International Journal of Computer Sciences and Engineering
Vol.**12**, Issue.**5**, pp.**01-10**, May **2024**
ISSN: 2347-2693 (Online)
Available online at: www.ijcseonline.org

ISSN: 2347-2693 (E)

Research Article

Adaptive Hybrid Case-Based Neuro-Fuzzy Model for Intrusion Detection and Prevention for Smart Home Network

Davies I.N.[1*], Taylor O.E.[2], Anireh V.I.E.[3], Bennett E.O.[4]

[1,2,3,4]Dept. of Computer Science, Faculty of Science, Rivers State University, Port-Harcourt, Nigeria

**Corresponding Author: isobo.davies@ust.edu.ng*

Received: 23/Mar/2024; **Accepted**: 25/Apr/2024; **Published**: 31/May/2024. **DOI**: https://doi.org/10.26438/ijcse/v12i5.110

Abstract: The advent of Internet-of-Things (IoT) technology has ushered in a new era of unprecedented interconnectivity, by transforming our living space into a dynamic ecosystem. The challenging part of this is the security risk it poses on the network. Due to the vulnerabilities that are usually associated with smart devices, integrating them within the smart home ecosystem presents significant concern for the need of preserving data privacy, network traffic classification, and proper management of trusted devices. Various techniques have been employed in the development of Network Intrusion Detection System (NIDS) to safeguard the network against the evolving nature of attack deployed by cyber-criminals. This paper presents an Adaptive Hybrid Case-Based Neuro-Fuzzy System (HCBNFS) technique to the development of a robust and efficient Intrusion Detection and Prevention System (IDPS). The HCBNFS technique deploys the CBR as a detection engine to easily detect already known traffic patterns on the network, while the NFIS was deployed as a tuning factor to the reverse phase of the CBR to further investigate unknown traffic to the detection engines case-base. Five network packet features were selected as input variables to the proposed model. These features are the source IP, destination IP, source port, destination port, and network protocol. The model was trained using the CIC-IoT2022 dataset. For this study, the CIC-IoT2002 and a synthetic dataset were used for both testing and evaluating the performance of the system. The experimental results of the system using the CIC-IoT2022 dataset achieved 99% accuracy rate in intrusion detection, and recorded 99.5% for precision, recall, and F1-Score. The empirical evaluation of the proposed model validates its effectiveness and contributes towards the development of a more robust intrusion detection and prevention system. By enhancing data confidentiality, privacy, and security, the model represents a significant step forward in the safeguarding IoT-based smart home network against cyber-criminals.

Keywords: Internet-of-Things, Artificial Intelligence, Neuro-Fuzzy Inference System, Case-Based Reasoning, Smart Home, Machine Learning

1. Introduction

In recent years, cyber security issues have not only affected our social live and development but have also threatened the security and stability of politics, military operations, economic landscapes, and cultural developments. This shift is attributed to the rapid proliferation of the IoT, which has introduced numerous cutting-edge gadgets and devices into our daily lives, transforming them into ubiquitous "smart" entities. These devices now possess the capability to connect and communicate with both humans and other devices, amplifying the scope and complexity of cyber threats [17].

However, as we enjoy this new era of interconnectivity, its benefits are often associated with huge number of devices or network vulnerabilities. IoT-based smart devices are particularly vulnerable due to their inadequate built-in security measures. Therefore, the integration of these devices within the IoT ecosystem presents significant challenges for efficient and secure connectivity. It creates new and complex security issues which must be adequately addressed so it will not hinder the applications of IoT.

Technically, one way to provide solution to this problem is to deploy an Intrusion Detection System (IDS) capable of monitoring systems that detect any potentially malicious endeavours and initiate an alarm. Once notifications or flags are received, a situation handler can proceed to examine the issue and implement the requisite measures to mitigate the detected threat [2].

However, various techniques had been employed to the design and development of an IDPS. Till date, no technique or method has been proven to precisely detect and fend off all anomalies or malicious traffic from the network. The main reason for deploying an IDPS on an IoT network is to inspect network traffics, detect, and mitigate those viruses or attacks that a traditional firewall or access rule may not be able to

detect. This is because most devices operate between layer-2 and layer-4 of the TCP/IP Systems. Nonetheless, researcher stated that any IoT driven by AI is regarded as an intelligent IoT [16].

