ARTIFICIAL INTELLIGENCE-BASED INTRUSION DETECTION MECHANISM FOR MICRO-GRID BUILDINGS

USAMA ALI	**2K18-ELE-67**
MUHAMMAD RUMMAN	**2K19-ELE-80**
SYED MUHAMMAD HAMZA	**2K19-ELE-76**

Supervised By: **Engr. Dr. Saqib Ali**

2023

DEPARTMENT OF ELECTRICAL ENGINEERING
NFC INSTITUTE OF ENGINEERING & TECHNOLOGY
MULTAN PAKISTAN

ARTIFICIAL INTELLIGENCE-BASED INTRUSION DETECTION MECHANISM FOR MICRO-GRID BUILDINGS

NAME OF CANDIDATE

USAMA ALI	**2K18-ELE-67**
MUHAMMAD RUMMAN	**2K19-ELE-80**
SYED MUHAMMAD HAMZA	**2K19-ELE-76**

THIS THESIS SUBMITTED IN FULFILMENT] OF THE REQUIREMENTS FOR THE DEGREE OF ELECTRICAL ENGINEERING BSC (HONS)

Engr. Dr. Saqib Ali

ELECTRICAL ENGINEERING BSC (HONS)

NFC INSTITUTE OF ENGINEERING & TECHNOLOGY
DECLARATION

Names of Candidates:

USAMA ALI	**2K18-ELE-67**
MUHAMMAD RUMMAN	**2K19-ELE-80**
SYED MUHAMMAD HAMZA	**2K19-ELE-76**

Name of Degree: Electrical Engineering

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Field of Study:

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;

(2) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, reference to or reproduction of any copyrighted work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;

(3) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyrighted work;

(4) I hereby assign all and every right in the copyright to this work to the NFC IET, Multan, who henceforth shall be the owner of the copyright in this Work, and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of NFC IET, Multan having been first had and obtained;

(5) I am fully aware that if during making this Work, I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by NFC IET, Multan.

- Candidate's Signature

- Candidate's Signature

- Candidate's Signature

Date________________

Acknowledgments

All praise to **ALLAH Almighty** lord of all the worlds the most beneficent the most merciful owner of the Day of Judgment. We ask Allah to bestow his blessing and salutations of peace upon our noble **Prophet Hazrat Muhammad (PBUH).** We are three members of this group we are enrolled in NFC Institute of Engineering and Technology Multan, Pakistan in this program we want to express our gratitude to all the people who have given their heart whelming full support in making this compilation a magnificent experience.

To ALLAH who is most merciful and beneficent, we thank you for the strength keepsakes us standing and for the hope that keeps us believing that this affiliation would be possible and more interesting. First, we intend to acknowledge every single individual who came across us during our research work. There are scores of people whose names and generosities would remain unrecorded but not unacknowledged; we apologize to all of them.

We are greatly indebted to our teacher and mentor Professor **Engr. Dr. Saqib Ali** who encouraged us to pursue this Project and with profound patience went through the manuscript and provided time to debate many of the issues concerning our topic. We also wanted to thank our family who inspired, encouraged, and fully supported us for every trial that comes our way in giving us not just financially but morally and spiritually.

ARTIFICIAL INTELLIGENCE-BASED INTRUSION DETECTION MECHANISM FOR MICRO-GRID BUILDINGS

This thesis is presented by:

USAMA ALI	**2K18-ELE-67**
MUHAMMAD RUMMAN	**2K19-ELE-80**
SYED MUHAMMAD HAMZA	**2K19-ELE-76**

Under the supervision of their project advisor and approved by the project Examination committee, has been accepted by the NFC Institute of Engineering & Technology, in partial fulfillment of the requirements for the four-year Degree of **B.Sc. Power Systems Engineering.**

____________________ ____________________

(Engr. Dr. Saqib Ali) **(Engr.)**

Supervisor External Examiner

NFC IET, Multan

____________________ ____________________

(Engr. Mughees Riaz) **(Dr. Kamran Liaqat Bhatti)**

Project coordinator Head of Department of Electrical Engineering

DATE: ____________________

NFC INSTITUTE OF ENGINEERING & TECHNOLOGY MULTAN

2022

ABSTRACT

Artificial Intelligence-Based Intrusion Detection Mechanism for Micro-grid Buildings

Smart grid systems have revolutionized traditional power networks, but they also bring along vulnerabilities to various cyber-attacks. These attacks can lead to network crashes and compromise the integrity and confidentiality of smart grid systems. To ensure secure and reliable services in a smart grid environment, it is crucial to have an intrusion detection system (IDS). This proposed Final Year Project (FYP) suggests the development of a feature-based IDS specifically designed for smart grid systems. The performance of this system is evaluated based on accuracy, intrusion detection rate (IDR), and false alarm rate (FAR). This result obtained from the evaluation demonstrate that the random forest and neural network classifiers outperform other classifiers in terms of performance. Remarkably, the achieved FAR is 0.5% on the KDD99 dataset and 0.08% on the NSLKDD dataset. Moreover, both datasets exhibit an IDR and testing accuracy of 99% on average. This suggested feature-based IDS's great efficiency and precision show how well it can protect smart grid systems from different online threats. The IDS can quickly detect and react to potential intrusions due to its low false alarm rate and high intrusion detection rate, thereby minimizing the risk of network breakdowns and intrusions in the context of the smart grid. The project efficiency on both the KDD99 and NSLKDD datasets shows that it is able to adapt to different kind of attacks and network circumstances. The IDS can effectively monitor and identify network traffic, differentiating between legal and malicious activities, by using cutting-edge classifiers like random forest and neural networks. In the end, the suggested feature-based IDS is a useful tool for improving the security posture of smart grid systems, securing important infrastructure, and ensuring a consistent supply of services to end users.

Keywords: Smart grid, Cyber-Attacks, Intrusion Detection System, Algorithm, Datasets.

1 Table of Contents

List of Figures

List of Abbreviations

VPP	Virtual Power Plant
WT	Wind Turbine
μG	Microgrid
DG	Distributed Generation
EMS	Energy Management System
PV	Photovoltaic
RES	Renewable Energy Sources
PSO	Particle Swarm Optimization
KNN	K Nearest Neighbor
IDS	Intrusion Detection System
IoT	Internet of Things
NN	Neutral Network
WIDS	Wireless Intrusion Detection System
DT	Decision Tree
AI	Artificial Intelligence
FAR	False Alarm Rate
DR	Detection Rate

Chapter 1
Introduction

1 Introduction

1.1 Background

Despite being commonly employed in many IDS, it often raises false-positive alerts. The preconfigured signatures from the previously discovered malware saved in a database are used in pattern matching by the signature-based IDS. This results in a small number of False Positive (FP) alerts, but it also permits fresh assaults to go undetected.1-2

As a result, a method must be created to lower false alarms in previously specified signatures and raise the detection rate (DR) for newer (also known as zero-day malware) assaults.

Figure 1 depicts the interaction between energy generating facilities, distribution hubs, and other entities such as enterprises, smart buildings, households, and so on. The smart grid substantially facilitates the successful distribution of the appropriate amount of power to these many enterprises. To make the power distribution process more flexible, the smart grid employs a range of artificial intelligence (AI) algorithms. Flexibility is a concern due to the need for dynamic power from diverse businesses.

In order to identify and categorize network intrusions using signature-based IDS while lowering the false alarm rate (FAR), this study utilizes an optimized feature selection approach. Real-time patterns and traffic often include a high-dimensional speed of characteristics. As a result, feature selection is often used to minimize the complexity of a dataset, make it simpler to find significant characteristics, and improve forecast accuracy. Real-time traffic may be made clear of noise and unnecessary features with the use of an effective feature selection.3-5 One method for selecting features that are often utilized is called particle swarm optimization (PSO).6-8 It is a popular option for academics because of its simple-to-encode features, support for global searching, low computing power requirements, few parameters, and simplicity of usage.9-11 As a result, we also employed PSO in our trials to choose features.

Various forms of assaults have been identified and detected using machine-learning techniques. Several machine-learning techniques have been incorporated into this FYP to categorize network packets as harmful or innocent.

This study's new contribution is the adjustment of PSO algorithm weights, which enables our weighted Particle Swarm Optimizer to choose the best features from datasets and generate high

DR, high accuracy, and enhanced FAR with those chosen features. NSLKDD22 and KDD99 are the two datasets utilized in this study.12A few prepossessing approaches are used after choosing the datasets. The min-max normalization approach is used to scale the data and normalize datasets. Data encoding transforms nominal and numeric values after data normalization since machine learning only works with numeric data. The accuracy, intrusion DR, and FAR of the proposed system are examined. The results show a neural network (NN) and a random forest.

Classifiers have worked more effectively. On the KDD99 dataset, we obtained a 0.5% FAR, and on the NSLKDD dataset, a 0.08% FAR. For both datasets, the average DR and testing accuracy are 99%.

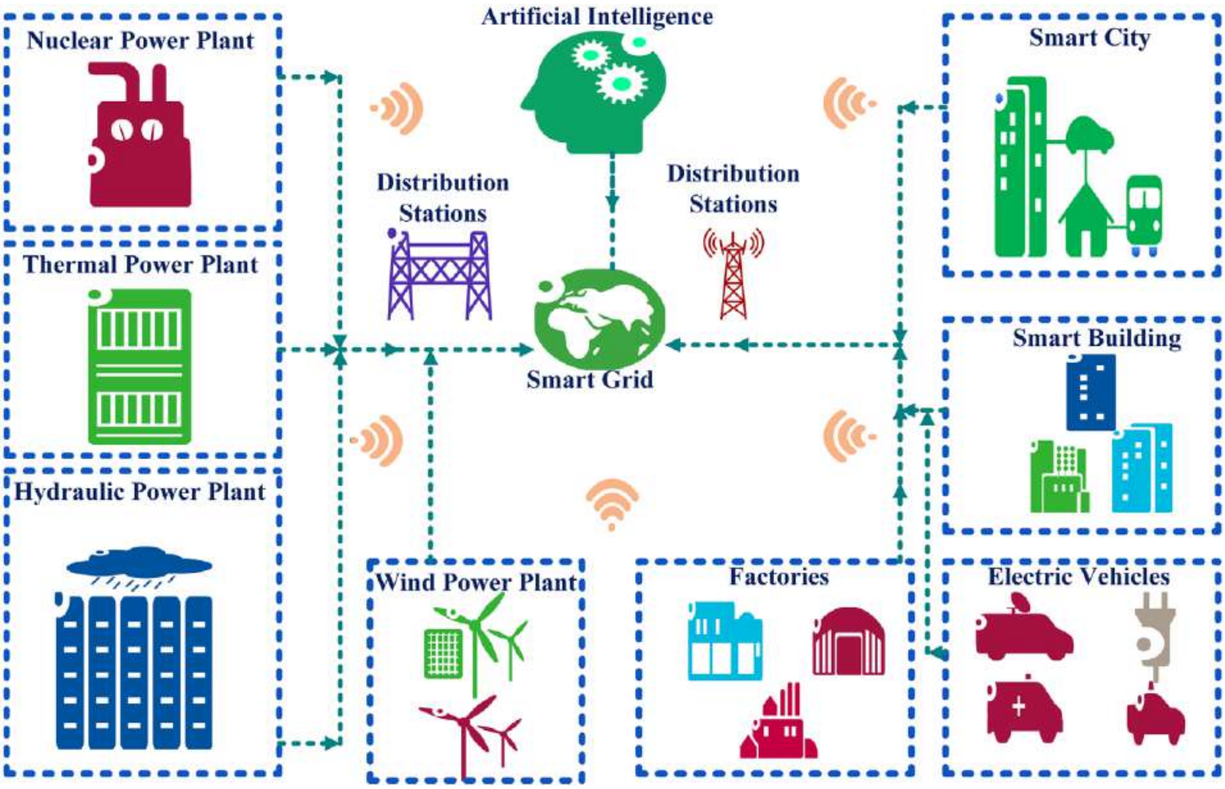


Figure 1.1 Smart Grid System Illustration

Problem Statement:

In this section, we contribute to the proposal.

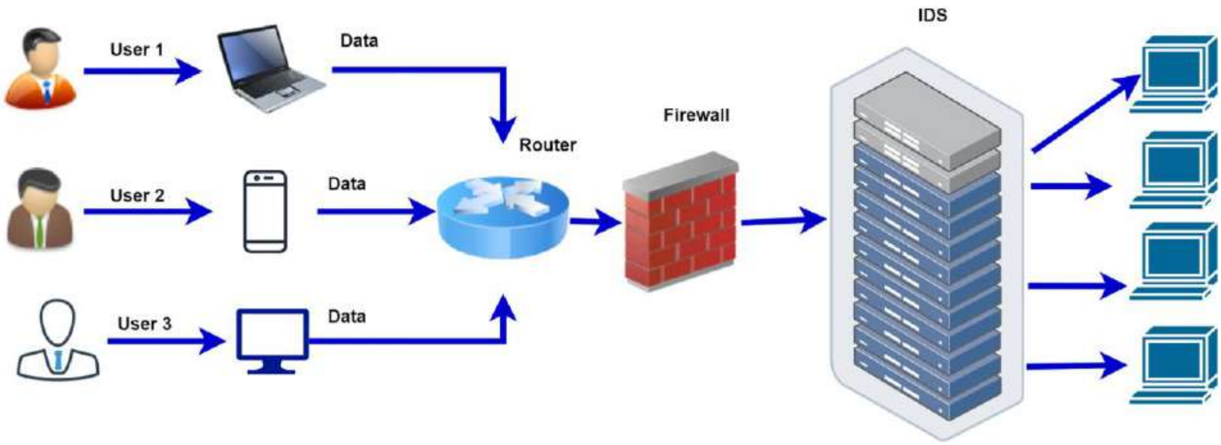


Figure 1.2 Intrusion Detection System Working

for detecting intrusions. The performance of the multiple criterion linear programming classifier may be enhanced using PSO. PSO offers a variety of optimum features for different datasets, including KDDCUP99.13 We have researched many methods for feature selection and examined the systems that may automatically categorize a packet into normal or abnormality groups. The following criteria were used to evaluate the literature that is currently accessible. Table 1 shows that, although offering the best feature set possible, PCA affects training effectiveness.14 Micro Grid and Gini-index have the drawback of producing biased findings for non-numeric data.15 Similar to fuzzy logic, optimum solutions need additional assurance from genetic algorithms.

16 As a result, more durable solutions are needed, as opposed to the genetic algorithm, which has a slow convergence rate and is dependent on the population utilized. These solutions must also offer optimum solutions and have a quick convergence rate.

17 To strengthen the system, we employed weighted PSO for feature optimization. Regardless of the dataset, PSO will automatically offer a set of optimum features. On various datasets, the aforementioned feature selection techniques either increased DR, accuracy, or FAR, but not all three metrics at once. These feature selection techniques rely on data.

As a result, a more ideal approach is needed, one that can address the aforementioned issues and perform well independent of the dataset. This is why we suggested weighted PSO in this study, which had promising outcomes in comparison to prior investigations.

Contributions

1.2 Problem Statement:

- The accuracy of information is critical to the optimal operation of EMS.
- Communication Links Usually have Cyber Security Issues.
- Create an efficient method for detecting and preventing internal and external intrusion attempts.
- Developed Intrusion Detection & Prevention System against Cyber-Attacks using Linux Operating System-based Smooth-sec Software (Continuously Monitor & Record the Traffic).

1.3 Motivation for Research:

Aim & Objectives

- Proposed Cybersecurity Mechanism for Energy Management Control in Microgrid Buildings.
- Provides software-based Simulations for Intrusion Detection & Prevention for Mini-And Microgrid Buildings.

1.4 Case Studies

- **CASE-1:** Without Proposed Intrusion Detection & Prevention Technique.
- **CASE-2:** With Proposed Intrusion Detection & Prevention Technique.

