages 2839-2846, ISSN 0031-3203,..

[71] Gaoxia Jiang, Wenjian Wang, Error estimation based on variance analysis of k-fold cross-validation, Pattern Recognition, Volume 69, 2017, Pages 94-106, ISSN 0031-320..

[72] Nima Shiri Harzevili, Sasan H. Alizadeh, Analysis and modeling conditional mutual dependency of metrics in software defect prediction using latent variables, Neurocomputing, Volume 460, 2021, Pages 309-330, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2021.05.043..

[73] Chao Ni, Xiang Chen, Fangfang Wu, Yuxiang Shen, Qing Gu, An empirical study on pareto based multi-objective feature selection for software defect prediction, Journal of Systems and Software, Volume 152, 2019, Pages 215-238, ISSN 0164-1212.

[74] Ruchika Malhotra, Shine Kamal, An empirical study to investigate oversampling methods for improving software defect prediction using imbalanced data, Neurocomputing, Volume 343, 2019, Pages 120-140, ISSN 0925-2312.

[75] Yerima, S. Y., Alzaylaee, M. K., & Sezer, S. (2019). Machine learning-based dynamic analysis of Android apps with improved code coverage. EURASIP Journal on Information Security, 2019(1), 1-24.

[76] Siddiqui, M., Wang, M. C., & Lee, J. (2008, February). Data mining methods for malware detection using instruction sequences. In Artificial Intelligence and Applications (pp. 358-363).

[77] N. Milosevic, A. Dehghantanha, K-K. R. Choo. "Machine learning aided Android malware classification", Computers & Electrical Engineering, 61, 266-274, 2017.

[78] A. Pektaş, M. Çavdar, T. Acarman, "Android malware classification by applying online machine learning", (ISCIS 2016) International Symposium on Computer and Information Sciences, Kraków, Poland, 72-80, October 27–28, 2016.

[79] L. Onwuzurike, M. Almeida, E. Mariconti, J. Blackburn, G. Stringhini, E. D. Cristofaro, "A family of droids: Analyzing behavioral model based Android malware detection via static and dynamic analysis", arXiv:1803.03448, https://arxiv.org/abs/1803. 03448, 2018.

[80] Xiangwen Wang, Xianghong Lin, Xiaochao Dang, Supervised learning in spiking neural networks: A review of algorithms and evaluations, Neural Networks, Volume 125, 2020, Pages 258-280, ISSN 0893-6080.

[81] Khalil Moshkbar-Bakhshayesh, Performance study of bayesian regularization based multilayer feed-forward neural network for estimation of the uranium price in comparison with the different supervised learning algorithms, Progress in Nuclear Energy, Volume 127, 2020, 103439, ISSN 0149-1970.

[82] Jiabin Zhang, Hu Su, Wei Zou, Xinyi Gong, Zhengtao Zhang, Fei Shen, CADN: A weakly supervised learning-based category-aware object detection network for surface defect detection, Pattern Recognition, Volume 109, 2021, 107571, ISSN 0031-3203.

[83]] Lu Han, Wenjun Li, Zhi Su, An assertive reasoning method for emergency response management based on knowledge elements C4.5 decision tree, Expert Systems with Applications, Volume 122, 2019, Pages 65-74, ISSN 0957-4174.

[84] X. Wang, C. Zhou, X. Xu, Application of C4.5 decision tree for scholarship evaluations, Procedia Computer Science, Volume 151, 2019, Pages 179-184, ISSN 1877-0509.

[85] Sunanda Gamage, Jagath Samarabandu, Deep learning methods in network intrusion detection: A survey and an objective comparison, Journal of Network and Computer Applications, Volume 169, 2020, 102767, ISSN 1084-8045.

[86] Fatsuma Jauro, Haruna Chiroma, Abdulsalam Y. Gital, Mubarak Almutairi, Shafi'i M. Abdulhamid, Jemal H. Abawajy, Deep learning architectures in emerging cloud computing architectures: Recent development, challenges and next research trend, Applied Soft Computing, Volume 96, 2020, 106582, ISSN 1568-4946.

[87] S. Hernández, Juan L. López, Uncertainty quantification for plant disease detection using Bayesian deep learning, Applied Soft Computing, Volume 96, 2020, 106597, ISSN 1568-4946.

[88] Hui Yin, Yuanhao Gong, Guoping Qiu, Fast and efficient implementation of image filtering using a side window convolutional neural network, Signal Processing, Volume 176, 2020, 107717, ISSN 0165-1684.