To achieve better accuracy in terms of detecting and preventing anomalies or malicious network traffic packets on the network, we proposed a novel Hybrid Case-Based Neuro-Fuzzy System (HCBNFS) approach which combines the learning capabilities of Case-Based Reasoning (CBR) and adaptive nature of Neuro-Fuzzy Inference System (NFIS) to develop a model for network Intrusion Detection and Prevention which will be suitable for IoT-based smart home network. The HCBNFS model employed the CBR as a primary intrusion detection engine, while the NFIS was deployed as a tuning factor to the revise phase of the CBR. With the increasing complexity of IoT infrastructure regarding to its security, reliability, and efficiency, our study aims to address the following problems of:

i. Data Privacy Preservation,
ii. Classification of network traffic patterns, and
iii. Management of trusted device identity.

2. Related Works

In the study of information system security, the development of Intrusion Detection System (IDS) had been known as a successful element in securing computer systems, networks, and devices. It had been mainly used for identifying network anomalies on the communication network. However, Nevertheless, with the present rise of ML models in recent years, researchers explored possibilities of utilizing them for detecting Endpoint MiTM attacks. They focused on employing ML algorithms in conjunction with ARP analysis to detect these attacks. Their proposed model combined signal processing and ML algorithms for accurate detection. They assessed the correctness of their model using linear-based ML classification models. Their empirical results showed that their technique attained an impressive accuracy rate of 99.72% in sensing MiTM attacks when utilizing linear-based ML classification. This demonstrates the effectiveness of using ARP analysis as a detection technique for Endpoint MiTM attacks [9].

Researcher conducted a study that explicitly examines the design of a smart home that integrates a smart grid with IoT technology. The system's functionality, applied protocols, and encountered challenges were also discussed. The study examines the function of IoT in enabling the operation of intelligent power grids and residences [13].

Technically, IoT & Cyber-Physical System (CPS) are both contributing significant roles in government operations and our day-to-day life activities. The need for security increases by the day as sensitive data are being sent via the internet. However, researchers had proposed a Mathematically Modified Adaptive Neuro-Fuzzy Inference System (MMANFIS). While Achieving 96% accuracy, their model surpassed other existing systems [10].

Development of smart application had been made possible due to the proliferation of IoT devices. Ahanger *et al.* [1] developed a system to detect personnel identity in smart home. In their proposed method, Fog computing was utilised to perform real-time examination of the load bearing capacity, dimensions, and gait. The task of their model prediction is realized by the predictive learning based AFIS. Their methodology was specifically built to produce warnings and critical notifications based on real-time data. Their proposed model outperformed other existing prediction model that was compared to it.

Butt *et al.* [5] proposed approach introduces an innovative anomaly-based IDS tailored for smart home networks. It tackles the common issues of overfitting and underfitting faced by traditional anomaly detection models. Additionally, the approach aims to deliver high-performance intrusion detection capabilities through a hybrid approach. Their model adopts the one-hot encoding with pandas' technique to extract features from database. The researchers evaluated their model using a pre-existing dataset and found that it significantly improved performance while reducing misclassification and other constraints compared as compared to other existing systems or solution.

Farhin *et al*., [6] presented a feedforward network for identifying attacks in the IoT utilising SDN and FNN. They utilised SDN controller to analyse the traffic flow, identify irregularities, and then obstruct incoming traffic and source nodes. They employed the FNN to identify cyber-attack in the SDN. Their proposed method is efficient of identifying several categories of cyber-threats such as the malicious code, MiTM, DDoS, and the side-channel attack. They utilised the NSL-KDD datasets to train and evaluate the FNN. They employed Matlab-Simulink to simulate their model. The examined performance of their system demonstrates that their approach is effective for accurately detecting both MITM and DDoS cyber-attacks exhibit a notable accuracy rate of 83%, indicating its ability to effectively detect a wide range of assaults in the IoT system.

Alrayes *et al.* [4] presented a novel Metaheuristics Features Selection with Fuzzy Logic enabled IDS (MFSFL-IDS). The main purpose of their approach was to recognize the presence of network or device intrusions and accomplish security in the IoT ecosystem. They employed the Henry Gas Solubility Optimization (HGSO) algorithms as a feature selection approach to derive useful feature vectors. Also, they applied the ANFIS technique to enhance the recognition and classification of malicious traffic patterns on the network. Further, they used the Binary Bat Algorithm (BBA) to adjust the parameters involved in the ANFIS model. They carried out a comprehensive experimental validation of their proposed model using benchmark dataset. The experimental results of their proposed MFSFL-IDS model showed superior accuracy of 99.80% with regards to performance when compared with other recent approaches.