1.5 Proposed Monitoring System

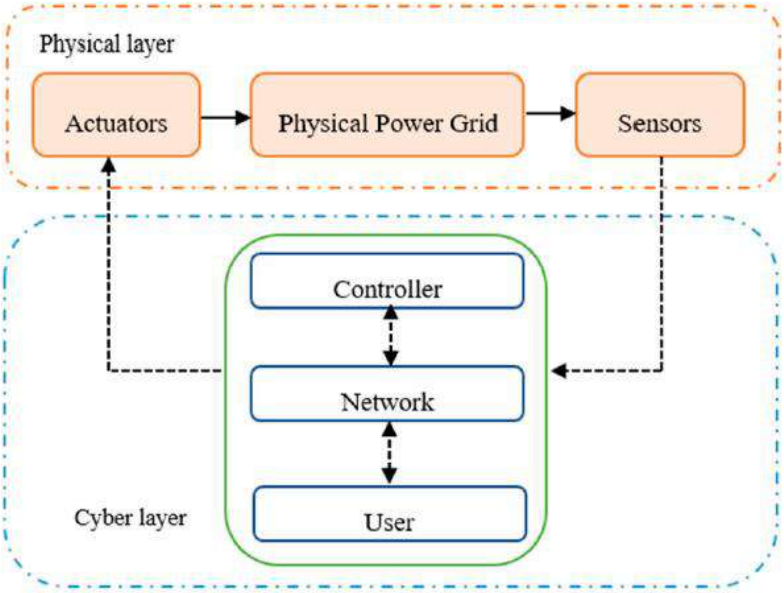


Figure 1.3: Proposed Monitoring System

Disclosure (confidentiality) attacks

Deception (integrity) attacks

Disruption (availability) attacks

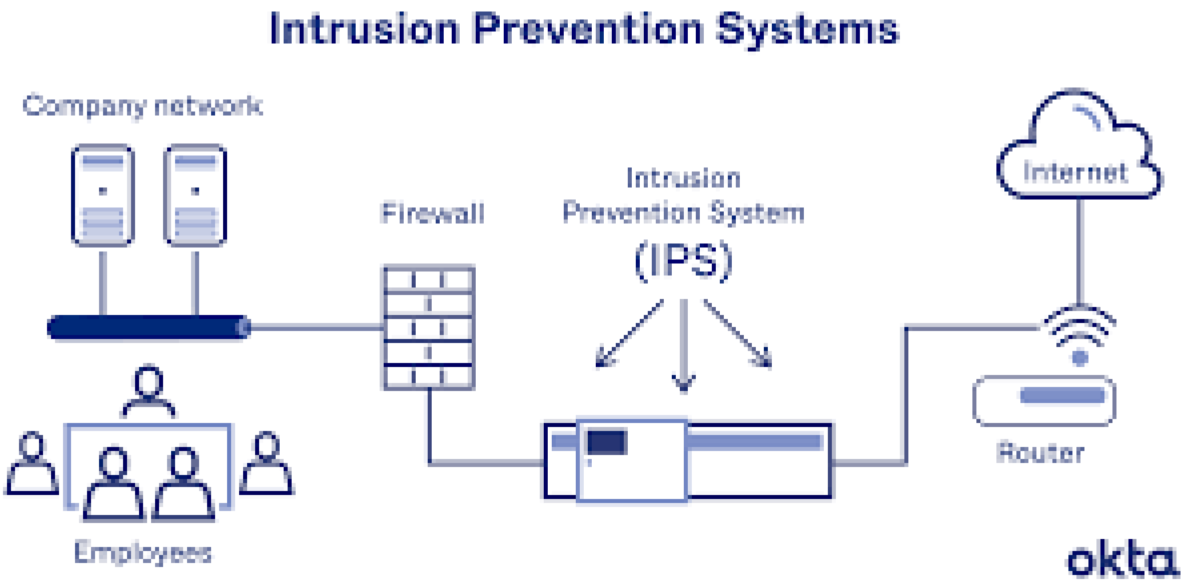


Figure 1.4: Intrusion Prevention Systems

1.6 Intrusion Detection System:

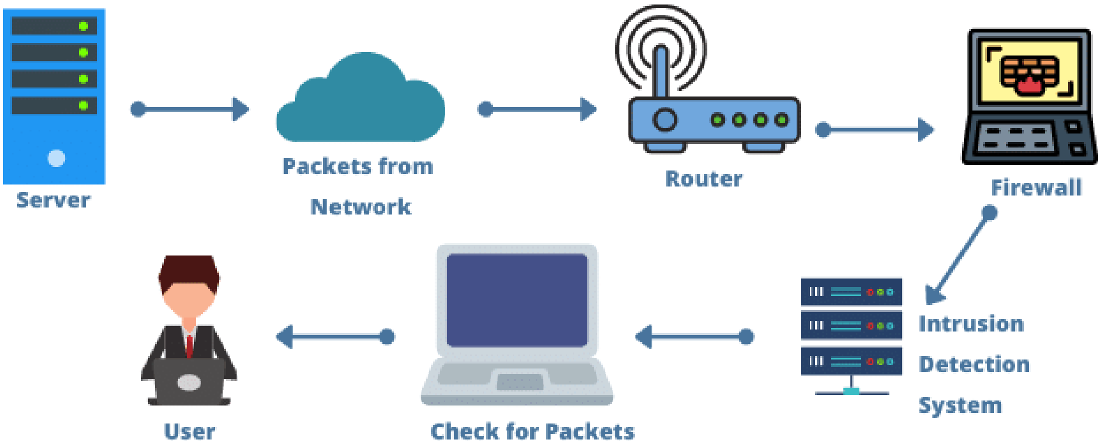


Figure 1.5: Intrusion Detection System

1.7 Flow Chart

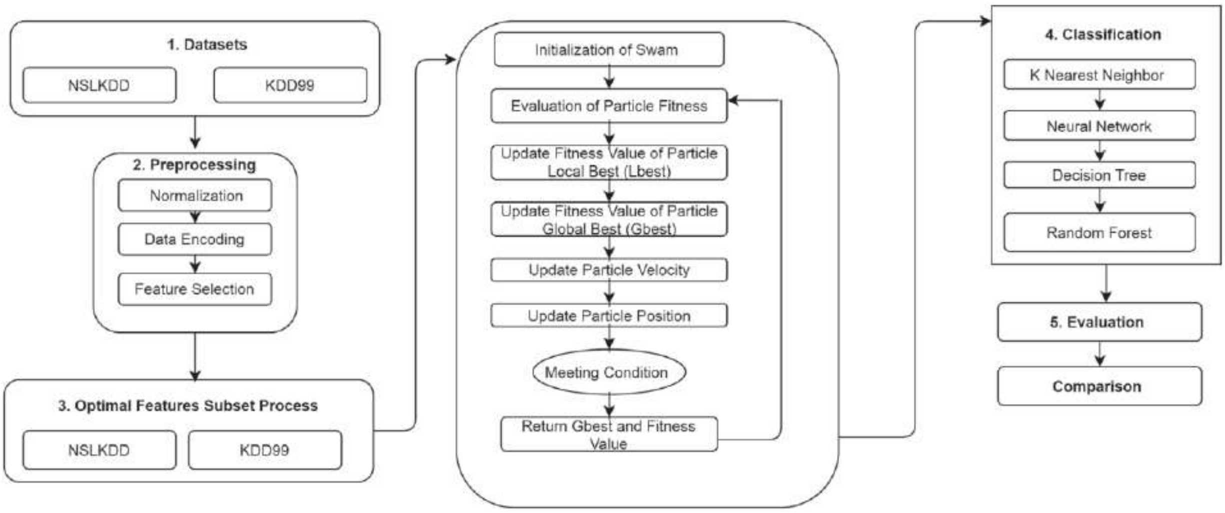


Figure 1.6: Flow Chart

1.8 Benefits of our Systems:

1. Thus project is used for the Security purpose of Grid-Station because there are Cyber Security threats.

2. This will increase Efficiency of the Station.

3. By algorithms we can detect abnormal patterns or behaviors very fast and effectively.

4.AI-based IDS have the ability to adapt and learn from new patterns and developing attack methods. Over time, they can improve their detection abilities and regularly update their knowledge base, strengthening the micro-grid system's overall security posture.

5. Artificial intelligence (AI) can automate the intrusion detection process, saving resources and labor. A proactive response to security issues is made possible by the system's ability to analyses and deal with data in real-time, freeing up resources to concentrate on other important tasks.

6. At the start of implementation of AI-based intrusion detection techniques can require an investment in resources and technology, this may result in cost savings in the long term.

7. Systems can operate in temperatures as low as -40˚C.

8. Simulate in Software

1.9 Thesis Organization

Chapter 1 introduction to the proposed work, background, problem statement, scope and aims of the work, and thesis organization.

Chapter 2 is a wide literature survey of hybrid power system concepts and background, renewable power sources, and stability analysis - an overview.

Chapter 3 presents the problem formulation, research methodology, and proposed research solution.

Chapter 4 presents the simulation and results, and conclusion.

1.10 Summary

In this chapter, we discuss the voltage stability problem, motivation for research, problem statement, current situation of the research topic, scope, and objective, and in the last, thesis organization. The problem statement was also clearly defined. Different installation techniques were compared, and a suitable technique is presented for this thesis. We also the scope of this thesis is presented in this thesis.

CHAPTER 2

Literature Review

2 Literature Review

2.1 Artificial Intelligence, or AI

Artificial intelligence (AI) describes computing technologies that can do certain tasks in place of human intellect. The deployment of clean, renewable energy sources in power systems throughout the globe might be facilitated and accelerated by artificial intelligence (AI), which has the potential to reduce energy waste, decrease energy expenses, and reduce energy costs. AI can also help in planning, running, and controlling electricity systems.

2.2 Intrusion Detection System

An Intrusion Detection System is a monitoring system that detects odd behavior and delivers alerts when it detects it? Based on these notifications, a security operations center (SOC) analyst or incident responder may analyze the situation and take the appropriate actions to eliminate the hazard.

2.3 Micro-Grid

A microgrid is a small power grid that may run by itself or in cooperation with other tiny power grids. Distributed, scattered, decentralized, district, or embedded energy generation are terms used to describe the use of microgrids. Over the next 20 years, it is predicted that the demand for energy would rise by 30% to 40%. The current power networks need to create more energy since they are extremely old, getting more crowded, and unreliable. A smart grid manages energy supply and consumption in an organized, analytical manner. Two methods are used by the smart grid to monitor and control the flow of electricity. Depending on the quantity of electricity consumed, users also had the choice to utilize an optimized algorithm to purchase the lowest energy at a certain time. Bidirectional communication between energy providers and their customers is made possible by the smart grid. The transition to the smart grid from the present electrical system necessitates fresh financing, ensuring the returned tremendous value. Facilities that are dependable, stable, affordable, effective, efficient, ecologically sustainable, and healthy are required for the smart grid. The smart grid has the following seven key characteristics: it enables active customer participation; it manages all production and storage options; it develops new products, utilities, and markets; it offers the best digital economy with power reliability; it uses energy, optimizes it; and it is reliable; it has the capacity to self-heal; and it is resistant to both physical and digital attacks. The creation of smart grids necessitates the fusion of several applications and technology. Customer acceptance, advanced distribution operations, advanced transmission operations, and advanced asset management are the four milestones for the smart grid. The smart grid enhances the monitoring and management of the power system coordinates by enhancing network-wide dependability and dynamic performance. For an electric power system to operate automatically, cyber protection is crucial.

On the DARPA1998 dataset, one of the first efforts to obtain a high DDR and a low FAR has been made.24 In this FYP, the authors have chosen features and NNs for classification using principal component analysis (PCA). Even though PCA offers the best feature set, it reduces training effectiveness while still producing accurate results.25 The algorithm of the feature vitality-based reduction approach is another technique for selecting the best features.19 The experiment employed the Naive Bayes classifier and 41 features from the NSLKDD dataset. Multiple feature selection strategies have been used in certain investigations. There are four feature selection strategies that were employed by Hee-Su et al. These methods include attribute ratio, information gain (IG), correlation-based feature selection, and gain ratio. The J48 classifier was used to classify the 22 characteristics that were chosen from the NSLKDD dataset. A support vector machine (SVM) classifier has been used to choose the best features from the KDDCUP99 dataset using the genetic principal component, 27 methods. Manekar et al. 28 employed parameters turning with PSO with an SVM classifier in Otoelop an intelligent IDS utilizing the NSLKDD dataset (Figure 2). The intrusion feature selection algorithm-based PSO is another kind of PSO.29,30, as opposed to a single numeric number, that shows velocity and location in intervals. On the KDD99 dataset, the method has been used together with random-based PSO.

The proper remedies are required for rising fuel prices, ecological system imbalances, and an abundance of energy resources.

Distributed energy resources (DERs) and demand response (DR) tactics are two possibilities that might work well to address these issues. The DERs are located in a microgrid (G) and act as the onsite generation for the linked loads while also employing net metering to sell any surplus electricity to the main grid [1]. Buildings may actively contribute to lowering the cost, emissions, and network burden associated with energy usage.

Demand response programs affect customer energy consumption patterns to transfer load from peak to off-peak hours for lower energy costs and lessening of overload.

These two options combined in microgrids might prove successful in solving issues with the energy distribution system.

2.4 Literature Survey on the energy management system (EMS)

A literature survey on the energy management system (EMS) of Micro Grids is conducted as follows:

By using multicarrier energy resources, the authors compared the single-cored and multi-cored processors for the home energy management system (HEMS) in terms of effective cost and performance. The system includes solar photovoltaics, a battery bank, a power source, and a load. Communication with the room controller and continual load evaluation are the responsibilities assigned to the first core [1].

The second core's functions include managing energy consumption, estimating battery state of charge (SoC), and encrypting communication data. Results reveal that employing single and dual-core processors, respectively, reduces the processor's execution time from 1.88 m/s to 1.42 m/s and raises its speed from 1 m/s to 1.32 m/s [2].

To reduce energy prices and load variations for residential users, the authors suggested a multi-objective EMS. The structure is equipped with a utility grid, a battery energy storage system (BESS), as well as important and movable appliances. Mixed-integer linear programming is used in MATLAB to address the issue for four buildings. According to the findings, clients may save a total of $565.75 per year and return their BESS investment in three years [3].

The authors saw a Palmdale, California, office building as a sizable Micro Grid that included a PV system, electrochemical energy storage, a life cycle analysis of energy, and an electrical network. The model is presented as a linear programming problem using mixed integers. According to simulations, the suggested framework saves 112,410 dollars in energy costs throughout its 20-year life cycle [4].

The authors suggested a household Micro Grid with solar PV, wind turbines, diesel generators, and battery energy storage that is linked to the grid. The created model tackles economic and environmental challenges by minimizing yearly energy consumption costs and emissions by controlling responsive loads (lights, shiftable loads, cooling, and heating) appropriately. This model was solved in MATLAB using mixed-integer linear programming. Results indicate that emissions and overall yearly cost decreased by 30.5% and 23.7%, respectively [5].

Amrr et al.'s (2018) suggestion for a residential EMS to handle building-owned components in the best possible way. Lead-acid batteries, load, national grids, and solar PV panels are all components of a microgrid. Mixed integer linear programming (MILP) was used in MATLAB to find a solution to the issue. According to analysis, the suggested EMS module chooses a cheap source from those offered to cheaply and dependably supply a building in either grid-connected or islanded mode. Later, the authors created the control in order to confirm the findings using real load patterns [6].

To examine the nature of the customer-utility interaction for residential consumers, the authors presented a two-step technique. Bi-level macro energy hub management was suggested in step one to benefit both consumers and the utility. The control module reduces cost, emissions, and fluctuations in network demand on the utility network at the customer level. The suggested model is solved using the flower pollination algorithm (FPA) and particle swarm optimization (PSO) in MATLAB, demonstrating FPA's superiority over PSO and MILP. Using bivariate regression analysis, the cost of energy consumption and fluctuations in network load are correlated in the second stage. A nonlinear link between a new consumer and utilities is shown. Biogas is added [7] to overcome this non-linearity for simultaneously achieving bi-lateral stacks.

To lower energy expenditures and emissions, the authors suggested a methodology of predictive control for the home Micro Grid. Electro-thermal storages, fuel cells (FCs), combustion engines, steam engines, Stirling engines, and micro combined heat and power (CHP) are all components of a microgrid. Using MATLAB and MILP, the suggested linear optimization framework is resolved. The heat produced by fossil fuel-based engines is caught, stored, and used for heating and cooling to increase energy efficiency. A building's load may be supplied by wind energy in the range of 36.93% to 61.84%, resulting in a drop in emissions of 833 tons to 783 tons [8].

The authors suggested an EMS hardware testbed for a home microgrid with two distribution buses (AC and DC), 8 kW of solar PV, 5 kW of wind turbines, 8.5 kW of micro-CHP, 2 kW of EV chargers, and a 110-kWh battery. EVs, lights, and elevators are categorized as considered loads. Using the general algebraic modeling software (GAMS), the linear framework is solved. Results indicate a 28% reduction in operating costs [9].

The authors suggested a DR approach for business buildings in order to plan HVAC, electric water heaters, and plug-in electric cars in the best possible way. The structure features a heating, ventilation, and air conditioning system, an electric water heater, a battery, an electric vehicle, solar panels, and a power grid. The goal is to reduce overall energy expenses while increasing consumer comfort. In order to create scenarios for random solar irradiation, the Monte Carlo approach is utilized. Results show that simply forgoing 20% of energy costs, home comfort levels rise from 40% to 100% [10].

For a commercial facility in Tshwane, South Africa, the authors presented a model predictive-based energy management and control system. A solar system, utility grid, and battery storage are all present in the structure. In MATLAB, the suggested method is resolved. The cost of imported energy is shown to have decreased by 46% as a consequence [11].

The authors suggested that Hong Kong City implement a hybrid renewable energy-based commercial Micro Grid that includes both WT and solar PV. Two scenarios are used to test the suggested concept. 2) National grid solely, and 1) onsite DG with the national grid. Through the use of the program RET-Screen, various instances were compared. As a consequence, the energy payback period for installing onsite DGs in a commercial building is 9.2 years, which is 6.4 times less than the payback period for the second scenario, which calls for expanding the national grid infrastructure. Additionally, in the first and second scenarios, CO2 emissions drop from 24.091 kg per day to 3.077 kg per day [12].

A microgrid designed by the authors includes energy storage, solar PV, wind, small hydro, biomass, geothermal, FC, and load. The authors reduced the payback time, net present value, and internal rate of return by using a decision support model as a strategy for a solution and mapping the issue in MATLAB. Simulations indicate a 42% and 15% reduction in energy expenditures and emissions, respectively [13].

The authors suggested a hybrid Micro Grid testbed with an 80-kWh lithium battery and 15.5 kW of solar power to provide a 145-kVA load for a commercial building at Griffith University's Nathan Campus in Australia. The Micro Grid also includes WTs, single-phase and three-phase stations with ten kVA and thirty kVA outputs for voltage control, an EV lot, two AC buses (200V-300V, 700V-800V), and a DC bus (300V-500V) in addition to these sources. In MATLAB/Simulink software, a rule-based scheduling technique was created to address the issue. The outcome reveals that the proposed arrangement adopts energy storage and renewable energy to reliably and sustainably serve the linked load. Sitcoms' existence guarantees that voltage regulation stays within acceptable bounds [14].

In order to increase interior comfort and save energy expenses for an office block at the University of Zagreb in Croatia, the authors presented a framework. A 1500 W solar PV array, a 2000 W WT, a 200 Ah lead-acid battery, a 93.7 F ultracapacitor, and 500 W FCs make up the proposed Micro Grid. The MATLAB issue was resolved using the model predictive control approach. Results reveal an average cost reduction of 115% while retaining the required degree of comfort [15].

While taking into account several choices, including diesel generators, BESS, and solar PV power for a glass plant in India, the authors offered a collection of commercial and Micro Gris. In HOMER Pro, the suggested issue is resolved. Results show a 45% and 19% reduction in fuel and energy consumption expenses, respectively [16].

For the purpose of lowering energy costs, the authors suggested a multi-agent optimization framework to quantify the flexible load in industrial facilities used for smelting aluminum and cement. The Nordic Electricity Market's Danish sector has been used to test the suggested strategy. Stochastic programming is implemented in GAMS, and the output is imported into MATLAB. According to the calculations, the cost of energy is reduced by 18% and 34%, respectively, when renewable energy is added to the cement and smelting sectors [17].

To assess the long-term effects of various energy sources in the Micro Grid, the authors suggested running a time series simulation for 10 years with the HOMER Pro program. The national grid, DGs, and WT make up the suggested model, which examines financial gains and carbon emission reduction. The suggested emission-based approach is more cost-effective, with a total energy cost of $6.5744 107 dollars. However, without the emission, the price jumps to 6.6827 107$. According to data, there are 10,946,355 kg of carbon emissions per year [18].

The authors described a Micro Grid for an Irish manufacturing site that was grid-connected and included WTs, batteries, and a CHP unit. The suggested model reduces expenses and emissions. The Levenberg-Marquardt training approach was used to solve the proposed framework in MATLAB. The mean absolute percentage error and coefficient of determination were used to assess the performance and accuracy between the actual and estimated outcomes. The suggested EMS approach reduces peak load by 8.3%, according to the results. Costs were cut by 73% by using the CHP unit, WT, and battery. In a similar manner, WT output cut carbon emissions by 88%, while WT and CHP units cut emissions by 97% [19].