[89] Samira Zare, Moosa Ayati, Simultaneous fault diagnosis of wind turbine using multichannel convolutional neural networks, ISA Transactions, 2020, ISSN 0019-0578.

[90] Amin Khatami, Asef Nazari, Abbas Khosravi, Chee Peng Lim, Saeid Nahavandi, A weight perturbation-based regularisation technique for convolutional neural networks and the application in medical imaging, Expert Systems with Applications, Volume 149, 2020, 113196, ISSN 0957-4174,.

[91] Jialin Tang, Qinglang Su, Binghua Su, Simon Fong, Wei Cao, Xueyuan Gong, Parallel ensemble learning of convolutional neural networks and local binary patterns for face recognition, Computer Methods and Programs in Biomedicine, Volume 197, 2020, 105622, ISSN 0169-2607.

[92] Zhiying Fang, Han Feng, Shuo Huang, Ding-Xuan Zhou, Theory of deep convolutional neural networks II: Spherical analysis, Neural Networks, Volume 131, 2020, Pages 154-162, ISSN 0893-6080.

[93] A. Singh, A. V. Agrawal and A. Kanukotla, "A method to improve application launch performance in Android devices," 2016 International Conference on Internet of Things and Applications (IOTA), Pune, 2016, pp. 112-115, doi: 10.1109/IOTA.2016.7562705.

[94] D. W. Aha, D. Aha, M. K. Albert, "Instance-Based learning algorithms", Machine Learning, 6, 37-66, 1991.

[95] Ghorbani, Ali & Lu, Wei & Tavallaee, Mahbod. (2009). Detection Approaches. 10.1007/978-0-387-88771-5_2.

[96] C. Cortes, V. Vapnik, "Support-Vector Networks", Machine Learning, 20, 273-297, 1995.

[97] Y. Siboopimpa, T. Chaipranitan, S. Chiewchanwattana and K. Sunat, ""I am in ASEAN": A learning media application on Android operating system," 2014 Third ICT International Student Project Conference (ICT-ISPC), Nakhon Pathom, 2014, pp. 9-12, DOI: 10.1109/ICT-ISPC.2014.6923206.

[98] Mudzingwa, D., & Agrawal, R. (2012, March). A study of methodologies used in intrusion detection and prevention systems (IDPS). In 2012 Proceedings of IEEE Southeastcon (pp. 1-6). IEEE.

[99] H. Peng, C. Gates, B. Sarma, N. Li, Y. Qi, R. Potharaju, C. Nita-Rotaru, I. Molloy, "Using probabilistic generative models for ranking risks of android apps", (CCS 2012) ACM Conference on Computer and Communications Security, New York, USA, 241–252, October 16-18, 2012.

[100] Y. Zhao et al., "An Experimental Study on Software Aging in Android Operating System," 2015 2nd International Symposium on Dependable Computing and Internet of Things (DCIT), Wuhan, 2015, pp. 148-150, doi: 10.1109/DCIT.2015.19.

[101] O. Sohail and T. Naqash, "Anti-theft cloud application for android operating system (Nougats)," 2018 IEEE International Conference on Applied System Invention (ICASI), Chiba, 2018, pp. 321-324, doi: 10.1109/ICASI.2018.8394598.

[102] Bezobrazov, A. Sachenko, M. Komar and V. Rubanau, "Artificial immune system for Android OS," 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Warsaw, 2015, pp. 403-407, doi: 10.1109/IDAACS.2015.7340767.

[103] Rajalakshmi and N. Anusha, "Sensor based application for malware detection in android OS(Operating System) devices," 2017 International Conference on Information Communication and Embedded Systems (ICICES), Chennai, 2017, pp. 1-6, doi: 10.1109/ICICES.2017.8070722.

[104] A. Utku and İ. A. Doğru, "Malware detection system based on machine learning methods for Android operating systems," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4, doi: 10.1109/SIU.2017.7960231.

[105] P. G. A. A. Putra, E. Sediyono and A. Setiawan, "E-land design of mobile application for land information system using Android-based Google Maps API V2," 2017 International Conference on Innovative and Creative Information Technology (ICITech), Salatiga, 2017, pp. 1-5, doi: 10.1109/INNOCIT.2017.8319145.

[106] M. Kovacev, K. Lazic, N. Jovanovic, A. Stefanovic and I. Petrovic, "Implementation of Disaster recovery agent within Android based set-top box," 2013 IEEE Third International Conference on Consumer Electronics ¿ Berlin (ICCE-Berlin), Berlin, 2013, pp. 1-3, doi: 10.1109/ICCE-Berlin.2013.6697996.