An anomaly-based IDS using Fuzzy Logic was proposed by Almseidin *et al* [3]. They implemented their fuzzy inference

as a detection method for Distributed Denial of Service (DDoS) attacks. They applied their proposed method to an open-source DDoS dataset. Their proposed system achieved the optimal outcome by employing the InfoGain feature selection technique by obtaining a 91.1% and 0.006% true and false positive rate respectively.

Imtiaz *et al.* [7] introduced an effective cybersecurity methodology founded on machine learning algorithms to identify anomalies employing a lightweighted Intrusion Detection Systems. They evaluated their proposed DeepNet using the UNSW-NB15 dataset and conducted comparisons with other existing system or solution. In classifying network-based anomalies, their model achieved an accuracy of 99.16% when utilizing every attribute and achieves a post-feature minimization accuracy of 99.14%.

Researchers [14] developed a technique to identify various type of malicious attacks using the triangular membership function of Fuzzy Logic Systems. Their work utilizes the KDD dataset for both training and testing of the system. They used standard deviation for the normalization the KDD dataset, and further matched it with the various predefined signatures. Their study employed the Centroid defuzzification method to convert the fuzzy values into crips values for identifying the attacks using a layered approach. Their results of their study achieved an accuracy of 61.53% and their system performance showed 44.44%. However, they argued that their rules were designed in such a way that they can be used for detecting maximum possible attacks towards the IoT network.

Moudni *et al*. [11] presented a novel approach for mobile ad hoc networks by employing ANFIS and PSO. By constructing an adjacent list that documents the activity of all neighbouring entities, they were able to retrieve a database from the mobile ad hoc network and use it to calculate their input parameters. Their experimental results indicated that the system achieved a high performance. To identify black hole attacks, they suggested that future research compare their IDS with other proposed IDSs. Further, they intend to broaden their investigation to identify additional assaults that have taken place on the mobile ad hoc network.

Rajput and Sikka [12] proposed a self-healing architecture for services that leverages agent technology to achieve autonomous capabilities. Their proposed system consists of multiple agents with distinct roles. Specifically, the planning agent serves as a vital function in the process of making informed decisions to restore the system from an ill-state at runtime. They incorporated a CBR (Case-Based Reasoning) fault recovery mechanism, where similarities of the perceived fault & previously logged, unsuccessful cases were computed to identify the most suitable solution. The case with the highest similarity index was considered the closest match to the failure. They employed various recovery strategies and validated their architecture utilizing an SOA-based application. Their performance of their approach was evaluated using various metrics.

3. Methodology

This study employs the Design Science Research Method (DSRM) for carrying out the research construct due to how it generally improves human understanding on how to design and build new artefact with an intention for solving real-world problems [8]. While the Object-Oriented Design Approach (OODA) is employed for the systems structural designing and development.

4. Dataset Description

We utilize the sample Canadian Institute for Cybersecurity (CIC-IoT2022) dataset for both training and testing the efficiency of our application. The dataset captured network traffic instances from various IoT smart devices traversing the gateway in six distinct categories of experiments utilizing Wireshake and dumpcap. The CIC-IoT2022 dataset has about 373,988 instances and aggregates data regarding was being deemed normal behaviour, 195,184 instances of MiTM attacks, 141,709 instances of DoS attacks, 240,314 Scannings, and 1,327, 748 MiraiBotnet attacks. Nonetheless, this study focuses more on the attack experiments, where several attacks were performed on IoT devices, and their attack traffic pattern were also captured. For this study, the pre-processed relevant packet features selected form this dataset is the source IP, destination IP, communication protocol, source port, and destination port [15].

5. System Design

The proposed IDPS was designed based on task division of various system components, into self-sufficient and individual reusable objects with each containing the data and behaviour relevant to itself. The architectural design of smart home and the proposed system for this study is captured in Figure 1 and 2 respectively.