The authors described a polymer processing facility complete with a CCHP unit, solar PV panels, a gas boiler, an electric boiler, an absorption chiller, a heat exchanger, a thermal oil system, a blown film extruder, a dryer, an air compressor, a compression chiller, and water storage. GAMS has been used to resolve the suggested linear optimization framework. After incorporating RERs, the results indicate that the energy consumption is reduced by 23.19% [20].

In Tehran, Iran's Shad-Abad industrial zone, the authors suggested a micro grid. The microgrid consists of WT, batteries, PV panels, fuel cells, diesel generators, and electrolyzes for producing hydrogen. The network's stability will be improved, and the costs associated with energy and emissions will be reduced. In HOMER Pro, the suggested issue has been resolved. Results indicate a \$1.87M net cost reduction and a 90000 kg/year reduction in CO2 emissions, respectively [21].

The authors provided an EMS for the Nordic power market's mechanical pulp producing process. The Micro Grid uses CCHP powered by natural gas, district heating, electricity, and heat as output carriers. The developed technique creates an operational strategy for pulp manufacturing that is optimized. A CPLEX solver has resolved the framework. Results show that industrial demand-side management (DSM) provides the flexibility needed to maintain network stability and reduce energy costs [22].

For the purpose of solving a DR-based two-stage energy management system for the Micro Grid in Florida, the authors created a multi-objective genetic algorithm-based optimization approach. Stage 1's optimizer adjusts the load that can be shifted, while stage 2's control continually monitors the load that can be controlled. With Pareto optimum settings, the simulation revealed decreased utility costs of 2% to 6% [23].

The "Gold Wind Smart Microgrid System" testbed for a Micro-Grid was suggested by the authors for Beijing, China. A 2500 kW WT, 480-kW solar PV system, 500 kW diesel generator, 1600 kW battery, and a 2000 kW supercapacitor are included in the proposed concept. A 10 kV bus connects each of these parts to the national grid. The suggested EMS's input parameters were loading demand, wind, PV irradiance, and energy price. Utilizing MATLAB regrouping particle swarm optimization (Reg PSO), simulations were run with the goal of minimizing fuel consumption, operating and maintenance expenses, energy purchase prices, and maximizing income. According to the findings, the Reg PSO produces the best overall solution in 2.1152 seconds as opposed to 16.3456 seconds for the genetic algorithm-based method. The cost of energy from wind, solar, diesel generators, and batteries, respectively, was 0.3767 cents per kWh, 0.2169 cents per kWh, 0.5767 cents per kWh, and 0.003 cents per kWh [24].

To guarantee grid-connected and islanded mode stability, the authors suggested voltage and frequency management for the microgrid. The industrial facility may isolate and run in islanded mode during utility grid outages, utilizing its on-site generating resources, such as DGs and batteries, to provide the whole load. However, the EMS module has a demand curtailment mechanism to maintain frequency and voltage if the load exceeds the capacity of on-site generation. MATLAB was used to resolve the mixed-integer nonlinear optimization issue. The outcome demonstrates that the proposed technique maintains stability in both normal and abnormal circumstances [25].

2.5 Approaches to Modelling Uncertainty and Risk Aversion

The literature on risk aversion and uncertainty modeling is reviewed in this part.

The two-stage hybrid optimization approach for scheduling power consumption in homes with DERs and storage was given by the authors. In order to reduce costs, a genetic algorithm was used to build a HEMS consisting of solar PV, WTs, EVs, BESS, and controllable and uncontrolled appliances. In the second stage, a neighborhood EMS flattens the load curve by reducing the peak load demand and its oscillations. The Bayesian Game approach was used to resolve the mathematical conundrum. In terms of power conservation, the results revealed a 25% rise in inefficiency and an 8.7% drop in cost [26].

In order to maximize their revenue through the use of CPLEX solver, the authors suggested a stochastic bi-level decision-making bidding strategy with the conflict between wind power producers and aggregators. Additionally, CVaR is taken into account to manage volatile market pricing, wind energy, and EV availability. In order to convert the stochastic bi-level issue into a linear stochastic single-level problem, the Karush-Kuhn-Tucker optimum conditions and duality theory are used. According to the findings, CVR rises to 33.81%, 40.79%, and 46.99%, respectively, under 0%, 60%, and 100% DR individuals [27].

The authors presented a linear bi-level risk-averse macro energy hub control for a collection of big commercial buildings Micro Grid under the uncertainties of unexpected solar irradiance and unscheduled electric and NG network disruptions. Internal combustion engines (ICEs), EVs, BESS, MT, and FC are all present in the structure. In order to achieve risk aversion, CVaR is introduced to the objective function and the MATLAB flower pollination algorithm (FPA) is used to solve the problem. According to the findings, for the fixed battery and EV lot, respectively, the charge keeping capability improved by 21.03% and 24.10% under the risk-averse mode [28].

A mathematical model for a stand-alone Micro Grid was suggested by the authors utilizing a risk-based, two-stage stochastic optimization method. Both the erratic supply of solar and wind energy as well as load demand are taken into account. CVaR is modeled to lessen the negative impact of RERs. Software called Wien Automatic System Planning (WASP) resolves the mathematical issue. To stop the load during load spikes and/or carrier disruptions, a price-based DR system is used. The incorporation of CVaR enhances the reserve energy with the growth in energy costs, according to the results. By strategically interrupting building loads, reserve energy has been saved for usage in the event of an unexpected increase in energy demand or a carrier outage [29].

In the presence of a plug-in EV lot, the authors presented an energy management control of a Micro Grid consisting of 10 buses and 7 manufacturing units to reduce operating costs by employing a receding horizon framework. The chance constraint method and convex relaxation techniques are used to address the uncertain availability of plug-in EVs lots. Under both deterministic and uncertain circumstances, the proposed technique is solved. Results indicate that, with a 30% EV penetration, energy costs are comparable for both techniques. As EV penetration rises over this limit, the network becomes unstable [30].

The writers gave an industrial park with a collection of Micro Grids some thought. Energy price and plant products have been included as unknown variables using robust optimization approaches. A Micro Grid is made up of DGs. Goal programming methods have been used to resolve the presented issue. Using a game-theoretic approach, the developed central controller lowers the energy costs of individual industrial buildings [31].

The authors suggested an ideal size method for a home energy hub with energy cost reduction as the aim function, including combined heat and power units, gas boilers, PV panels, and storage. The cost of energy includes construction expenditures as well as operating and maintenance costs. Both deterministic and random solar irradiance were used to verify the developed system. Monte Carlo simulations are used for scenario development and reduction. A CONOPT solver in the General Algebraic Modelling System (GAMS) is used to resolve the suggested model. According to the findings, the cost rises from 48083 to 48115 dollars with and without solar irradiance uncertainty, respectively [32].

For a commercial Micro Grid, the authors recommended a two-stage EMS with plug-in EVs as a DR tool to lower peak demand, economic efficiency, and environmental issues. The framework intends to lower energy costs. One MW of baseload electricity, 100 EVs, and 1.5 MW of wind power make up the Micro Grid. An autoregressive integrated moving average model [12] was used to account for the uncertainty associated with wind speed, load demand, and electricity pricing. To handle uncertainties, the CVaR was implemented into the model predictive control. The outcome shows that CVaR's addition improves the battery's capacity to store energy. To prevent Loll, retained energy serves as a backup during blackouts of wind and solar power. Additionally, operating costs were 206.6003$ and 128.7805$, respectively, with a 37.67% improvement with the suggested control [33]. A cost-risk tradeoff model was developed by the authors for a home Micro Grid with integrated heat and electricity production. A two-stage stochastic programming approach is used to represent the CVaR framework. Wind energy, solar irradiance, load demand, and electricity price uncertainties are all mitigated by conditional value at risk. Results indicate that adding CVaR causes costs to increase. However, a risk aversion component makes the system more stable and ensures a constant flow of energy [34].

To simulate the uncertainties of DERs, such as solar PV arrays, WTs, microturbines (MT), fuel cells (FC), and battery storage, the authors took into account risk-based, two-stage stochastic linear as well as nonlinear optimization programming. For issue analysis, the framework uses nonlinear economic models, conditional value at risk (CVaR), and incentive-based DR programs. SCENRED was used to produce 10,000 scenarios at first, which were then quickly condensed down to 15. A 24h time frame was used for the analysis of a collection of 60 residential dwellings. Results indicate a decrease in consumption costs and peak load [35].

The authors put up a paradigm for energy procurement that would allow for cost reduction in the face of volatile market pricing, solar radiation, and wind energy. The objective function measures a risk factor in terms of CVaR. The developed method is solved using GAMS in CPLEX. The cost of energy consumption with and without price-based DR is $36,945 and $40,253, respectively, according to the results [36].

According to a review of the literature, EMS modules for residential and commercial buildings have previously been developed to achieve goals including lowering energy costs and emissions. Similar to this, small- and medium-scale Micro Grid EMS architectures have also been developed. However, EMS has not yet been developed and verified under deterministic and stochastic situations. EMS is capable of properly managing bidirectional energy transactions with the national grid by scheduling the sources, loads, and storages for large-scale Micro Grid. Therefore, this FYP designs a big Micro Grid's energy management control. This FYP includes stochastic characteristics including random solar irradiation, uncertain electric network presence, and random natural gas (NG). These uncertainties could be brought on by small variations in the weather, natural catastrophes including earthquakes, floods, tornadoes, and tsunamis, terrorist attacks, system errors, energy demand shortages, overloading, load shedding, human error, and cybersecurity breaches. Uncertainty may make the likelihood of loss of load (LoL) during energy transport disruptions more likely. In order to change the control from a risk-neutral to a risk-averse condition, CVaR is introduced to the goal function [37].

In FPA, the developed framework is resolved. The flower pollination method outperforms particle swarm optimization (PSO), fuzzy logic, cuckoo search, genetic algorithm differential evolution PSO, and random search algorithm [FYP] in terms of convergence speed, computing complexity, and procedural complexity. In contrast to deterministic approaches [building solution method], metaheuristic algorithms solve large-scale issues but have a long execution time and tend to converge on local optimal solutions. On the other hand, deterministic methods quickly solve the issue without being stuck in the local optimum. However, they struggle to solve significant issues. A tri-layered hybrid modified FPA-Mix integer linear programming (MILP) approach has been suggested to overcome these problems.

2.6 Internet of Things (IoT)

The Internet of Things (IoT) is a smart network that allows everything to connect to the Internet and share information using specified protocols. [1]. Allowing the user to access everything remotely at any time [2]. IoT is built on the use of smart sensors to link physical items together through a standard platform. The aim of these intelligent sensors is to operate with little to no physical contact with people [3]. Using addressing methods, physical things and smart gadgets communicate and collaborate with one another. When "Smart" is used as a prefix to a variety of physical products, it denotes IoT use. Smart surroundings, smart watches, smart homes, smart TVs, and smart water are a few examples [4]. IoT is sweeping the globe, and by 2020, it's expected that there will be six times as many smart devices as people. IoT [5] is the driving force behind all technology advancement and is recognized for its rapid expansion.

IoT devices are becoming more vulnerable to security flaws as they evolve and innovate, which is critical for IoT devices. These "smart" gadgets are susceptible to a variety of malicious assaults since they may be simply utilized from any location, any network, and at any distance. The networks and objects are not secured against numerous invasions since they are accessible from anywhere over the Internet. Keeping the IoT network secure is thus the most important responsibility for academics.

Here there are some important security points:

1)Privacy of data: Due to its modifiability, the data sent between the transmitting and receiving nodes is vulnerable to hacking and will lose its secrecy. As a result, protecting IoT data is crucial.[6]

2) Integrity: Information or data shouldn't be changed while being sent in any way. Throughout the whole transmission procedure, the message should preserve its correctness. Integrity has to be guaranteed in the IoT context, according to the work [6].

3) Availability: The work [6] demonstrates that the availability of resources is still essential for the transmission of sensitive information in the future. The bandwidth is purposefully loaded by the attackers to limit the sources that are accessible using a variety of techniques, including blackhole intrusion, flooding, denial-of-service intrusion, etc.

4) Authenticity: The paper [7] makes the case that only authorized people should have access to both devices and data. During an engagement, end users can identify other participants. According to the study, the verification process must be error-free to prevent unauthorized people from accessing the devices or information. [8]

5) Non-repudiation: The study demonstrates that this provides assurances that the transmitters or receivers won't, respectively, interfere with the receiving or broadcasting of the data. [9]

6) Information Recentness: The information or data must meet the threshold for novelty. According to the study, IoT should ensure that time-stamped information is not re-delivered by an unauthorized party. [10]

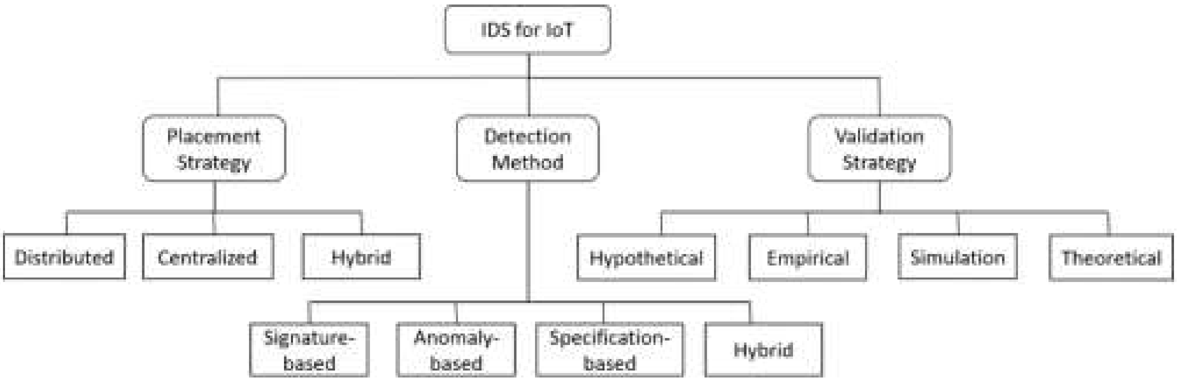


Figure 2.1 IOT types

IoT or "smart" device use has made the intrusion detection system a crucial tool for protecting against numerous invaders. The IDS scenarios utilized in IoT systems today are insufficient to handle the variety of IoT situations. IDS for IoT development is a challenging and difficult process. Therefore, a thorough investigation in this specific area is necessary. IDS has a high resource need, which restricts its use on IoT devices. IDS must thus use compact solutions that provide a high level of security. The literature review that is being provided is based mostly on [11], a resource that is often used in the field of computer science and has a wealth of material for the study and the preceding of research efforts will be presented.

Three portions make up the remainder of the FYP. Important terms for intrusion detection systems in the Internet of Things are defined in Section II. In the literature study in Section III, a taxonomy for assaults on IoT systems is examined. The research's results are presented in part IV, which is the last portion.

2.7 Taxonomy of intrusion detection in the Internet of Things

The presently suggested IDS for IoT is reviewed in the section that follows. According to the given classification: location, detection, and validation methods, the works have been categorized. Figure 1 depicts the Types for Intrusion Detection in IoT.

2.7.1 Distributed IDS Placement Strategy: The placement process is known as distributed IDS placement when an IDS is installed on each physical item. Oh et al. [12] proposed a substitute employing a distributed lightweight IDS for detecting attacks on the CMVIT 2020 Journal of Physics: Conference Series 1518 (2020) 012040 IOP Publishing doi:10.1088/1742-6596/1518/1/012040 3 IDS system. The suggested strategies may be roughly divided into two categories: early decision techniques and auxiliary shifting techniques. The goal was to lower the number of matches necessary to identify an assault. The outcomes of this approach were contrasted with those of the Wu Manber algorithm. Lee et al. [13] proposed monitoring the energy usage by each node for detecting intrusions using a similar paradigm of utilizing a lightweight IDS. By controlling the node energy, this technique managed the incoming and outgoing traffic. The model's objective was to reduce the number of resources needed to identify intrusion. In the event that the IDS was attacked, a broadcast message was delivered to every node.

2.7.2 Centralized IDS Placement Strategy: As the name implies, the IDS is positioned in the centralized portion of the IDS in the event of centralized IDS deployment.

Cho et al. [14] proposed a way to analyze the packets transiting the border router's physical and network domains. Only the packets in the border router were the subject of the effort. Kasinathan et al. [15] implemented the analysis engine together with the IDS report creation system in a powerful dedicated server while keeping the same philosophy in mind. Powerful sensors were installed in the Low power and Lossy Network, or LLN, to monitor the traffic in the network infrastructure. The LLN would then transfer the data to the IDS engine for processing. Since the host and the IDS sensor are physically connected, IDS data cannot be sent over the same wireless network.

2.7.3 Hybrid IDS Placement: Hybrid IDS placement combines centralized and dispersed placement strategies, allowing us to maximize the benefits of each. In hybrid placement, the network is divided into several portions, and each section's significant nodes host an IDS instance. Additionally, the conspicuous node is used to inspect each section's nodes individually. In comparison to dispersed IDSs, the hybrid placement IDS may be able to use a wider variety of resources due to its architecture. Using this idea, Amaral et al. [16] presented an IDS for the Internet of Things. IDS is only hosted by a select few nodes in this case. The selected nodes search for probable incursions among their neighbors. Every other node is verified by the authorized node using established protocols. Due to the changing nature of network components, each of these specified nodes uses a unique protocol. Depending on how flexible distinct protocols are created for each network segment, this notion provides advantages. 2.2 Method of Detection.

2.7.4 Signature-Based IDS: This method checks the profile of the current network using a collection of previously identified attack patterns and signatures as references. In order to assess if an incursion has taken place, a set of guidelines is followed. The aforementioned attack signatures or patterns are kept in a database so that any attack may be recognized using them. Utilizing this strategy is simple. Since this method depends on pattern matching, precise details about every attack must be recorded. This is a pricey strategy since the amount of storage space needed increases as more threats need to be detected. This approach does not identify new attacks unless the signatures are actively added to the database. As a result, the database must be updated often [17]. As a result, it is a static strategy. As a result, the disadvantages of signature-based IDS are as follows: a) It requires specialized pattern knowledge to identify assaults. b) Unknown or recent assaults cannot be discovered.[18]

2.7.5 Anomaly-Based IDS: This method is referred to as event-based detection since it examines occurrences in order to find an intrusion. After a monitoring period, the network's regular or typical behavior is identified. Any occurrence that is not consistent with routine behavior is referred to as an incursion [19]. This strategy outperforms signature-based IDS in terms of effectiveness. A host-based IDS employing Software Defined Technology (SDN) was suggested by Nobakht et al. [20]. Three basic requirements for an IoT intrusion detection system are listed in the work: an inconspicuous methodology, minimal overheads, and scalability. An anomaly-based IDS was described by Chordia and Gupta in [21]. The goal of this IDS, which is based on a data mining approach, is to reduce false positives while also increasing the effectiveness of detection. K-Means, Decision Table, and K-NN The aforementioned IDS uses methods like the majority rule-based system to keep track of traffic.