[107] N. T. Cam, T. Nguyen, K. Nguyen, T. Nguyen and V. Pham, "Detect Malware in Android Firmware Based on Distributed Network Environment," 2019 IEEE 19th International Conference on Communication Technology (ICCT), Xi'an, China, 2019, pp. 1566-1570, doi: 10.1109/ICCT46805.2019.8947099.

[108] Jaiswal, Y. Malik and F. Jaafar, "Android gaming malware detection using system call analysis," 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, 2018, pp. 1-5, doi: 10.1109/ISDFS.2018.8355360.

[109] Damodaran, A. Troia, F.D. Visaggio, C.A. Austin, T.H. Stamp, M. (2017). A comparison of static, dynamic, and hybrid analysis for malware detection J Comput Virol Hacking Tech, 13 (1), pp. 1-12

[110] Burnap, P., French, R., Turner, F., Jones, K. (2017). Malware classification using self-organising feature maps and machine activity data. Computers & Security. 73. 10.1016/j.cose.2017.11.016

[111] Yerima, S. Y., Alzaylaee, M. K., & Sezer, S. (2019). Machine learning-based dynamic analysis of Android apps with improved code coverage. EURASIP Journal on Information Security, 2019(1), 1-24.

[112] Alzaylaee, M. K., Yerima, S. Y., & Sezer, S. (2020). DL-Droid: Deep learning based android malware detection using real devices. Computers & Security, 89, 101663.

[113] https://archive.ics.uci.edu/ml/datasets/Detect+Malacious+Executable%28Anti

Virus%29 , Accessed 25-8-2021.

[114] M. Kovacevic, B. Kovacevic, B. Pavlovic and N. Bandic, "Visualization of interactive digital television services content in Android operating system," 2012 20th Telecommunications Forum (TELFOR), Belgrade, 2012, pp. 1357-1360, doi: 10.1109/TELFOR.2012.6419469.

[115] Gupta R, Kewalramani MA, Goel A. "Prediction of Concrete Strength Using Neural-Expert System". Journal of Materials in Civil Engineering, 18(3), 462-466, 2006.

[116] Topcu IB, Sarıdemir M. "Prediction of Compressive Strength of Concrete Containing Fly Ash Using Artificial Neural Networks and Fuzzy Logic". Computational Materials Science, 41(3), 305-311, 2008.

[117] Raghu Prasad BK, Eskandari H, Venkatarama Reddy BV. "Prediction of Compressive Strength of SCC and HPC with High Volume fly ash Using ANN". Construction and Building Materials, 23(1), 117-128, 2009.

[118] Khan MI. "Predicting Properties of High Performance Concrete Containing Composite Cementitious Materials Using Artificial Neural Networks". Automation in Construction, 22, 516-524, 2012.

[119] Atıcı U. "Prediction of the Strength of Mineral Admixture Concrete Using Multivariable Regression Analysis and an Artificial Neural Network". Expert Systems with Applications, 38(8), 9609-9618, 2011.

[120] Nikoo M, Zarfam P, Sayahpour H. "Determination of Compressive Strength of Concrete Using Self Organization Feature Map (SOFM)". Engineering with Computers, 31(1), 113-121, 2015.

[121] Fazel-Zarandi MH, Türksen IB, Sobhani J, Ramezanianpour AA. "Fuzzy Polynomial Neural Networks for Approximation of the Compressive Strength of Concrete". Applied Soft Computing, 8(1), 488-498, 2008.

[122] Yerima, S. Y., Alzaylaee, M. K., & Sezer, S. (2019). Machine learning-based dynamic analysis of Android apps with improved code coverage. EURASIP Journal on Information Security, 2019(1), 1-24.

[123] Siddiqui, M., Wang, M. C., & Lee, J. (2008, February). Data mining methods for malware detection using instruction sequences. In Artificial Intelligence and Applications (pp. 358-363).

[124] N. Milosevic, A. Dehghantanha, K-K. R. Choo. "Machine learning aided Android malware classification", Computers & Electrical Engineering, 61, 266-274, 2017.

[125] A. Pektaş, M. Çavdar, T. Acarman, "Android malware classification by applying online machine learning", (ISCIS 2016) International Symposium on Computer and Information Sciences, Kraków, Poland, 72-80, October 27–28, 2016.

[126] L. Onwuzurike, M. Almeida, E. Mariconti, J. Blackburn, G. Stringhini, E. D. Cristofaro, "A family of droids: Analyzing behavioral model based Android malware detection via static and dynamic analysis", arXiv:1803.03448, https://arxiv.org/abs/1803. 03448, 2018.