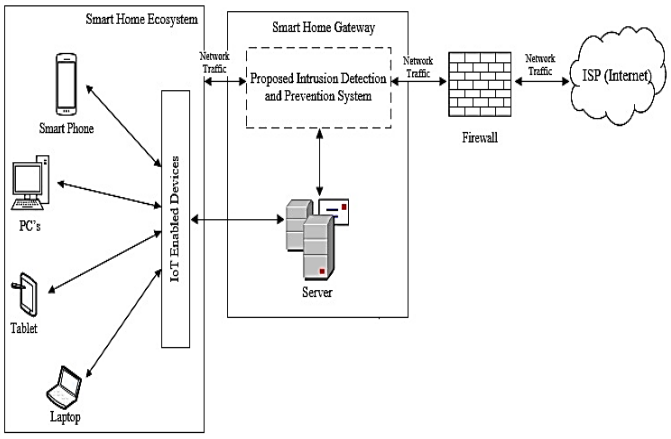

Figure 1. Architectural Design for Smart Home Network

Figure 1 captures the architectural design of smart home network. It depicts various communication between the devices and the network. However, this study is restricted to developing an adaptive IDPS for smart home networks. Hence, it will be limited to developing mechanism to safeguard the network from cybercriminals. The architecture of our system is capture in Figure 2.

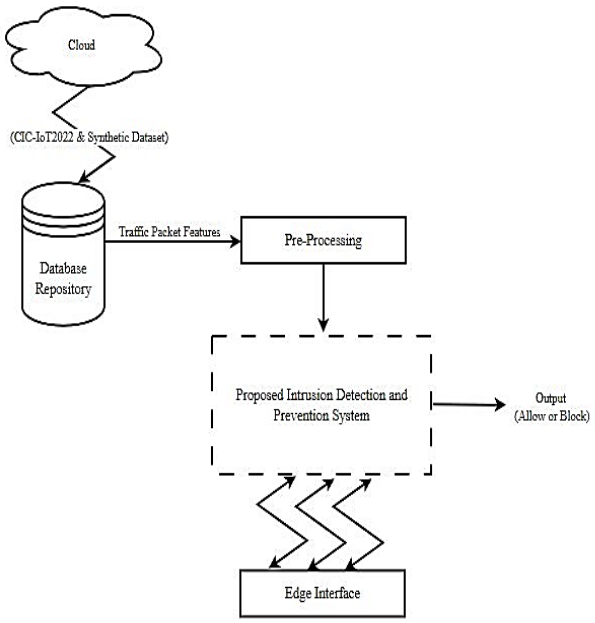

Figure 2. Architecture of the System.

Figure 2. represents a high-level architectural design of the system for intrusion detection and prevention. It captures how the selected (CIC-IoT2022 dataset) packet features will be pro-processed to an acceptable format of the model. Further, the pre-processed packet features will serve as input to the system. The output of the system will be a defuzzified value ranging from zero to one [0 - 1] representing either to allow or block the device or network traffic packet. It is worth mentioning that there are several additional modules embedded in the proposed Intrusion Detection and Prevention System module for this study. Hence, in the subsequent sessions of this paper, these modules or components and their functionalities will be further broken down into smaller pieces for better readability and understanding.

Pre-processing Module

This phase transforms the raw data file into a well suitable format for the training and testing of the system. The CIC-IoT2022 dataset was downloaded as "Packet Capture (PCAP)" file, and then converted to Comma-Separated Value (CSV) files using Wireshark. The essence of the pre-processing phase for cleaning the dataset, normalization, and feature selection. The Synthetic Minority Over-Sampling Technique (SMOTE) was employed used balancing the collected CIC-IoT2022 dataset. The min-max method was adopted for the normalization of the collected dataset. Five network traffic packet features were selected for this study. Further, the dataset was divided into 70% and 30% with regards to training and testing respectively.

Intrusion Detection and Prevention System (IDPS) Module

Embedded in the proposed IDPS module are the device management and the Hybrid Case-Based Neuro-Fuzzy System (HCBNFS) modules. The device manager module is deployed for proper management of trusted network devices, while the HCBNFS was deployed as a novel hybrid machine learning technique for detecting and preventing anomalies or malicious traffics on the network. We adopted the anomaly-based method for the development of the proposed IDSP. This method is adopted because of its effectiveness in detecting unknown attacks on the network. The schematic overview of our IDPS model is captured in Figure 3.