2.7.6 Specification-based IDS: In 1997, Ko et al. [22] were the first to present a specification-based IDS. Their suggested method operates similarly to the anomaly-based approach in that it looks for occurrences that are related to the specified regular behavior. CMVIT 2020 Journal of Physics: Conference Series 1518 (2020) 012040 IOP Publishing doi:10.1088/1742-6596/1518/1/012040 4. By taking into account the system's security features and rules, this behavior definition was created. Security infractions are defined as actions that do not adhere to these requirements [23]. The practical usefulness of this approach is limited, however. The difficulties in analyzing the requirements are to blame for this. The specification-based method combines the misuse-based IDS and anomaly-based IDS approaches to offer protection against known attacks as well as novel or unidentified threats [24–26].

2.7.8 Hybrid approaches: As suggested by its name, the Hybrid approach combines the previously mentioned methods to provide a high effects detection system. The goal of hybrid systems is to maximize the positive aspects of several methodologies while minimizing their negative aspects. In their study, Raza et al. [27] developed a hybrid model referred to as SVELTE. IDSs have also been explored and analyzed by Krimmling and Peter [28] using a framework that was presented to them. The outcomes demonstrate that no strategy is entirely impregnable since they all failed for one assault or another. Therefore, their approach is combining all of these strategies to provide a safe range of assaults. To combat sinkhole attacks, Cervantes et al. [29] suggested INTI IDS. The specification-based technique and the anomaly-based method are combined by INTI in place of the signature-based method used by SVELTE. 2.3.3 Validation Plan Validation, as defined by D. Chrun [30], occurs when a model that has been thus created or implemented behaves reasonably accurately and in accordance with the objective of the research. Although there are many different validation approaches, they may be generically divided into experts and data. As opposed to the data source's objectivity and quantitative validation, the instance of an expert source of knowledge awards subjective and qualitative model certification. Studying the validation approach used in intrusion detection is objective. Based on Hypothetical, Vierendeel [31] categorized validation techniques as follows: - Hypothetical, as the name implies, occurs when the observations are based on a model but not on actual data. Empirical: - Data from different operating settings is collected to carry out testing in a methodical manner. Some IoT models are evaluated by simulation.

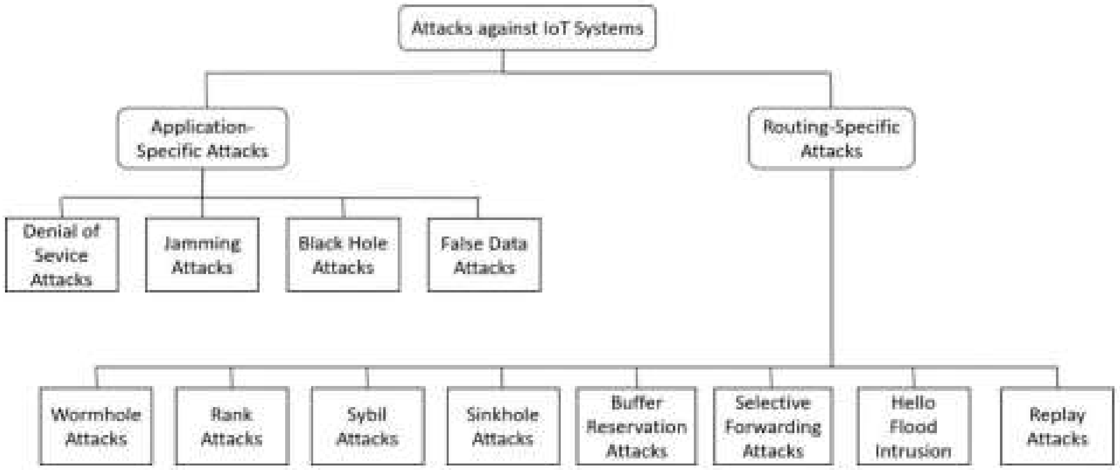


Figure 2.2 Classifications of attack against IOT

2.8 Taxonomy of attack for IoT system

Although the IoT model is a relatively new paradigm, the existing software paradigms served as the foundation for the development of the IoT software. As can be observed, the foundational elements for IoT-specific stacks (such Routing Protocol for Low-Power and Lossy Networks, or RPL, Zigbee, and 6LoWPAN) remain in the current infrastructure, such as IPv6 and IPv5. IoT networks are vulnerable to several assaults from outside and inside attackers. We cover a few cyber-attacks against IoT applications in the part that follows. The potential attacks against IoT systems are shown in Figure 2. Conference Series 1518 (2020) 012040 IOP Publishing, CMVIT 2020 Journal of Physics doi:10.1088/1742-6596/1518/1/0120405

2.8.1 Wormhole attack: It is a kind of assault where the target node is infiltrated from two different angles; in this case, attacker location is critical. The hacked nodes proclaim to one another that they are close to the base station. Additionally, a wormhole attack might be used to simulate a neighborhood setting between two clearly separated nodes [32], [33].

2.8.2 Rank attack: In the Routing Protocol for Low Power and Lossy Networks (RPL), ranking is a notion. The revised values for these node rankings are sent to every node in the network when they are changed. The RPL use the rank rule to control overhead, prevent the formation of loops, and design the ideal topology. The rank data is messed up by this kind of assault, as detailed in [34]– [36]. The attacker modifies the rank data to force the selection of the node with the lowest rank value, which affects the network's structure and impairs transmission.

2.8.3 Sybil Attack: In this kind of assault, the infected node would fake several routing protocols, detection algorithms, and cooperative techniques to attack them. Sybil attacks may be roughly categorized according to the skill of the attacker. For a thorough study, the classification is further based on graph-based detection known as the social graph-based Sybil attack (SGSD) and behavior classification-based Sybil attack (BCSD). The author concludes by discussing the difficulties and potential for security protection in IoT devices in the future.

2.8.4 Sinkhole attack: Data from all other nearby nodes is drawn to the compromised node during a sinkhole attack [38]. In this attack, the malicious node shows all other nodes that it has the lowest routing cost by routing all network traffic via itself. A bogus node is inserted into the network to initiate this kind of attack. Using the RPL as a routing protocol, R. Stephen et al. [39] suggested an Intrusion Detection System (IDS) to identify the sinkhole attack in the network. The measure employed by the proposed technique to verify the Intrusion Ration (IR) by the IDS agent is the quantity of packets sent and received. In order to identify the false node in the ensuing transmission if a malicious node is found, the IDS sends alert signals to the leaf nodes. Reduce the Intrusion Ratio using the recommended work.

2.8.5 Buffer reservation attack: After a successful fragment duplication assault, a buffer reservation attack could result. As mentioned in [40], this attack takes use of the fact that, in the event of fragmented packet transmission, the receiver is unsure of whether all of the pieces have been accurately received or not. Therefore, in accordance with the 6LoWPAN header, the receiving node reserves a buffer space.

2.8.6 Denial of Service (DoS) Attack: As the term implies, a malicious node prevents neighboring nodes from accessing resources, which prevents them from providing service as usual due to a shortage of resources. The same malicious nodes are used in a DDoS assault as in a DoS attack. This kind of attack degrades service by rendering resources inaccessible to authorized users. By providing data in certain samples or simply by flooding the network with a lot of requests or packets, a DDoS/DoS attack may create considerable chaos and have a detrimental impact on how usable the network is. Attackers often have the ability to impede the remote service [41]. Tasnuva Mahjabin et al. [42] go through a thorough analysis of distributed denial-of-service attacks, how to avoid them, and how to lessen the danger. Their research offers a critical examination of these assaults with a focus on their development, attack analysis, countermeasures, and mitigation strategies.

2.8.7 Selective forwarding attack: In this kind of assault, a hostile node impersonates a trustworthy node. This attack aims to interfere with normal transmission and routing [43]. The infected node selectively transmits inbound messages while sinking certain data packets. The corrupt node prevents certain packets from being sent, interfering with routing processes. An attacker may, for instance, decide to transmit all control messages while blocking the rest [44]. When combined with a sinkhole attack, this assault has the potential to seriously harm a network. Research in multi-stage IoT attacks is motivated by these interconnected dependencies between assaults and their effects, creating a significant research topic with a broad reach.

2.8.8 Hello Flood Intrusion: To let adjacent nodes know it is there, the routing protocol sends a hello message. In the CMVIT 2020 Journal of Physics: Conference Series 1518 (2020) 012040 IOP Publishing, the nearest node that receives this greeting message is regarded as the source. The corrupt node is shown in the node's list of neighboring nodes since it is close by [45].

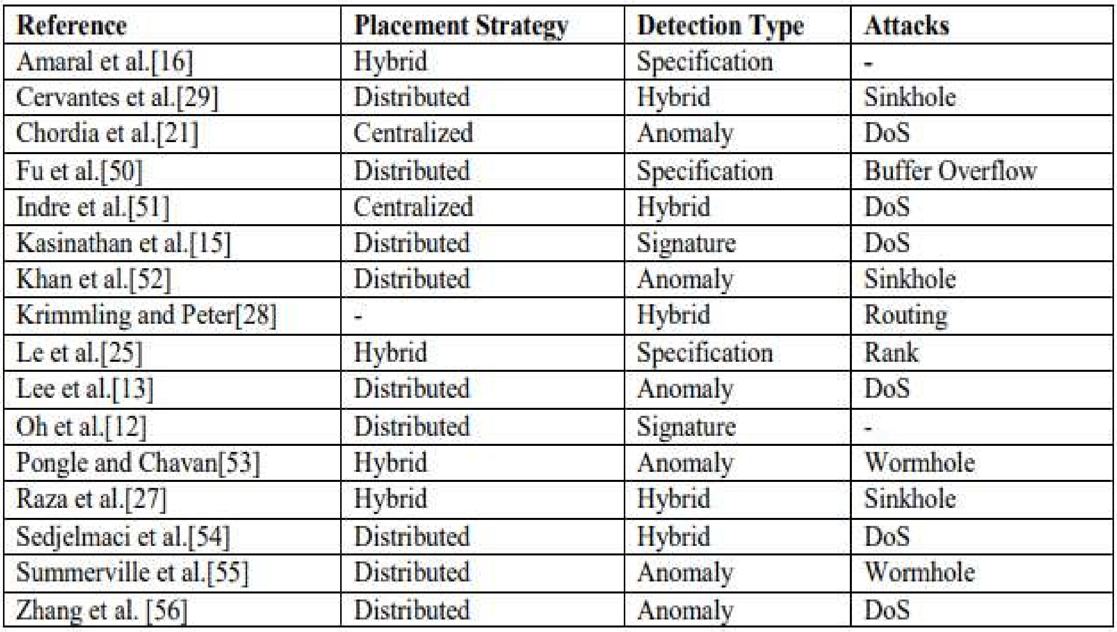
2.8.9 Replay attack: This attack involves intermittent intrusions during which data is gathered. Then, this gathered data is played back [46].

2.8.10 Jamming attack: In these kinds of assaults, the attacker keeps an eye on the wireless media. This is done to allow the destination node to regulate the frequency of the signal it receives from the sender [47]. The attacker then broadcasts a signal on this frequency in an effort to interfere with the error-free receptor [47].

2.8.11 Blackhole attack: In a blackhole attack, the attacker watches the request packets. This is accomplished by keeping an eye on the routing protocol's dynamic features, which enables the reply to be transmitted using a false reply packet [48].

2.8.12 False data attack: In this kind of attack, the attacker's initial move is to identify the present network organization and structure. The system is then compromised by the addition of manipulated measurements, which alter the estimates [49]. The Internet has become into a necessity and a part of everyday life. It occurs in a variety of spheres of human existence, including business, education, and entertainment.

It is a crucial element of daily business life [1]. The Internet has become into a necessity and a part of everyday life. It occurs in a variety of spheres of human existence, including business, education, and entertainment. It is a crucial element of daily business life [1].



Reference	Placement Strategy	Detection Type	Attacks
Amaral et al.[16]	Hybrid	Specification	-
Cervantes et al.[29]	Distributed	Hybrid	Sinkhole
Chordia et al.[21]	Centralized	Anomaly	DoS
Fu et al.[50]	Distributed	Specification	Buffer Overflow
Indre et al.[51]	Centralized	Hybrid	DoS
Kasinathan et al.[15]	Distributed	Signature	DoS
Khan et al.[52]	Distributed	Anomaly	Sinkhole
Krimmling and Peter[28]	-	Hybrid	Routing
Le et al.[25]	Hybrid	Specification	Rank
Lee et al.[13]	Distributed	Anomaly	DoS
Oh et al.[12]	Distributed	Signature	-
Pongle and Chavan[53]	Hybrid	Anomaly	Wormhole
Raza et al.[27]	Hybrid	Hybrid	Sinkhole
Sedjelmaci et al.[54]	Distributed	Hybrid	DoS
Summerville et al.[55]	Distributed	Anomaly	Wormhole
Zhang et al. [56]	Distributed	Anomaly	DoS

Table 2.1: Overview of IDS for various IOT devices

In other words, as technology progresses, network usage pervades all aspects of our lives. The probability of a network attack rises as network usage grows in popularity. One of the most pressing issues in recent years has been computer network security.

The most effective method of network security is to establish a robust security system. A firewall is one of the solutions, however since it can only detect attacks from outside the network, it is ineffective at protecting the network from attacks.

In recent years, network attacks have increased tremendously. As a result, researchers' interest in intrusion detection systems (IDSs), an alternative security method, has grown. [2].

2.9 Overview of IDS

IDS is software that keeps an eye out for criminal activity on computer networks, such as information theft, network protocol violations, and censorship. Wei Huang was the assistant editor who handled the submission's review and granted final clearance for publishing. IDSs are often used to detect network attacks from both known and unknown internal and external intruders. The bulk of current IDS techniques are incapable of dealing with the dynamic and complex nature of assaults on computer networks. Because of the constant growth of these damaging attacks, network security measures now in use are insufficient to protect computer systems. As a result, it has become critical to develop new ways and improve existing technology in this field. This study focuses on in-depth assessments of IDSs, current development approaches, easily available datasets, and unsolved challenges. This entails a thorough analysis of the primary intrusion detection methods, methodologies, tools, and strategies published in the literature [3–10]. The goal of this FYP is to investigate and analyze the present state of IDSs via a comprehensive literature study. First, a description of the system and the essential components of any IDS are supplied. IDSs are then classified depending on how they monitor network traffic, collect flow data, detect attacks, and provide alarms. Every IDS technology, technique, and approach has been thoroughly assessed within this scope. A full analysis of the work done in each area is offered, as well as a discussion of their benefits and drawbacks. Finally, popular intrusion detection technologies used to identify assaults by people, institutions, and organizations are discussed. The datasets utilized extensively in the testing and evaluation phase of the proposed intrusion detection systems were evaluated. Each intrusion detection tool's technique of detection, as well as its benefits and drawbacks, are evaluated. In many ways, the FYP for this review is different from the FYPs for the other surveys. Previous research has mostly concentrated on one or two topics, such as intrusion detection techniques or the datasets that have been used. The many IDSs elements are explored in this paper, however. Additionally, there are several recommendations for each topic. The FYP also offers assistance to commercial businesses that seek to use IDSs more successfully, in addition to researchers.

The contributions of this study are given below:

- The status, weaknesses, and new technological breakthroughs of intrusion detection systems are examined in this context.

- Technologies, methodologies, and techniques for intrusion detection are given, as well as a summary of recent work in these disciplines.

- A description of datasets that are widely utilized in intrusion detection systems.

- There is a list of well-known and popular intrusion detection tools.

- Current difficulties are investigated, and new theories for intrusion detection systems are proposed.

- Provides a comprehensive overview of intrusion detection technology and research approaches.

2.10 Intrusion Detection Systems

Monitoring computer/network events for signs of prospective intrusions, such as threats or breaches of use guidelines or accepted security practices, is known as intrusion detection. IDS primarily focus on anticipating occurrences and documenting data TABLE 1. The paper's most frequently used words and their abbreviations. relating to these occurrences and informing security administrators of collected data. IDSs are also used to identify problems with security policy, notify current risks, and deter people from security assaults, among other things. In general, components must be adequately safeguarded for an effective and successful IDS. IDS is made up of users, sensors, database servers, management servers, networks, and other components. IDS component security is critical because attackers attempt to prevent IDSs from collecting crucial data, exploiting known weaknesses, or detecting assaults. All software-based intrusion detection system components must be protected from assaults and have their operating systems and initiatives updated. Using several IDS systems may also be an option for comprehensive and accurate attack detection. Many IDS technologies are now in use, including host-based, network-based, and wireless. Each has distinct information collecting, recording, detection, and prevention capabilities.

Each technology also has benefits, such as the ability to identify specific occurrences more accurately or efficiently. As an example, host-based and network-based IDSs may be combined to provide a productive solution. In other words, the distinctive characteristics and benefits of each IDS technology should be taken into account while making a decision. Figure 1 lists the most popular intrusion detection system technologies, strategies, and techniques in the literature. In conclusion, IDSs are now required for the security of practically every individual, institution, and organization owing to the expansion of assaults, their potential for harm, and the rising reliance on technology and information systems.

2.10.1 Network-Based IDS

A network-based intrusion detection system (NIDS) monitors network device security through analyzing the protocols that have been utilized to identify unusual activity [20], [21]. TCP/IP is used by many networks to ease communication. TCP/IP is made up of four interconnected layers. When a user sends data, it is transported from the top to the bottom layer, with more information added to each layer as time passes. Data is transported from the lowest layer to the final destination through the physical network. To transport data between hosts, the four TCP/IP layers collaborate. The application layer is where the bulk of the analysis in network-based intrusion detection systems takes place. Certain network-based intrusion detection systems do minimal hardware layer analysis as well. Sensors, one or more management servers, database servers, and several consoles are often incorporated in network-based intrusion detection systems. Except for the sensors, all of the components specified are interchangeable with different IDS technologies. Network-based intrusion detection sensors monitor and assess network activities.

2.10.1.1 Security Features of NIDS Network-Based IDS

They provide many different security capabilities. The following lists common security features, roughly categorized into three different types: gathering, logging, and detection of information.

2.10.1.2 Information Collection Network-based

IDSs are unable to obtain data from communication networks. The information gathered is typically about linked hosts and network activity.

Some of the collected info features are:

- **Identifying Hosts**: An intrusion detection system (IDS) may generate a list of network hosts.

- **Identification of Operating System:** It is possible to detect which operating systems and versions are utilized by hosts. This information may be useful in finding susceptible hosts on a network. Security experts may examine the possible hazards and vulnerabilities associated with a certain operating system version by understanding the version number. This information helps them to put in place suitable security measures and to issue patches or upgrades to guard against known vulnerabilities. The identification of the operating system version is a critical step in establishing a safe network environment.

- **Identification of Applications:** By monitoring ports in use and application communication, an IDS sensor may detect application versions. This data is used to detect potentially susceptible apps and the unauthorized usage of such applications.

- **Determining Network Characterization**: Data is gathered on certain IDS sensors, network setup, and traffic. Any modifications in the network setup are quickly recognized thanks to this information. b: LOGGING Network-based intrusion detection systems record detailed information about identified events. This information is used to authenticate alerts and to investigate and correlate incidents.

Data types commonly used by network-based intrusion detection system are:

- Date and time;
- Number of connections;
- Event type;
- Protocols;
- Source and destination IP addresses;
- Number of transmitted packets;
- Application requests and responses.

2.11 Detection

Network-based IDSs offer broad detection power. A lot of network-based IDS blend signature- and anomaly-based methods to perform in-depth analysis and improve detection rates. The requests and replies are analyzed and compared with the fingerprints of known attacks when the anomaly-based method studies odd behavior. In other words, the methods' application is hierarchical.

CONNECTED WORK IDSs A broad range of monitoring skills are offered by network-based systems. In addition to NIDS, the bulk of studies use several attack detections methods to achieve high attack detection success rates. In other words, intruder detection methods often cross tracks. Wattanapongsakorn et al. [22] suggested a network-based Intrusion Detection and Prevention System (IDPS). The goal of this system is to quickly and consistently recognized known attack types and respond to strikes. The suggested method may be tried in an open network setting and used in combination with other machine-learning techniques. The results show that the proposed IDPS can identify between normal events and attacks and can also automatically avoid attacks on the victim's computer network. Furthermore, they applied the C4.5 Decision Tree algorithm in combination with a suggested way to identify unknown attack types, and they found that this system worked effectively when dealing with unknown network attack types. However, by building a way for spotting both new and known threats, this work might be further improved. Amaral et al. [23] suggested a network-based intruder detection solution for IPv6-enabled wireless sensing networks. To identify attacks, the suggested method uses traffic patterns and unusual actions. The proposed method consists of two components: PPP Sniffer and Finger2IPv6. In the suggested system, the intruder detection system locates network nodes marked as watchers. Potential attack efforts may be found by watching the packets that peers trade. The received messages are compared to the NIDS rule set. If a match is found, a warning is created and sent to the event management system. Instead of identifying intentional attacks, this suggested system may detect possible wrongdoing. To improve the system, new detecting factors should be added. Kumar et al. [24] created and tested network-based intrusion detection systems based on machine learning to identify network risks. This study uses datasets, named examples of network traffic features made by both innocent and harmful programs, and other data to build a variety of trained machine learning models. Because of the increase in mobile malware and its market draw, the main focus of this study is malware that runs on Android smartphones.

To put the proposed method to the test, traffic was created. A range of malware cases, including Premium SMS sender, backdoor, spamming, bots, ransomware, information theft, and fake antivirus, were used to create this traffic. According to the data, the suggested method has a 99.4% accuracy recognition rate for both known and new threats.

This study might be improved by growing the created dataset and putting it into the previously described current breach detection systems. According to Qassim et al. [25], an anomaly-based intrusion detection system (AIDS) may spot unfriendly network activity. It raises a warning if it sees a behavior that is out of the ordinary. Managing IDS alerts and distinguishing false positives from true threats becomes a substantial load as a result. 2021, VOLUME 9 OF 157732 M. Ozkan-Okay.

2.12 Literature Review on Intrusion Detection Systems

A two-step approach was created in this study. First, they offered a list of traits for network data that are thought to be the most useful for spotting network problems. Second, a packet header-based anomaly detection system is proposed to automatically identify behaviors using an AIDS warning algorithm.

The suggested method, which is based on machine learning techniques, is effective and efficient in identifying malicious behavior, say the writers. Machine learning methods may be used to improve this study and boost accuracy. Mazzini et al. [28] offer a unique mixed network-based IDS system that uses AdaBoost and artificial bee colony (ABC) methods to identify errors. The features were found using the ABC method. The AdaBoost method was used to assess and classify the chosen traits. The method's precision was tested using the NSL-KDD and ISCXIDS2012 datasets. The accurate score is 98.9%. According to the authors, when applied to the same dataset, the new method beat other IDSs. Accuracy may be improved further in subsequent tests, and performance can be measured using a variety of datasets. Meftah et al. [29] built an anomaly-based method for network attack detection using the UNSW-NB15 dataset. Their method is split into two major parts. They apply a range of methods, including Recursive Feature Elimination and Random Forests, to choose key features for machine learning. Then, to identify unusual traffic, they use various data mining methods, including Support Vector Machine, Gradient Boost Machine, and Logistic Regression. The Support Vector Machine gave the best accuracy, 82.11%. They then put the SVM result into a number of polynomial models to improve the accuracy of finding attack types. They assessed the effectiveness of Naive Bayes, Decision Trees, and Polynomial SVM in particular. The two-stage hybrid classification was utilized, which improved the accuracy of the data to 86.04%. This work might be improved on different datasets by creating a new classification method or using deep learning methods. NIDSs trained on weak data tend to perform badly against certain attack types, resulting in unnoticed or erroneously classified intrusions. Previous study has handled the problem of class imbalance using data-level methods that either boost minority-class instances or reduce majority-class instances. Although these balancing methods indirectly improve the performance of NIDSs, they do not solve the fundamental problem. To solve the problem of class mismatch, Bedi et al. [31] suggested a two-layer Improved Siam-IDS (I-Siam

IDS) method. I-Siam IDS describes methods for both minority and majority classes without applying any data level balancing approaches. The first layer of I-Siam IDS screens input data using a binary ensemble of Siamese Neural Networks, extreme Gradient Boosting, and Deep Neural Networks (DNN). These strikes are then passed to the second layer, where they are sorted into various attack types using the multiclass extreme Gradient Boosting algorithm (m-XG Boost).

When compared to earlier studies, I-Siam IDS showed significant gains in memory, accuracy, F1 score, precision, and AUC values for the CIDDS-001 and NSLKDD datasets. The numerical cost analysis of the suggested method is also presented to make the findings more understandable. While doing so, this study might be improved by reviewing the results of many datasets.

2.13 Evaluation of Network-Based

False negative and positive detection rates are often quite high for IDSS network-based IDSs. The bulk of the initial network-based intrusion detection systems used signature-based detection to identify known easy assaults. By integrating detection methodologies, innovative technologies have broadened the sorts of attacks that may be detected and achieved high accuracy. False positive and false negative rates are therefore reduced. Another difficulty is that they often need a great deal of tuning and modification in order to take into account the properties of the observed environment. Network-based intrusion detection systems have great detection capabilities, but they also have a number of important drawbacks. The most important ones include handling huge traffic volumes, decrypting encrypted communication, and guarding against IDS assaults. In high-demand scenarios, NIDSs are unable to perform a comprehensive analysis and so cannot detect assaults on encrypted network traffic. Furthermore, IDS sensors may impede the identification of a variety of events, particularly when stateful protocol analysis is performed. B. IDSs based on hosts in order to find possible threats, host-based IDSs (HIDS) watch the behaviors and attributes of the host. A host-based IDS keeps track of information including traffic statistics, system logs, and file access and modification [32, 33]. The majority of HIDS have agents—detection programs—installed on target hosts. Each agent keeps an eye on one host's behavior. Agents provide information to management servers, which are able to access database servers. Management and monitoring are done on consoles. Instead, then installing the agent software on specific hosts, some host-based IDSs employ specialized hardware.

Each device is placed to keep an eye on traffic on a certain host. In a technical sense, these gadgets are network-based IDSs.

Each device is specifically designed to protect one of the following:

2.13.1 Server: The agent can monitor certain programs in addition to watching the server's operating system.

2.13.2 Client Host: The operating system and popular apps like email clients and web browsers are often observed by agents meant to monitor users' hosts.

2.13.3 Application Service: Some agents are solely intended to watch a single application, such as a web server or database server program. These agents are also known as application-based intrusion detection systems (IDSs). Security features of HIDS Host-based IDSs provide a variety of security features. These include logging, detection, and other functions. a: LOGGING Host-based intrusion detection systems often record comprehensive data on observed events. Among other things, this data may be used to verify warnings, explore incidents, and correlate events.

The data fields saved by host-based intrusion detection systems are

•Date and time

• Type of event or alert

• IP address

- Port information

- Application information

- Filenames/paths and user IDs.

2.14 Detection

The majority of host-based intrusion detection systems can identify a wide range of dangerous behavior. They often mix signature-based detection methods when finding known attacks, and policy or rule sets with anomaly-based detection techniques when discovering previously unknown attacks. 2) Related Work offers a group of host-based intrusion detection methods. shows the main theme of each study as well as notable aspects of the investigations. Ou et al. created and implemented a host-based intrusion detection system with two detection methods. Back propagation neural network technology and log file analysis are two examples. Log file analysis and the BP neural network were used to spot misuse and abnormalities, respectively. By combining these two detection methods, the suggested HIDS is meant to successfully improve the accuracy and speed of intrusion detection. The results show that the proposed method improved the effectiveness and accuracy of attack detection. Creech and Hu [35] suggested a unique host-based abnormal intrusion detection method based on a semantic algorithm. This study uses irregular system call patterns to boost detecting rates while decreasing false alarm rates. The fundamental idea is to utilize a semantic framework on kernel-level system calls to aid in the discovery of odd behavior. This unique method was tested using three different datasets. The KDD98 dataset and the new ADFA Linux dataset (ADFALD) were used to measure core performance, while the UNM dataset was used to test transfer and reliability. M. Ozkan-Okay et al., 2021: A Comprehensive Systematic Literature Review on Intrusion Detection Systems (157734 VOLUME 9). According to the writers, a unique semantic-based algorithm significantly beat earlier algorithm. This study might look at cutting-edge methods for improving the stability of semantic traits and lowering training costs. According to Catherine et al. [36], host-based intrusion detection systems are not fast enough to identify attacks since they apply the whole feature set. As an answer to the aforementioned problem, this work suggested a unique Host-based IDS called CPDT (Correlation-based Partial Decision Tree Algorithm). The CPDT uses a Partial Decision Tree with Correlation Feature Selection for traffic sorting and feature selection. The method was tried and reviewed using the KDD '99 dataset, giving a 99.9458% success rate. According to the authors, CPDT gives better results than current algorithms. A new approach for finding unexplained attacks might improve this study. Subba et al. [37] presented a new HIDS structure that is resource-intensive and saves processing. The suggested method turns system call traces into n-gram vectors before compressing the input feature vectors via dimensionality reduction. A number of machine learning-based predictor models were used to measure the shortened feature vectors. The success of the proposed model was tested using an ADFA-LD dataset. The results showed that it correctly and effectively finds attacks with a low false positive rate. A variety of things might be changed to improve the results of this study. Chawla et al. [38] suggested a unique Host-based IDS to recognize a system's normal behavior based on system call patterns. This study gives an effective anomaly-based intrusion detection system based on recurrent neural networks in terms of computing. Instead of using normal LSTM networks, gated repeated units may offer considerable benefits while needing less training time. When GRUs and CNNs are joined, Anomaly IDS improves. The suggested method is applied to the ADFA dataset. As a result of faster convergence during training, the results showed that stacked CNN/GRU is about 10 times faster than LSTM. They also got a 100% True Detection Rate and a 60% False Alarm Rate using the proposed method. This work might be better by applying extra training samples or other datasets. According to Byrnes et al. [39], due to the fast growth of operating systems and the resulting complexity, decades-old records may stay out-of-date for a very long time. They tried to bridge the gap between theory models and application settings by studying the most recent Linux kernel, 5.7.0-rc1. This setting examines the usefulness of system call-based HIDS in current operating systems as well as the limits placed on HIDS writers. Following a study of recent kernel advancements, a unique method for making data and improving the detecting model is suggested. It also included a set of speed and memory limitations that must be met in order for the models to work to their full potential. Park et al. [42] performed a trial with the Leipzig Intrusion Detection Dataset (LID-DS), a host-based IDS dataset built in 2018. An intrusion detection method with host-based analysis, vector-to-image processing, training, and testing steps is also given to improve system performance. During the training and testing steps, a Siamese Convolutional Neural Network (Siamese-CNN) was made using a learning method that covered multiple stages and got good success with a small amount of data. Siamese-CNN identifies the kind of attack based on the similarity score of each assault sample translated to a picture. Accuracy is determined using a couple of shot-learning systems. The performance of the Vanilla Convolutional Neural Network (Vanilla-CNN) and Siamese-CNN was compared. In terms of accuracy, precision, recall, and F1-score measures, the suggested Siamese-CNN model beat the vanilla-CNN model by around 6%. The suggested model may be improved by improving the hyperparameter parameters to boost the accuracy of intrusion detection for known or new threats. 3) ANALYSIS OF HOST-BASED IDSS the attack types that host-based IDSs can spot are set by the system's monitoring methods. While some host-based intrusion detection systems focus on one or more of these detection methods, others mix several. Because host-based IDSs have in-depth knowledge of host features and settings, an IDS agent can often predict whether an attack on a host would be successful if it is not stopped. As with other IDS methods, host-based IDSs often cause fake positives and rejections. The efficiency of identification for host-based IDSs may be more tough. due to the fact that the majority of IDSs are ignorant of the context in which observed events, including log analysis and file system tracking, take place. Installing a new program or restarting the computer, for example, are both possibly damaging activities.

Even when the events are correctly named, it is sometimes hard to determine whether they are a regular occurrence or an attack without further details. In general, host-based IDSs that mix several detection methods may give a greater detection rate than those that use just one methodology.

Using several ways helps us to understand more about the activities taking place since each approach may watch different parts of the host. This gives a more thorough account of what happened and may also offer additional information to help identify the events' goals.

2.15 Wireless IDS

In order to spot suspicious behavior, Wireless IDSs (WIDS) track wireless network traffic and examine wireless network protocols [43], [44]. It is unable to detect questionable behavior in the application or higher layer network protocols (like TCP and UDP) that wireless network data is traversing. THE 9TH ISSUE 2021 157735 Comprehensive Systematic Literature Review on Intrusion Detection Systems via, M. Ozkan-Okay et al. It is commonly used for monitoring across the wireless network's service area. A wide range of security features are available with wireless IDS. As a relatively new category of intrusion detection system, Wireless IDS functionalities are presently rather inconsistent amongst vendors. Three areas of common security capabilities—information gathering, logging, and detection—are specified. The majority of wireless IDSs can gather data on wireless devices.

2.16 Logging Wireless: IDSs often log a large amount of information about observed events.

This information may be utilized to verify warnings, analyze occurrences, and correlate events from IDS and other sources of logging.

The following data types are routinely captured by wireless intrusion detection systems:

- Date and time
- Alert type
- Priority
- Source MAC address

2.17 Detection Wireless

By initially looking at IEEE 802.11a, b, g, and I protocol traffic, IDSs may identify attacks, incorrect settings, and policy violations at the WLAN protocol level. Higher-level communications, such as application payloads and IP addresses, are not examined by wireless IDSs. While some solutions solely carry out straightforward signature-based detection, others combine approaches for signature-based detection with those for anomaly-based detection and situational protocol analysis. Each study's core premise as well as key elements of the publications are explored. Using a Bayesian model, Meng and Li [45] create a trust-based intrusion detection system and assess its effectiveness in spotting hostile activity. This Bayesian model generates a map of trust levels among various sensor nodes. They examined the effects of a fixed and a dynamic trust threshold as well as assessed the performance of the suggested method in a wireless sensor environment. The outcomes of the experiment demonstrated the Bayesian model's potential for identifying malevolent behavior. It would be better to use many models in this investigation. There is no unique and widely used open-source WIDS solution for the detection of de-authentication and the evil twin attack, according to Afzal et al. [46]. For identifying these OSI layer assaults, they suggested an open-source WIDS in this work. These two frequent criticisms of the standard are carefully examined. These assaults are then put into practice in order to study attack behavior. Finally, a Wireless Intrusion Detection System (WIDS) is devised and built using innovative attack signatures and approaches to identify these assaults. For the two assaults used in the proposed system's assessment, accuracy rates of 89% and 93% were attained. The authors claim that while the suggested attack signatures were successful, they might yet be enhanced. Term ID is a distributed intrusion detection system that Kolias et al. [47] devised that is appropriate for wireless networks.

The method uses classification rule induction and swarm intelligence concepts to create an effective training model for intrusion detection without trading data. The suggested model's readability is another advantage. These are the major tenets of the suggested strategy.

The AWID dataset was used to evaluate this strategy. This research successfully built a model that met the needs of a contemporary WIDS in terms of user privacy and a small network footprint. Examining the accuracy rate and spotting unknown assault kinds might help these investigations. Feature selection is important for a better wireless intrusion detection system based on machine learning classifiers, according to Abdul Hammed et al. [48].

This research makes two major contributions:

- Choosing effective characteristics
- Improving a wireless intrusion detection system based on machine learning.

2.18 Evaluation of Wireless Ids

By initially looking at IEEE 802.11a, b, g, and protocol communication, wireless IDSs may identify WLAN protocol-level assaults, incorrect setups, and policy violations.

Denial-of-service attacks and physical assaults are also possible targets for them. While some Wireless IDS systems solely do signature-based detection, others include approaches for anomaly-based detection, situational protocol analysis, and signature-based detection.

Wireless IDSs provide the best level of intrusion detection accuracy when compared to other IDS types. Although wireless IDSs have strong detection capabilities, they have significant drawbacks, including:

They are unable to recognize some forms of attacks on wireless networks, including those that include passive monitoring and offline processing of wireless data.

2.19 Network Behavioral Analysis

A network behavioral analysis system (NBAs) studies network traffic or numbers on network traffic to detect abnormal traffic patterns such as Distributed Denial of service (DDoS) attacks, viruses (such as worms and backdoors), and violations of policy.[55], [56].

Sensors and consoles are often included in NBA systems, and some systems also feature administration servers.

NBA sensors are often only accessible as hardware. In that they sniff packets to keep an eye on network activity in one or more network segments, some sensors are comparable to network-based IDS sensors.

NBA sensors employ network flow data supplied by routers and other network devices rather than actively monitoring networks.

2.19.1 Security Features of NBA Kason Go

This system has a wide range of security features. These include gathering data, logging, and detection. Some systems can also handle events and give security information.

2.19.2 Information Collection NBA

These systems have tremendous ability to gather info, but detection requires an understanding of the mainframes' peculiarities. NBA sensors are capable of creating and storing lists of hosts that communicate on networks which are monitored.