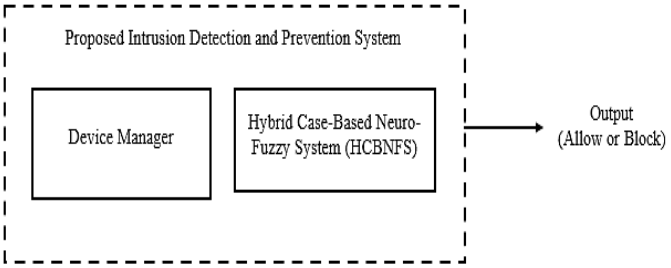

Figure 3. Intrusion Detection and Prevention System

Device Manager Module

Management of trusted network devices is very critical for ensuring security and reliability for modern network. This study deploys SQL server as a comprehensive model for managing network device credentials. As new IoT device to join the network, it sends its information, such as device ID, type, and initial authentication credentials to the device manager module. The device manager module now stores this information in the SQL Server database. The information includes device metadata, ownership details, and any other relevant attributes. For authentication, when devices try to connect to the network, it sends its authentication credentials (e.g., username/password or certificate) to the SQL server for authentication. These credentials are validated against the database's stored data by the server. The device is authenticated and granted network access if the credentials agree; otherwise, access is denied.

Hybrid Case-Based Neuro-Fuzzy System (HCBNFS) Module

The proposed novel HCBNFS technique combines the abilities of Case-based Reasoning (CBR) to apply previous solution to new similar problem together with the adaptability nature of the Neuro-Fuzzy Inference System (NFIS) to adapt to the evolving nature of cyber-attacks. Embedded in the proposed HCBNFS model for this study is an:

i. Intrusion Detection System (IDS) using CBR as the model's detection engine, and
ii. A tuning factor which serves as an Intrusion Prevention System (IPS) using NFIS.

The schematic and well detailed overview of the proposed Hybrid Case-Based Neuro-Fuzzy System (HCBNFS) module for this study is captured in Figure 4.

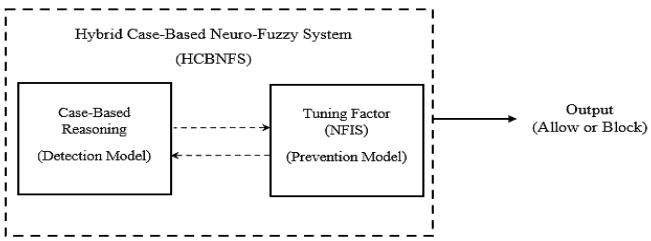

Figure 4. Hybrid Case-Based Neuro-Fuzzy System (HCBNFS) Module

Case-Based Reasoning (CBR) Module

This module serves as a detection engine to the HCBNFS. It is used to classify the input network traffic packet features into several classes based on the similarity to their predefined classes. Further, all future behaviour will be compared to the available cases in its case-base repository. If any suspicious traffic pattern, or anomalies is identified, it will label it as potential threat and log it to the tuning factor (NFIS) for further investigation and possible actions. For this study, a case is represented as a certain experienced intrusion attempt or situation which had been captured and acquired knowledge that can be applied again in order to solve future intrusion problems tailored to the smart home network ecosystem. A case contains both the possible intrusion's problem specification (attribute or packet features) together with its recommended solution (Allow or Block). The typical case structure used for this study is captured in Figure 5. Here the case is divided into two parts with the problem part labelled "Possible Intrusion Attempt" and the recommended solution is labelled "Solution".

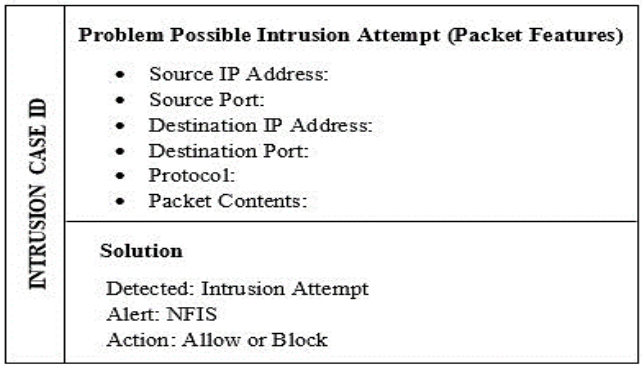

Figure 5. Basic Case Structure of the CBR Module

For this study, the problem specifications are the basic network packets features of the given base case, while the recommended solution is either to allow the traffic packet or to deny and block its access to the smart home network. The CBR module in this study goes through several processes to accomplish a given task at hand. These processes or steps are commonly referred to as the "4R's" of the CBR circle. The pictorial representation of the consecutive processes of the CBR module for this study, the is captured in Figure 6.