The data collected by NBA in order to identify assaults:

- IP address
- OS
- IP protocols
- TCP and UDP ports

2.19.3 Logging NBA

These devices record detailed information about identified occurrences. Among other things, this data may be used to verify warnings, explore incidents, and correlate events.

NBA software often records the following data types:

- Date and time

- Activity or alert type

- Protocols

- Source and destination IP addresses

- TCP or UDP ports

- Additional package header fields

- Number of bytes and packets.

2.20 Detection NBA

Detection NBA systems are normally capable of detecting a wide range of harmful behavior. Most solutions analyze network flows largely via anomaly detection and certain stateful protocol analysis methods.

The majority of NBA technologies do not provide signature-based detection.

2.21 Related Work

Traditional intrusion detection systems, In Youssef and Emam's [57] research, have limited capabilities and often are unable to identify hostile behavior or sound alert (false positive) when there is no irregularity in the network. In this research, it is hypothesized that the employment of a Network Behavior Analysis system and the application of Data Mining (DM) methods to network traffic data would aid in the development of improved intrusion detection systems. A hybrid technique to intrusion detection was suggested by Youssef and Emam. Examining DM and NBA methodologies for network intrusion detection, it is said that combining both approaches may more successfully identify network intrusions. A network behavior analysis system, an IDPS technology, was examined and assessed by Nitin et al. An IDPS (intrusion detection and prevention system) technology called a network behavior analysis system (NBAS) analyses network traffic to find threats including distributed denial of service (DDoS) assaults and certain kinds of malware that produce irregular traffic patterns. This article offers a thorough assessment of NBA technology.

2.21.1 Evaluation of Network-Based Analysis NBA

This method operates by identifying major departures from expected behavior, providing excellent accuracy in identifying assaults that produce unusually high levels of network activity in a short period of time. While NBA systems have strong assault detection capabilities, they also have severe limitations. One of them is that small-scale assault detection by NBA systems is not particularly good. The detection rate is not high, particularly if the assault is carried out carefully and does not break the established regulations.

2.21.2 Intrusion Detection Methodologies

Intrusion detection system (IDs) procedure have classifications:

1. Signature-based model
2. Anomaly-based model
3. Stateful protocol analysis

Different intrusion detection systems (IDSs) employ distinct strategies to detect network attacks. Signature-based detection is known for its speed and effectiveness in identifying known attack patterns, but it falls short in detecting zero-day attacks, which are previously unknown and therefore lack a signature. On the other hand, anomaly-based technology is capable of detecting novel network-based threats that have not been seen before. However, it has a drawback of generating false alarms by considering normal traffic as an attack. This can lead to an excessive number of alerts that need to be investigated.

Furthermore, there is a stateful protocol approach that focuses on monitoring the state and behavior of network connections to identify potential attacks. While this approach may be able to detect certain new types of attacks, it is often time-consuming, complex, and may not be effective against sophisticated and cleverly designed attacks.

It is important to understand that each IDS approach has its strengths and limitations. By combining multiple detection techniques and employing intelligent analysis, it is possible to enhance the overall effectiveness of an intrusion detection system and reduce false positives while detecting a wide range of network threats.

2.22 Signature-Based Model

A signature serves as a recognizable pattern associated with a known attack. Signature-based detection involves the correlation of these signatures with observable events to identify potential threats [67]. When a match is found during the comparison process, the system generates a warning or additional report. Examples of indicators include attempts to compromise network security using the login "root," emails with subjects like "Free programs" that are characteristic of well-known and pervasive malware, or system logs indicating the deactivation of the host controller.

Signature-based detection is the most fundamental method used, which compares observed events against a set of predefined signatures. If any of the specified attack criteria are present in the list, a warning is triggered. Intrusion Detection Systems (IDSs) that employ signature-based detection excel at identifying known threats but struggle to detect new attacks or variations of known threats. For the instance, if an attacker renames an infected file from "prog.exe" to "prog2.exe," a signature designed to detect "prog.exe" would not match.

Researchers have conducted comprehensive evaluations of signature-based intrusion detection approaches and related methodologies, assessing the efficiency of each study and the underlying concepts of the proposed methods. In a study by Kumar and Sangwan [68], Snort, a popular Network Intrusion Detection System (NIDS), was employed for identifying signature-based attacks. The study focused on analyzing anomalous connections detected during transmission using Snort and the DARPA Dataset. Snort analyzes network packets by cross-referencing them against a database of known threat signatures, which is periodically updated. This IDS system has demonstrated its capability to identify and investigate intrusions in real network traffic. According to the authors, their research will aid new users in comprehending the concept of a Snort-based IDS. However, it is worth exploring and testing different intrusion detection systems to enhance the study's findings.

One of the primary challenges with signature-based IDSs is efficiently handling large volumes of incoming traffic, as each packet must be examined against every signature in the database. When an intrusion detection system faces high traffic loads and loses packets, potential attacks may go unnoticed. In their work, Uddin et al. [157740 VOLUME 9, 2021] proposed a comprehensive systematic review of the literature on intrusion detection, addressing this issue and providing valuable insights into the field.

CHAPTER 3

Problem Formulation

3 Problem Formulation

3.1 Proposed Model

This study proposes an AI-based solution for the data-driven security component of a smart grid system, aiming to develop a machine learning model that can efficiently and accurately detect network traffic packets with a low False Acceptance Rate (FAR) and a high Detection Rate (DR). The selection of optimal features is crucial in achieving this objective, and the study employs the Particle Swarm Optimization (PSO) search method to identify the best features from a given feature group. The NSLKDD and KDD99 datasets are utilized for binary classification (normal vs. anomaly) as well as multiclass classification to predict attack categories such as denial of service (DoS), R2L, U2R, Probe, and Normal class. Once the attacks are successfully categorized, further classification is performed to precisely identify the specific type of anomaly.

The suggested model consists of six steps. Firstly, the KDD99 and NSLKDD datasets are separately read in the data reading stage. Then, data preparation is carried out, which involves replacing missing values with the mean and eliminating outliers. Data normalization is applied to scale the data, followed by data encoding to convert non-numeric values into numeric ones after normalization.

The ideal feature selection step, utilizing PSO, is performed as part of the data preprocessing phase. The subsequent part describes the working mechanism of PSO. In the third step, machine learning models are supplied with the selected features. In the fourth step, several models are trained using 70% of the data and labels, while the remaining 30% is used for testing. The experiment phase constitutes the fifth step, and the assessment phase concludes the model evaluation.

By leveraging the optimal feature subset and AI models, this study proposes an AI-based solution to enhance the data-driven security aspect of smart grid systems. It emphasizes the importance of feature selection and employs the PSO search method for this purpose.

The NSLKDD and KDD99 datasets are utilized for binary and multiclass classification, enabling the identification of various attack categories. The proposed model follows a systematic approach involving data reading, preparation, normalization, encoding, feature selection, machine learning model training, experimentation, and assessment

.

3.2Datasets

3.2.1 KDD99 dataset

One of the most well-known datasets used for IDS in network security is KDD99. A modified version of the 1998 DARPA is KDD99. It was created in an MIT research lab and used as a standard by IDS designers to assess different approaches and methods. In addition to containing 4900000 rows and 41 features with binary labels, the KDD99 dataset also lists 22 network assaults. DoS, Probe, U2R, R2L, and Normal are the four main assaults that make up a class designation. It represents all the packets carrying the KDD99 dataset that were utilized, both normal and anomalous. For the anomalous and normal classes, respectively, 97 277 and 396 731 packets are utilized to create ensemble machine learning classifiers that can be trained and tested.

The remaining 30% of the dataset is utilized for testing and validation, with the remaining 70% of this dataset is used for training.

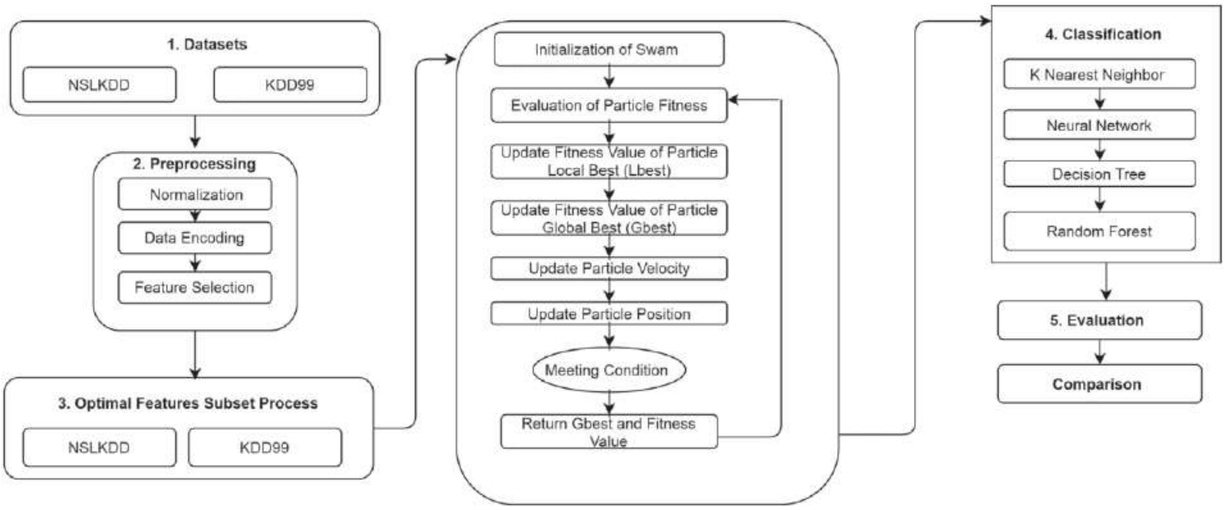


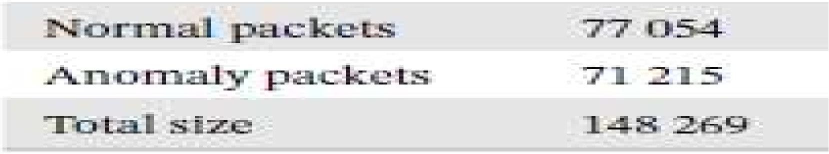
Figure 3.1 Proposed methodology

3.2.2 NSLKDD dataset

The KDD99 dataset has been updated and is known as NSLKDD. The KDD99 dataset contains duplicate Values, however NSLKDD does not. Additionally, there are no conflicting values in NSLKDD. In all, 148 517 instances and 41 features are available in NSLKDD for training and testing purposes.

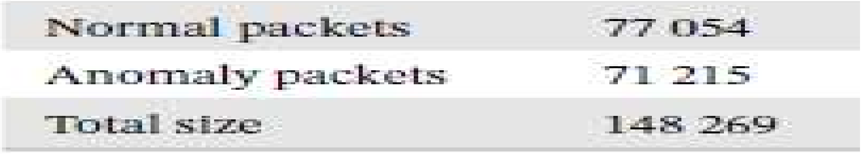
The overall quantity of regular and anomalous packets that included the study's NSLKDD dataset. To train and evaluate machine learning models, 71 215 anomaly packets and 77 054 regular packets were employed, respectively. The remaining 30% of the KDD99 dataset is utilized in testing, while the remaining 70% is used for training. Lists the 103 789 and 44 481 anomalous and normal packets that were utilized to test and train the models of the machine learning, respectively.

3.3 Dataset for Mathematical Models & Calculations



Normal packets	77 054
Anomaly packets	71 215
Total size	148 269

Table 3.1: NSLKDD datasheet Normal and Anomaly packets



Normal packets	77 054
Anomaly packets	71 215
Total size	148 269

Table 3.2: NSLKDD datasheet training and testing packets

3.4 Normalization

Data cleaning activities are carried out on the dataset once it has been chosen in order to normalize the features and eliminate noise from the dataset.

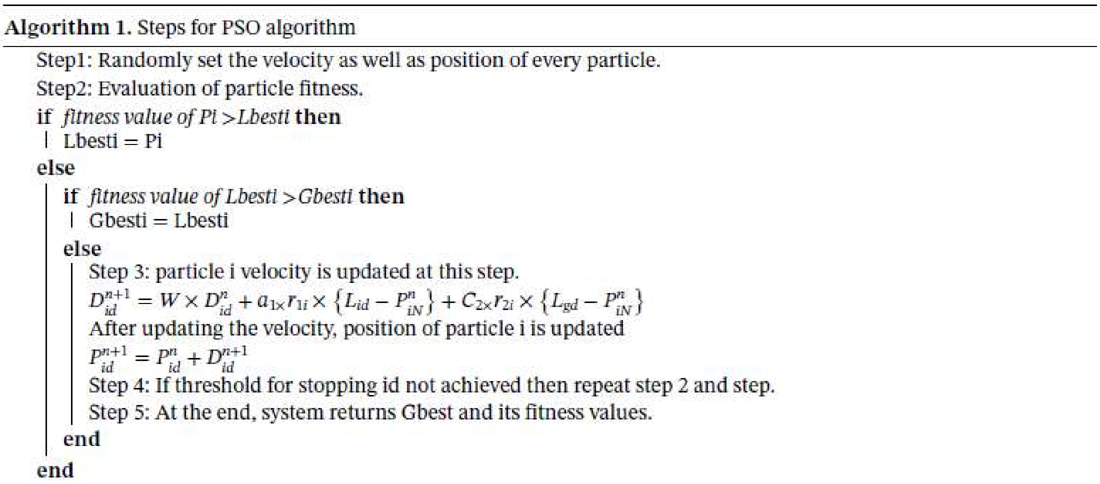
Different methods of normalization are used, however in this study the min-max normalization strategy is utilized since it scales more well and addresses the problems associated with outliers that z-score normalization has.

Values in the [0, 1] range are normalized using the min-max scale. Below is the equation for min-max normalization.

$$Z_i = \frac{Y_i - \min(Y)}{\max(Y) - \min(Y)}.$$

3.5 Data encoding

We eliminate redundant and incompatible values from the datasets prior to data encoding. The nominal qualities are then changed to numeric values, which is why machine learning algorithms base their computations on numeric values rather than nominal ones. So, before supplying data to the suggested model, this stage is completed.



```
Algorithm 1. Steps for PSO algorithm
Step1: Randomly set the velocity as well as position of every particle.
Step2: Evaluation of particle fitness.
if fitness value of Pi >Lbesti then
    Lbesti = Pi
else
    if fitness value of Lbesti >Gbesti then
        Gbesti = Lbesti
    else
        Step 3: particle i velocity is updated at this step.
        $D_{id}^{n+1} = W \times D_{id}^{n} + a_{1\times} r_{1i} \times \{L_{id} - P_{iN}^{n}\} + C_{2\times} r_{2i} \times \{L_{gd} - P_{iN}^{n}\}$
        After updating the velocity, position of particle i is updated
        $P_{id}^{n+1} = P_{id}^{n} + D_{id}^{n+1}$
        Step 4: If threshold for stopping id not achieved then repeat step 2 and step.
        Step 5: At the end, system returns Gbest and its fitness values.
    end
end
```

Figure 3.2: Algorithm of PSO

3.6 Feature selection

The crucial stage after feature normalization is feature optimization. Optimal characteristics enhance DR and FAR in addition to accuracy. Finding feature subsets that can cooperate with many classifiers to give very good results is the core goal of feature optimization. For feature selection in this study, we used the PSO approach. In 1995, Eberhart and Kennedy [54] created PSO, a generic optimization method that was motivated by the movement patterns of fish and bird flocks. PSO is one of the most effective methods for dealing with non-smooth global issues.51 PSO also has a very high convergence rate and provides the best solutions quickly.[55] Additionally, genetic algorithms are applied for the best feature selection, which results in excellent DR. However, the problem with the convergence rate of genetic algorithms is quite sluggish, and it can become very bad if individuals from the public are included.[56] The swarm particle is randomly initialized before being sent to the search arena, where we may find the quickest route by adjusting the velocity and location of the particles.

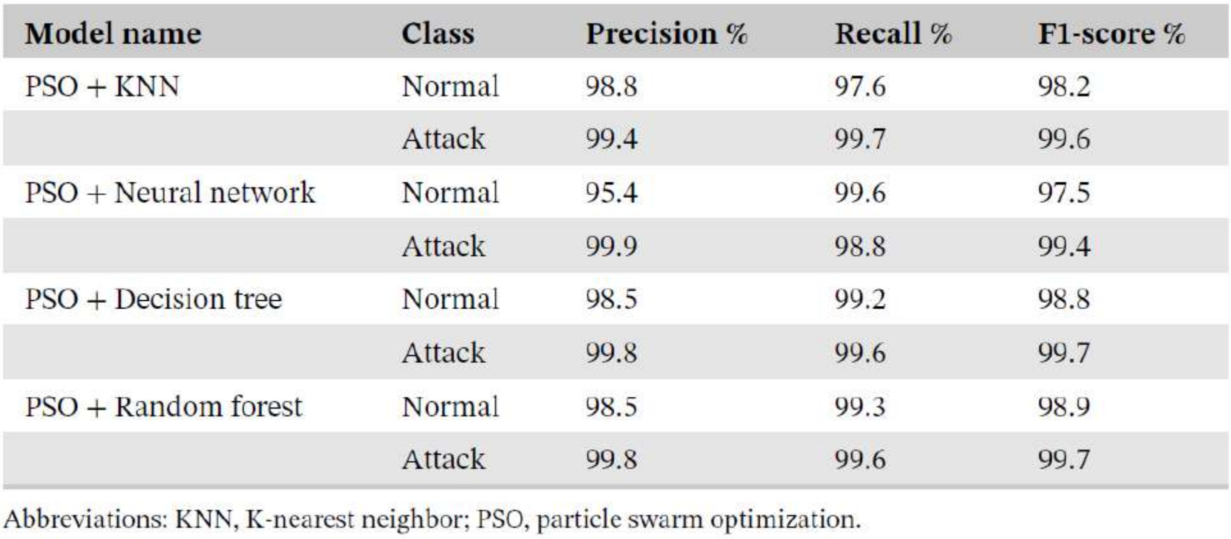
CHAPTER 4

Simulation and Results

4 Simulation and Results

4.1 Simulation and Results

The experimental results obtained from the KDD99 and NSLKDD datasets are discussed in this section. All the experiments were conducted using Google Colab. We can observe the performance metrics of different machine learning models. For the KNN algorithm, the accuracy, recall, and f1-score for the normal class were determined to be 98.89%, 97.60%, and 98.20% respectively. Similarly, for the anomalous class, the accuracy, recall, and f1-score were found to be 99.40%, 99.70%, and 99.60% respectively. When applying the Random Forest algorithm, the performance metrics for the normal class showed an accuracy of 98.50%, a recall of 99.30%, and an f1-score of 98.90%. For the assault class, the precision, recall, and f1-score were respectively calculated as 99.80%, 99.60%, and 99.70%. The Decision Tree (DT) algorithm achieved an accuracy score of 98.50% for the normal class, while the Neural Network (NN) algorithm achieved an accuracy score of 95.40% for the same class. The recall and f1-scores for the normal class using DT were found to be 99.30% and 98.40% on average respectively. Similarly, the recall and f1-scores for the normal class using NN averaged at 99.30% and 98.40% respectively. Regarding the attack class, the precision, recall, and f1-scores were reported as 97%, 99.60%, and 99.50% respectively when utilizing the DT and NN algorithms.



Model name	Class	Precision %	Recall %	F1-score %
PSO + KNN	Normal	98.8	97.6	98.2
	Attack	99.4	99.7	99.6
PSO + Neural network	Normal	95.4	99.6	97.5
	Attack	99.9	98.8	99.4
PSO + Decision tree	Normal	98.5	99.2	98.8
	Attack	99.8	99.6	99.7
PSO + Random forest	Normal	98.5	99.3	98.9
	Attack	99.8	99.6	99.7

Abbreviations: KNN, K-nearest neighbor; PSO, particle swarm optimization.

Figure 4.1: Results of Simulations

DT and random forest obtained 0.8% and 0.6% FAR, respectively, whereas the KNN classifier using the KDD99 dataset scored 2.4% FAR, which is high when compared to other classifiers. In terms of FAR, NN beat other classifiers for KDD99, achieving 0.5% FAR.

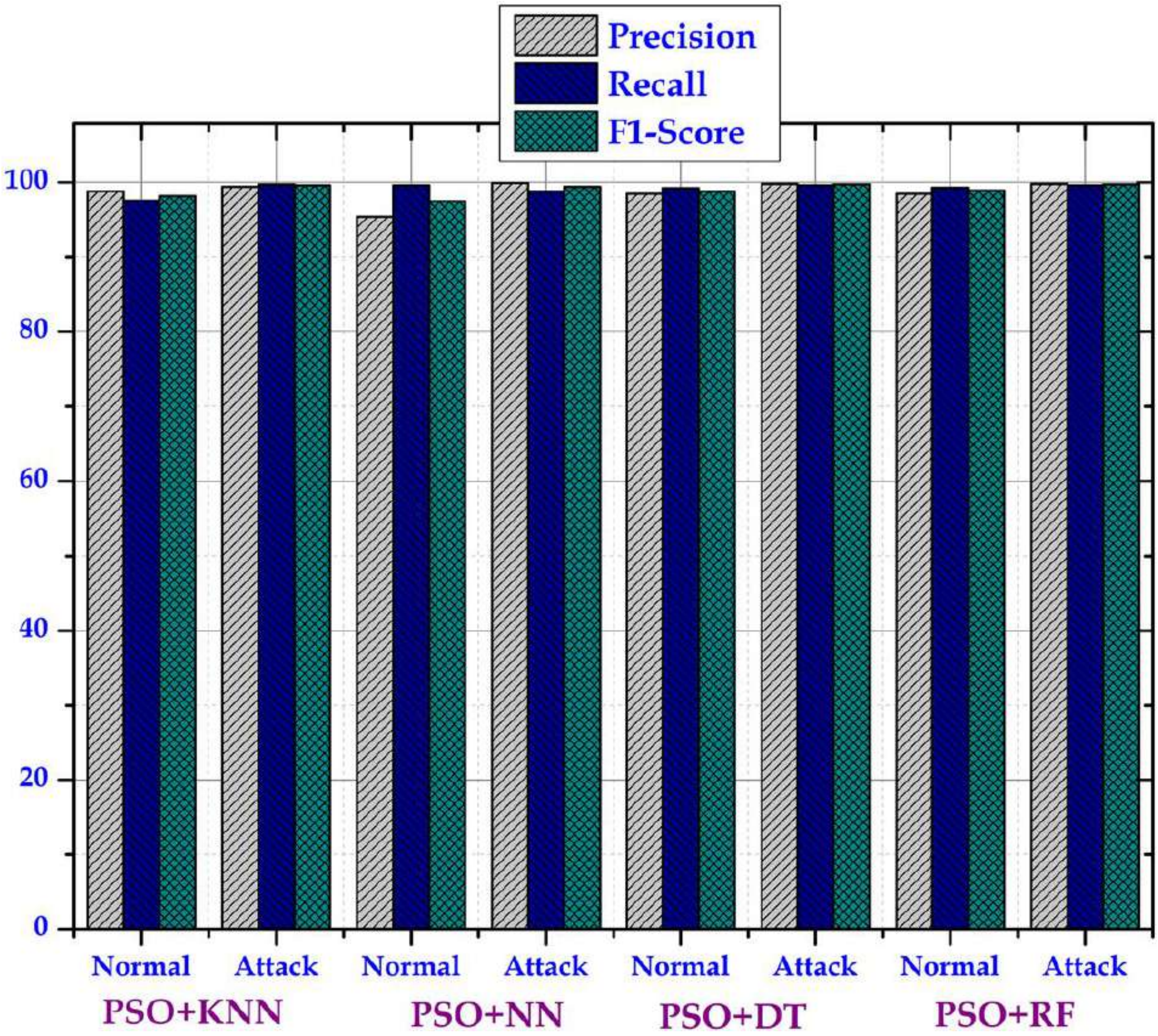
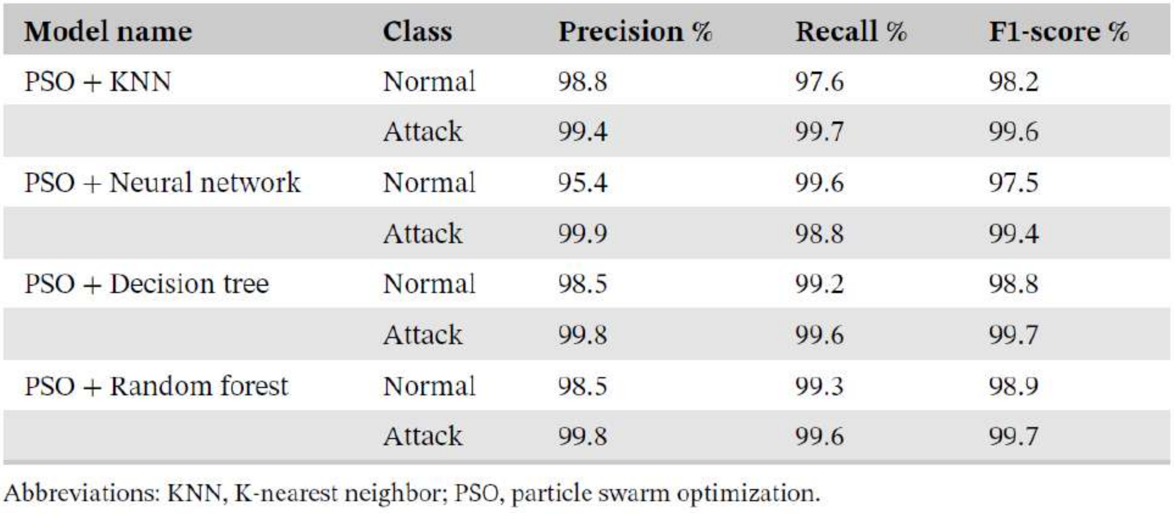


Figure 4.2: Graph of Results

The superior performance of the Neural Network (NN) algorithm can be attributed to its effectiveness in handling large datasets, as demonstrated by the KDD99 dataset, which contains more data compared to the NSLKDD dataset. On the other hand, the Random Forest algorithm showcased promising results in terms of False Alarm Rate (FAR) when applied to the NSLKDD dataset. Due to its ensemble nature, Random Forest combines multiple Decision Trees (DTs) to achieve an impressively low FAR of 0.08%. This outperformed other classifiers such as DT, K-Nearest Neighbors (KNN), and NN in terms of overall outcomes.

To elaborate further, when the KNN classifier was employed on the KDD99 dataset, it correctly identified 118,779 packets as attacks, with only 337 out of 119,116 packets being incorrectly categorized. This resulted in an accuracy of 97.60% for the normal class and 99.70% for the attack class. Out of the total 29,090 packets, the KNN classifier successfully recognized 28,390 packets, while mistakenly identifying 700 packets. The classifiers in question achieved a Detection Rate (DR) of 99.70%. Furthermore, the DT and Random Forest classifiers achieved True Positives (TPs) of 118,672 and 118,680, respectively. The Random Forest algorithm also obtained True Negatives (TN) of 28,902.



Model name	Class	Precision %	Recall %	F1-score %
PSO + KNN	Normal	98.8	97.6	98.2
	Attack	99.4	99.7	99.6
PSO + Neural network	Normal	95.4	99.6	97.5
	Attack	99.9	98.8	99.4
PSO + Decision tree	Normal	98.5	99.2	98.8
	Attack	99.8	99.6	99.7
PSO + Random forest	Normal	98.5	99.3	98.9
	Attack	99.8	99.6	99.7

Abbreviations: KNN, K-nearest neighbor; PSO, particle swarm optimization.

Table4.1: Result of False Alarm Rate

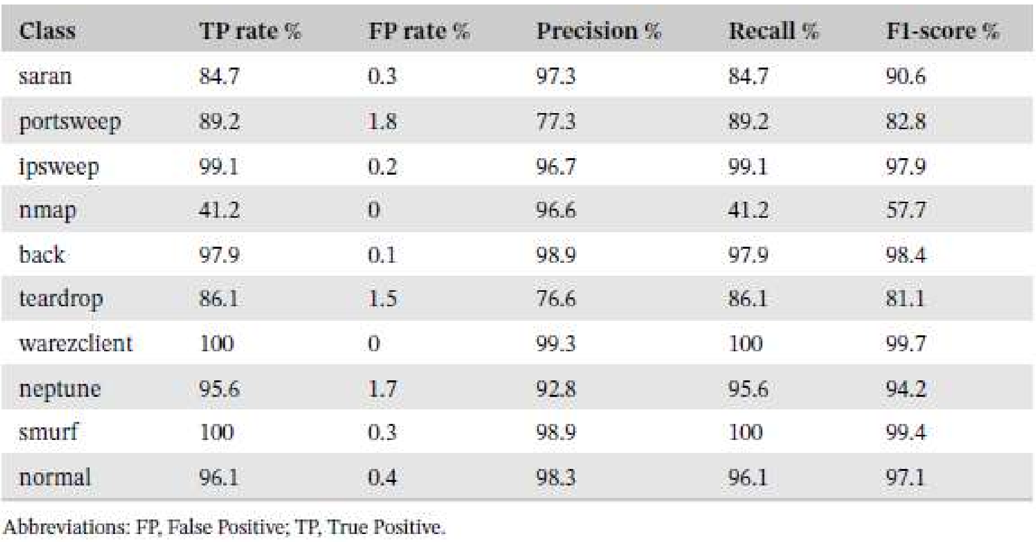
The random forest classifier achieved a False Positive (FP) score of 188 and a False Negative (FN) score of 444, while the decision tree (DT) classifier had a True Negative (TN) count of 28,850. However, the DT misclassified a total of 676 packets. Both the random forest and DT classifiers achieved high Detection Rates (DR) of 99.60%. The neural network (NN) algorithm also showed promising results with a DR of 99.20%, accurately recognizing 118,161 attack packets with a 99.20% accuracy rate and 28,927 regular packets with a 99.40% accuracy rate.

By combining NN and random forest with Particle Swarm Optimization (PSO), an accuracy of 99.65% and a DR of 99.30% were achieved. However, there were misclassifications of 95 packets in the attack class and 163 packets in the normal class. The precision, recall, and f1-scores for the normal class were 99.40%, 99.90%, and 99.70% respectively. Similarly, the anomaly class achieved an accuracy, recall, and f1-score of 99.90%, 99.40%, and 99.60% respectively. The overall accuracy of the model was 99.51%, with 99.80% accuracy for the normal class and 99.20% accuracy for the attack class.

For the DT classifier, it accurately recognized 21,307 out of 21,431 anomaly packets with a 99.40% accuracy rate. Similarly, 23,093 out of 23,124 regular packets were properly classified as normal traffic with a 99.90% accuracy rate.

The precision, recall, and f1-score for the attack class were 99.80%, 99.40%, and 99.70% respectively, while for the normal class, they were 99.50%, 99.90%, and 99.70% respectively.

Using a multilayer perceptron (MLP), an accuracy of 97.90% for the anomalous class and 99.50% for the normal class was achieved.



Class	TP rate %	FP rate %	Precision %	Recall %	F1-score %
saran	84.7	0.3	97.3	84.7	90.6
portsweep	89.2	1.8	77.3	89.2	82.8
ipsweep	99.1	0.2	96.7	99.1	97.9
nmap	41.2	0	96.6	41.2	57.7
back	97.9	0.1	98.9	97.9	98.4
teardrop	86.1	1.5	76.6	86.1	81.1
warezclient	100	0	99.3	100	99.7
neptune	95.6	1.7	92.8	95.6	94.2
smurf	100	0.3	98.9	100	99.4
normal	96.1	0.4	98.3	96.1	97.1

Abbreviations: FP, False Positive; TP, True Positive.

Table4.2: Tables of Results

Other studies reported overall accuracies of 98.5% and 97.87% using MLP classifiers. However, MLP showed poorer performance compared to the DT and random forest algorithms, possibly due to class imbalance and insufficient training and testing data.

When using the NSLKDD dataset with an MLP classifier, the precision, recall, and f1-score for the normal class were 95.10%, 99.90%, and 97.40% respectively. For the anomalous class, the accuracy, recall, and f1-score were 99.90%, 94.50%, and 97.10% respectively.

The normal class achieved precision, recall, and f1-scores of 98.30%, 96.10%, and 97.10% respectively. Rates of True Positive (TP) and False Positive (FP) were 96.10% and 0.4% respectively.

The Warezclient and Smurf attacks achieved a 100% detection rate with FP rates of 0% and 0.3% respectively. The recall for both attacks was 100%, and the average f1-score for both assaults exceeded 99%. For the Port Sweep attack, when compared to other attacks using a DT, it had a higher FP rate of approximately 1.8%.

Its precision and f1-score were 77.20% and 82.80% respectively. The IP sweep attack achieved average scores of 98.50% accuracy, recall, f1-score, and TP rate.

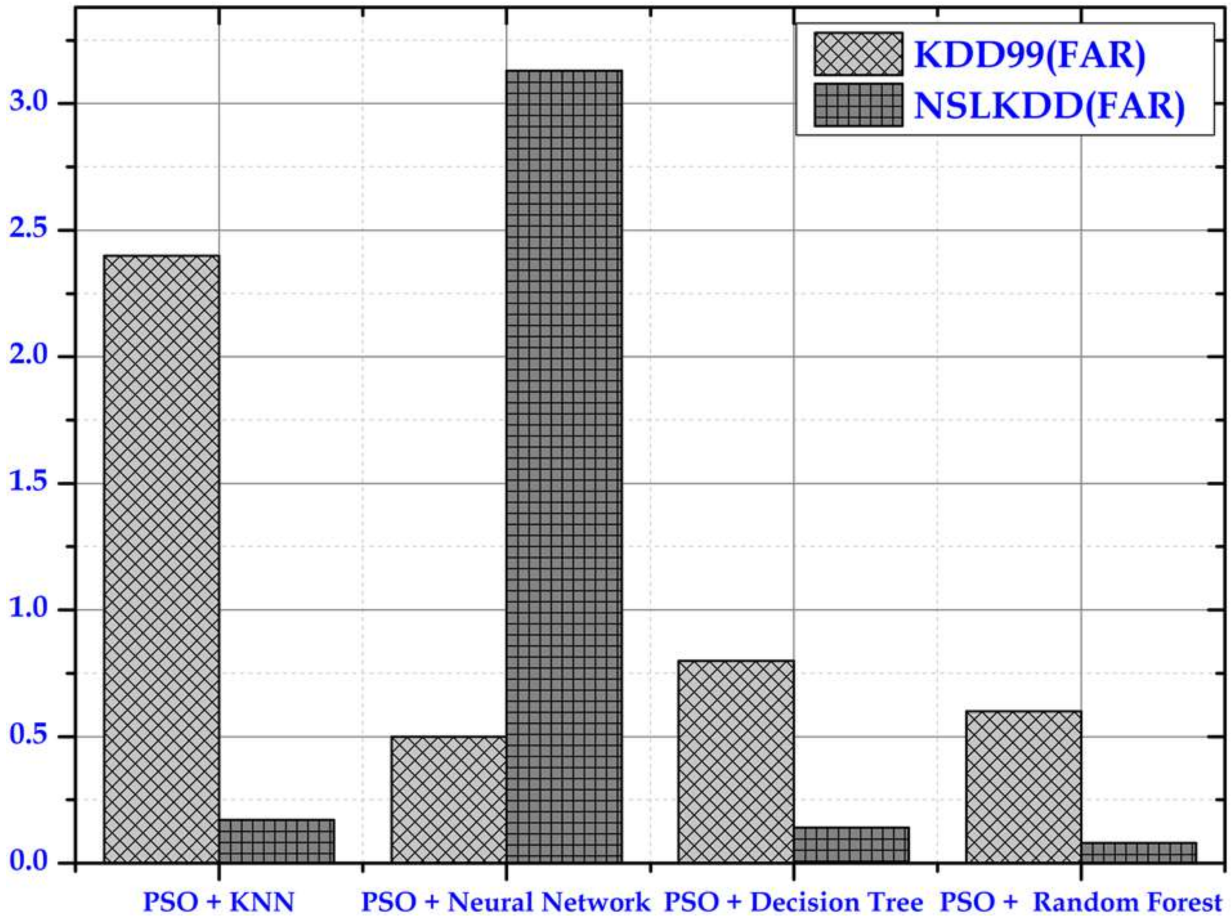


Figure 4.3: For KDD99 and NSLKDD datasets. FAR, false alarm rate

4.2 Conclusion

This proposed Final Year Project (FYP) suggests the implementation of an Intrusion Detection System (IDS) for smart grid systems, utilizing feature selection techniques. The project employs weighted Particle Swarm Optimization (PSO) to improve the False Alarm Rate (FAR) in the IDS. The selection of the best features is based on the KDD99 and NSLKDD datasets, and once the feature selection is done, machine learning models are employed for further analysis. During the project, various machine learning methods were applied to the KDD99 and NSLKDD datasets. These datasets were categorized into two classes: attack class and normal class, considering multiple types of attacks. The KDD99 dataset includes nine different types of attacks, while the NSLKDD dataset consists of 21 types. To prepare the datasets for analysis, preprocessing steps were performed, such as converting non-numeric values into numeric encoding. The data was then normalized using the min-max method. Feature selection was carried out using the PSO algorithm, which helped identify the most relevant features for the IDS. After feature selection, several machine learning methods were applied to both datasets. In terms of accuracy, training time, and FAR, Random Forest and Neural Networks (NN) outperformed other approaches.