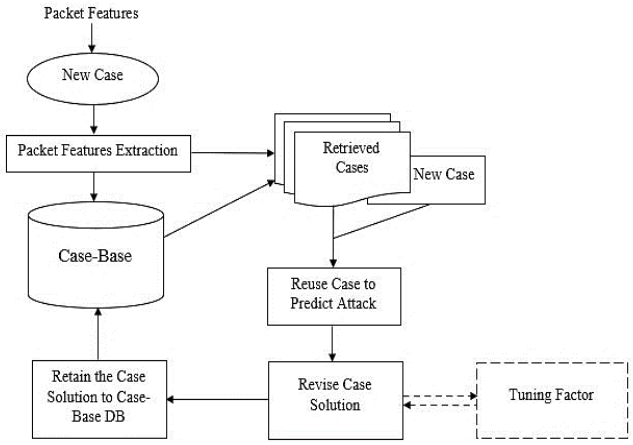

Figure 6. CBR Processes for Detecting Intrusion

Tuning Factor for the Reverse Phase of the CBR

We employed for the NFIS as a tuning factor for the revise phase of the CBR. The NFIS is as a fuzzy controller agent with the learning capabilities of neural network. The selected input features or variables will be fuzzified to various degrees given their degree of membership. The formula for fuzzification is as follows:

$$\mu A_i(x) = f(x) \quad (1)$$

Where μ is membership of the input variable x, A_i is linguistic term, $\mu A_i(x)$ is the degree of membership of x to the linguistic term A_i, and $f(x)$ is the membership function for the linguistic term A_i. The membership partition ranges are determined based on the characteristics of the input variables and the desired granularity of categorization. In this study, the ranges are chosen to provide meaningful distinctions between different levels of the input variables while keeping the partitioning relatively simple and intuitive.

This study employed the Class-C IPv4 format, which is represented by four decimal numbers separated by periods (e.g., 192.168.1.0). It is important to comprehend the structure of IPv4 addresses for accurate fuzzification for this study. The Class-C IPv4 consists of four octets, each ranging from 0 to 255. Note, that the fuzzification process will involves assigning each octet of the IP address to one of the linguistic variables (Low, Medium, or High), depending on its numerical value. Therefore, the fuzzification of IP address field involve converting the crips IP addresses into predefined fuzzy sets. This is achieved by dividing the crips IP address range into the stated linguistic variables based on the four different octets of IP addressing. For this study, "Low" represents IP addresses below a threshold value (0-63), Medium" for IP addresses within a moderate range (64-191), and "High" for those IP addresses above another threshold (192-255). This division enable the system in handling imprecision and variability. This aids the IPS module to reason about the incoming IP addresses in a more flexible way.

In this study, a combination of linguistic variables (High, Medium, or Low) obtained from the values in each of the four octets is used to express the IP address variable. Using a preset mapping or rule set, the linguistic variables from each individual octet are aggregated to determine the linguistic value for the entire IP address.

The overall linguistic variable for the IP address could be mapped to a specific combined value (such as "High-Medium-Low-High") based on predefined rules, for instance, if the first octet has a "High" value, the second octet has a "Medium" value, the third octet has a "Low" value, and the fourth octet has a "High" value. Then, using the specified criteria, the IP address's overall linguistic variable might be mapped to a certain combined value (such as "High-Medium-Low-High") according to the system's predetermined set of regulations.

Table 1. capture the input variables to the NFIS together with their respective membership functions.

Table 1. Input Variable for the proposed NFIS.

Input Variable	Membership Function (MF)	Range
Source IP	Low, Medium, and High	Low = 0-63, Medium = 64-191, and High = 192-255.
Destination IP	Low, Medium, and High	Low = 0-63, Medium = 64-191, and High = 192-255.
Source Port	Low, Medium, High	Low = 0-16383, Medium =16383-32767, and High = 32767-65535.
Destination Port	Low, Medium, High	Low = 0-16383, Medium = 16383-32767, and High = 32767-65535.
Protocol	Low, and High	Low = Unknown, and High = Known

The Takagi-Sugeno fuzzy inference was employed for specifying actions to every situation in the network depending on the degree of membership given the rules for detecting possible intrusions on the home network. The model was be trained offline (training the system with collected dataset) using the Backpropagation (BP) algorithm. The schematic overview of the NFIS for this study is captured in Figure 7.