Furthermore, the proposed technique was compared to existing research in the field. Experimental results demonstrated that the suggested method showed superior performance in terms of Detection Rate (DR), FAR, and accuracy when applied to the KDD99 and NSLKDD datasets.

4.3 Future Work

We want to replicate this experiment in the future utilizing deep learning algorithms and several classes using feature selection techniques. With that, we will be able to get values for our outcomes that are more exact and accurate, which will lead to increases in quality.

Chapter 5

Sustainable Development Goals

5 Sustainable Development and Goals

5.1. Introduction

Microgrid buildings, with their interconnected systems, have become an integral part of the modern infrastructure landscape. However, these buildings are not immune to cyber threats and intrusions, which can jeopardize their security, reliability, and overall sustainability. To address this challenge, the integration of AI-based intrusion detection and prevention techniques in microgrid buildings has emerged as a promising solution. This chapter explores the sustainable development goals (SDGs) associated with AI-based intrusion detection in microgrid buildings and analyzes the implications of implementing or lacking such techniques.

The sustainable development goals provide a comprehensive framework for global development, encompassing various dimensions such as energy, infrastructure, innovation, and resilience. By examining two case studies—one without the proposed intrusion detection and prevention technique, and the other with it—we can identify how these techniques align with specific SDGs. This analysis sheds light on the potential benefits of implementing AI-based intrusion detection and prevention techniques, as well as the risks associated with neglecting them. Ultimately, this exploration emphasizes the significance of integrating sustainable practices into cybersecurity measures to achieve long-term social, economic, and environmental sustainability.

5.2. Sustainability Development Goals and their Relevance

5.2.1 Affordable and Clean Energy

By detecting and preventing intrusions, an AI-based intrusion detection and prevention technique can help maintain the stability and reliability of energy supply in microgrid buildings, supporting the objective of ensuring access to affordable, reliable, sustainable, and modern energy for all.

5.2.2 Industry, Innovation, and Infrastructure

The absence of an intrusion detection and prevention technique can leave microgrid buildings vulnerable to cyber threats, compromising the security and resilience of critical infrastructure. Implementing such techniques promotes resilient infrastructure, inclusive and sustainable industrialization, and innovation in cybersecurity.

5.2.3 Sustainable Cities and Communities

Without intrusion detection and prevention measures, the security and resilience of microgrid buildings are compromised. This can negatively impact the goal of creating inclusive, safe, resilient, and sustainable cities and human settlements.

5.2.4 Climate Action

The absence of intrusion detection and prevention techniques can pose risks to the stability and performance of microgrid systems, hindering efforts to combat climate change and promote renewable energy integration. Implementing effective cybersecurity measures supports the goal of climate action.

5.2.5 Affordable and Clean Energy

The implementation of an AI-based intrusion detection and prevention technique in microgrid buildings can contribute to maintaining the stability and reliability of energy supply, ensuring access to affordable, reliable, sustainable, and modern energy for all.

5.2.6 Industry, Innovation, and Infrastructure

Implementing intrusion detection and prevention techniques in microgrid buildings enhances the security and resilience of critical infrastructure. This aligns with the objective of developing resilient infrastructure, promoting inclusive and sustainable industrialization, and fostering innovation.

5.2.7 Sustainable Cities and Communities

With the proposed intrusion detection and prevention technique, microgrid buildings become more secure and resilient, enabling the creation of inclusive, safe, resilient, and sustainable cities and human settlements.

5.2.8 Climate Action

The implementation of an effective intrusion detection and prevention technique contributes to the stability and performance of microgrid systems, supporting climate action goals by promoting renewable energy integration, reducing greenhouse gas emissions, and ensuring the long-term sustainability of these systems.

5.3. Conclusion

The integration of AI-based intrusion detection and prevention techniques in microgrid buildings presents a vital opportunity to address the cybersecurity challenges that these complex systems face. By examining the two case studies, it is evident that implementing such techniques aligns with several sustainable development goals.

The SDGs related to AI-based intrusion detection and prevention in microgrid buildings include SDG 7 (Affordable and Clean Energy), SDG 9 (Industry, Innovation, and Infrastructure), SDG 11 (Sustainable Cities and Communities), and SDG 13 (Climate Action). These goals highlight the importance of securing microgrid buildings to ensure access to affordable and clean energy, promote resilient infrastructure and innovation, create sustainable cities and communities, and support climate action efforts.

By leveraging AI technologies for intrusion detection and prevention, microgrid buildings can enhance their energy efficiency, strengthen cybersecurity, foster innovation, and contribute to the development of inclusive and secure communities. Moreover, these techniques play a crucial role in supporting climate action by safeguarding renewable energy integration and reducing greenhouse gas emissions.

In conclusion, the implementation of AI-based intrusion detection and prevention techniques in microgrid buildings is not only crucial for ensuring their security but also aligns with the broader agenda of sustainable development. By addressing cybersecurity risks and promoting the sustainable use of energy and infrastructure, these techniques contribute to a safer, more resilient, and sustainable future for microgrid buildings and the communities they serve.

Chapter 6

REFERENCE

6 References

1. Otoum Y, Liu D, Nayak A. DL-IDS: a deep learning–based intrusion detection framework for securing IoT. *Trans Emerg Telecommun Technols*. 2019; e3803. https://onlinelibrary.wiley.com/doi/full/10.1002/ett.3803.

2. Yousaf A, LoanA, Babiceanu RF, YousafO. Physical-layer intrusion detection system for smart jamming attacks. *Trans Emerg Telecommun Technol*. 2017;28(11):e3189.

3. Irshad O, Khan MUG, Iqbal R, Basheer S, Bashir AK. Performance optimization of IoT based biological systems using deep learning. *Comput Commun*. 2020; 155:24-31.

4. Vora J, Kaneriya S, Tanwar S, Tyagi S, Kumar N, Obaidat M. TILAA: Tactile Internet-based ambient assistant living in fog environment. *Futur Gener Comput Syst*. 2019; 98:635-649.

5. Uppal HAM, Javed M, Arshad M. An overview of intrusion detection system (IDS) along with its commonly used techniques and classifications. *Int J Comput Sci Telecommun*. 2014;5(2):20-24.

6. Bhattacharya S, Kaluri R, Singh S, Alazab M, Tariq U, et al. A novel PCA-firefly based XGBoost classification model for Intrusion detection in networks using GPU. *Electronics*. 2020;9(2):219.

7. Alazab M, Khan S, Siva Rama Krishnan S, Pham Q, PraveenKumarReddy M, Gadekallu TR.Amultidirectional LSTMmodel for predicting the stability of a smart grid. *IEEE Access*. 2020;8:85454–85463.

8. Almseidin M, Alzubi M, Kovacs S, Alkasassbeh M. Evaluation of machine learning algorithms for intrusion detection system.

FYP presented at: Proceedings of the 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY); 2017:000277-000282; IEEE.

9. Iwendi C, Maddikunta PKR, Gadekallu TR, Lakshmanna K, Bashir AK, Piran MJ. A metaheuristic optimization approach for energy efficiency in the IoT networks. *Softw Pract Exp*. 2020. https://doi.org/10.1002/spe.2797.

10. Jyothsna V, Prasad VR, Prasad KM. A review of anomaly-based intrusion detection systems. *Int J Comput Appl*. 2011;28(7):26-35.

11. Kaur H, Kumar N, Batra S. ClaMPP: a cloud-based multi-party privacy preserving classification scheme for distributed applications. *J Supercomput*. 2019;75(6):3046-3075.

12. Aujla GS, Kumar N, Singh M, Zomaya AY. Energy trading with dynamic pricing for electric vehicles in a smart city environment. *J Parall Distrib Comput*. 2019;127:169-183.

13. AghdamMH, Ghasem-Aghaee N, BasiriME. Application of ant colony optimization for feature selection in text categorization. FYP presented at: Proceedings of the 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence);

2008:2867-2873; IEEE.

14. Aghdam MH, Tanha J, Naghsh-Nilchi AR, Basiri ME. Combination of ant colony optimization and Bayesian classification for feature selection in a bioinformatics dataset. *J Comput Sci Syst Biol*. 2009;2(3):186-199.

15. Nguyen VG, Brunstrom A, Grinnemo KJ, et al. 5G mobile networks: requirements, enabling technologies, and research activities. *A Comprehensive Guide to 5G Security*. Hoboken, New Jersey: John Wiley & Sons; 2018:31-57.

16. Qasim OS, Algamal ZY. Feature selection using particle swarm optimization-based logistic regression model. *Chemom Intell Lab Syst*. 2018;182:41-46.

17. Ma T, Xu C, Zhou Z, Kuang X, Zhong L. SE-PSO: resource scheduling strategy for multimedia cloud platform based on security enhanced virtual migration. FYP presented at: Proceedings of the 2019 15th International Wireless Communications & Mobile Computing

Conference (IWCMC); 2019:650-655; IEEE.

18. Iwendi C, Khan S, Anajemba JH, Mittal M, Alenezi M, Alazab M. The use of ensemble models for multiple class and binary class classification for improving intrusion detection systems. *Sensors*. 2020;20(9):2559.

19. Xue B, Zhang M, Browne WN. Particle swarm optimization for feature selection in classification: a multi-objective approach. *IEEE Trans Cybern*. 2012;43(6):1656-1671.

20. Xue B, Zhang M, Browne WN. Particle swarm optimization for feature selection in classification: Novel initialization and updating mechanisms. *Appl Soft Comput*. 2014;18:261-276.

21. Reddy T, Swarna Priya RM, Parimala M, et al. A deep neural networks-based model for uninterrupted marine environment monitoring. *Comput Commun*. 2020;157:64-75

22. Revathi S, Malathi A. A detailed analysis on NSL-KDD dataset using various machine learning techniques for intrusion detection. *Int J Eng Res Technol (IJERT)*. 2013;2(12):1848-1853.

23. TavallaeeM, Bagheri E, Lu W, Ghorbani AA. A detailed analysis of the KDD CUP 99 data set. FYP presented at: Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications; 2009:1-6.

24. LiuG, Yi Z, Yang S.Ahierarchical intrusion detection model based on the PCA neural networks. *Neurocomputing*. 2007;70(7-9):1561-1568. KHAN et al. **23 of 24**

25. Heba FE, Darwish A, Hassanien AE, Abraham A. Principle components analysis and support vector machine-based intrusion detection system. FYP presented at: Proceedings of the 2010 10th International Conference on Intelligent Systems Design and Applications;

2010:363-367; IEEE.

26. Chae H, Jo B, Choi SH, Park TK. Feature selection for intrusion detection using NSL-KDD. *Recent Adv Comput Sci*. 2013;20132:184-187.

27. Ahmad I, Hussain M, Alghamdi A, Alelaiwi A. Enhancing SVM performance in intrusion detection using optimal feature subset selection based on genetic principal components. *Neural Comput Appl*. 2014;24(7-8):1671-1682.

28. Manekar V, Waghmare K. Intrusion detection system using support vector machine (SVM) and particle swarm optimization (PSO). *Int J Adv Comput Res*. 2014;4(3):808.

29. Tong L, Wu Q. Intrusion feature selection algorithm based on particle swarm optimization. *Int J Comput Sci Netw Sec (IJCSNS)*. 2014;14(12):40.

30. Zhang T, Kuang X, Zhou Z, Gao H, Xu C. An intelligent route mutation mechanism against mixed attack based on security awareness. FYP presented at: Proceedings of the 2019 IEEE Global Communications Conference (GLOBECOM); 2019:1-6; IEEE.

31. Patel R, Bakhshi D, Arjariya T. Random particle swarm optimization (RPSO) based intrusion detection system. *Int J Adv Technol Eng Expl*. 2015;2(5):60.

32. Bamakan SMH, Amiri B, Mirzabagheri M, Shi Y. A new intrusion detection approach using PSO-based multiple criteria linear programming. *Proc Comput Sci*. 2015; 55:231-237.

33. Syarif AR, Gata W. Intrusion detection system using hybrid binary PSO and K-nearest neighborhood algorithm. FYP presented at: Proceedings of the 2017 11th International Conference on Information & Communication Technology and System (ICTS); 2017:181-186;

IEEE.

34. Mukherjee S, Sharma N. Intrusion detection using naive Bayes classifier with feature reduction. *Proc Technol*. 2012; 4:119-128. 35. Tesfahun A, Bhaskari DL. Intrusion detection using random forests classifier with SMOTE and feature reduction. FYP presented at: Proceedings of the 2013 International Conference on Cloud & Ubiquitous Computing & Emerging Technologies; 2013:127-132; IEEE.

36. Shrivas AK, Dewangan AK. An ensemble model for classification of attacks with feature selection based on KDD99 and NSL-KDD data set. *Int J Comput Appl*. 2014;99(15):8-13.

37. Ahmad I. Feature selection using particle swarm optimization in intrusion detection. *Int J Distrib Sens Netw*. 2015;11(10):806954.

38. Eesa AS, Orman Z, Brifcani AMA. A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. *Expert Syst Appl*. 2015;42(5):2670-2679.

39. Rai K, Devi MS, Guleria A. Decision tree based algorithm for intrusion detection. *Int J Adv Netw Appl*. 2016;7(4):2828.

40. AmbusaidiMA, He X, Nanda P, Tan Z. Building an intrusion detection system using a filter-based feature selection algorithm. *IEEE Trans Comput*. 2016;65(10):2986-2998.

41. Bamakan SMH, Wang H, Yingjie T, Shi Y. An effective intrusion detection framework based on MCLP/SVM optimized by time-varying chaos particle swarm optimization. *Neurocomputing*. 2016;199:90-102.

42. Thaseen IS, Kumar CA. Intrusion detection model using fusion of chi-square feature selection and multi class SVM. *J King Saud Univ-Comput Inf Sci*. 2017;29(4):462-472.

43. Pajouh HH, JavidanR,Khayami R, AliD,Choo KKR.Atwo-layer dimension reduction and two-tier classification model for anomaly-based intrusion detection in IoT backbone networks. *IEEE Trans Emerg Topics Comput*. 2016; 7:314-323

44. Shone N, Ngoc TN, Phai VD, Shi Q. A deep learning approach to network intrusion detection. *IEEE Trans Emerg Topics Comput Intell*. 2018;2(1):41-50.

45. Naseer S, Saleem Y, Khalid S, et al. Enhanced network anomaly detection based on deep neural networks. *IEEE Access*. 2018; 6:48231-48246.

46. Sakr MM, Tawfeeq MA, El-Sisi AB. Network intrusion detection system based PSO-SVM for cloud computing. *Int J Comput Netw Inf Sec*.

2019;11(3):22.

47. Woo Jh, Song JY, Choi YJ. Performance enhancement of deep neural network using feature selection and preprocessing for intrusion detection.

FYP presented at: Proceedings of the 2019 International Conference on Artificial Intelligence in Information and Communication

(ICAIIC); 2019:415-417; IEEE.

48. Cadima J, Cerdeira JO, Minhoto M. Computational aspects of algorithms for variable selection in the context of principal components.

Comput Stat Data Anal. 2004;47(2):225-236.

49. Greselin F, Zitikis R. From the classical Gini index of income inequality to a new Zenga-type relative measure of risk: a modeler's perspective. Econometrics. 2018;6(1):4.

50. Goldberg DE, Holland JH. Genetic algorithms and machine learning; 1988.

51. Bai Q. Analysis of particle swarm optimization algorithm. Comput Inf Sci. 2010;3(1):180.

52. Kayacik HG, Zincir-Heywood AN, Heywood MI. Selecting features for intrusion detection: A feature relevance analysis on KDD 99

intrusion detection datasets. FYP presented at: Proceedings of the 3rd Annual Conference on Privacy, Security, and Trust, vol. 94;2005:1723-1722.

53. Nykvist C, LarssonM, Sodhro AH, Gurtov A. A lightweight portable intrusion detection communication system for auditing applications.

Int J Commun Syst. 2020;33: e4327.

54. Kennedy J, Eberhart R. Particle swarm optimization. FYP presented at: Proceedings of the Proceedings of ICNN'95-International

Conference on Neural Networks; vol. 4, 1995:1942-1948; IEEE.

55. Shi Y, Eberhart R. A modified particle swarm optimizer. FYP presented at: Proceedings of the 1998 Ieee International Conference on

Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No. 98TH8360); 1998:69-73; IEEE.

24 of 24 KHAN et al.

56. Ahmad I, Amin F. Towards feature subset selection in intrusion detection. FYP presented at: Proceedings of the 2014 IEEE 7th Joint

International Information Technology and Artificial Intelligence Conference; 2014:68-73; IEEE.

57. Bre F, Gimenez JM, Fachinotti VD. Prediction of wind pressure coefficients on building surfaces using artificial neural networks. Energy Build. 2018; 158:1429-1441.

58. Quinlan J. Program for machine learning. C4 5; 1993.

59. Kim G, Lee S, Kim S. A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. Expert Syst Appl.

2014;41(4):1690-1700.

60. Mulay SA, Devale P, Garje G. Intrusion detection system using support vector machine and decision tree. Int J Comput Appl. 2010;3(3):40-43.

61. Shakil M, Fuad Yousif Mohammed, Arul R, BashirAK, Choi JK. A novel dynamic framework to detect DDoS in SDN using metaheuristic clustering. Trans Emerg Telecommun Technol. 2019; e3622. https://doi.org/10.1002/ett.3622.

62. Iwendi C, Khan S, Anajemba JH, Bashir AK, Noor F. Realizing an efficient IoMT-assisted patient diet recommendation system through machine learning model. IEEE Access. 2020; 8:28462-28474.

63. Peng K, Leung V, Zheng L, Wang S, Huang C, Lin T. Intrusion detection system based on decision tree over big data in fog environment. Wireless Communication Mob Computer. 2018; 2018:10

64. Chung YY, Wahid N. A hybrid network intrusion detection system using simplified swarm optimization (SSO). Appl Soft Comput. 2012;12(9):3014-3022.

65. Tang TA, Mhamdi L, McLernon D, Zaidi SAR, Ghogho M. Deep learning approach for network intrusion detection in software-defined networking. FYP presented at: Proceedings of the 2016 International Conference on Wireless Networks and Mobile Communications (WINCOM); 2016:258-263; IEEE.

66. Dash T. A study on intrusion detection using neural networks trained with evolutionary algorithms. Soft Comput. 2017;21(10):2687-2700.