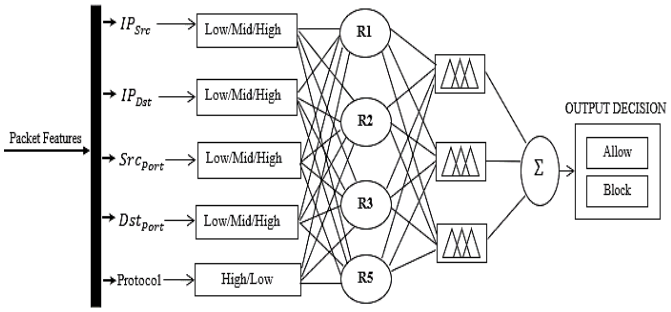

Figure 7. NFIS Module of the proposed HCBNFS.

The weighted average defuzzification method was be adopted to transform the fuzzy output (0-1) into a crisp or meaningful linguistic description (Allow or Block) based on the mapping of their output MF. The formular for Weighted Average defuzzification method is obtained as:

$$D^* = \frac{\sum(\mu(\bar{a}) \times \bar{a})}{\sum \mu(\bar{a})} \quad (2)$$

Where D^* is the crisp output value obtained by defuzzification. $\mu(\bar{a})$ is the MF of the fuzzy output value $\bar{a}$, which describes the degree of the membership of the input value to the linguistic variable $\bar{a}$. Σ denoted the algebraic summation of all elements with MF in the fuzzy set, and $\bar{a}$ is the input value to the MF.

Lastly, we classify the defuzzified crips numeric values for the HCBNFS model to represent final decision either to allow or block the incoming traffic using the follow equation.

$$FinalDecison_{Output} = \begin{cases} Normal & 0 \le y \le 0.3 \\ MiTM & 0.3 \le y \le 0.5 \\ Scanning & 0.5 \le y \le 0.7 \\ DoS & 0.7 \le y \le 0.9 \\ MiraiBotnet & 0.9 \le y \le 1 \end{cases} \quad (3)$$

In equation (3), the boundaries are defined such that each range is exclusive of the upper bound and inclusive of the lower bound. This means that if the value of y falls exactly on a boundary, it will be assigned to the corresponding category.

Technically, the novel HCBNFS in this study combines the strengths of CBR and NFIS to enhance the detection and response capabilities of the proposed model. The mathematical representation of flow between the CBR and the NFIS is given as follows:

Let C_i represent a historical case in case-base, N denotes the number of case instances in the case-base, and *EDSim* denote the Euclidian distance metric for retrieving similar cases in the systems case-base. S be the severity associated with C_i. X_1 and X_2 be the source and destination IP address respectively. X_3 and X_4 be the Source and Destination Ports. Finally, X_5 be the network Protocol. The input and output of the CBR module is given as follows:

$$InputPF_{CBR} = (x_1, x_2, \ldots, x_n) \quad (4)$$

$$CBR = \frac{1}{N}\sum_{i=1}^{N} EDSim(InputPF_{CBR}, C_i) \times S \quad (5)$$

The output of the CBR module will represent the likelihood of an intrusion or attack for known traffic patterns. However, unknown network traffic packets will be further investigated using NFIS as a tuning factor for the reverse phase of the CBR module. For this paper, the Neuro-Fuzzy Inference System (NFIS) is employed as a complementary mechanism for handling unknown network packet traffics or cases.

The NFIS module dynamically adjusts the reverse phase of the CBR process by fuzzifying the input packet features parameters based on their degree of membership and employing the Takagi-Sugeno fuzzy inference techniques to analyse incoming traffic. This will enable the system to adaptively tune its response based on the uncertainty and impression inherent in detecting new network traffic patterns. In tuning the reserve phase of the CBR using the NFIS, the NFIS module takes the output of the CBR module which is a possible intrusion case or unknown traffic pattern as its input and produces an output. This can be mathematically expressed as follows:

$$InputPF_{NFIS} = NFIS\big(CBR(InputPF_{CBR})\big) \quad (6)$$

Where the $InputPF_{CBR}$ represents the input packet feature vector that serves as input to the CBR component of the HCBNFS. The CBR, represents the Case-Based Reasoning module, which takes the input packet feature $InputPF_{CBR}$ and processes it using past intrusion experiences stored in its case base. The output of this process will serve as a response or decision related to a likelihood of an intrusion occurring. The $NFIS\big(CBR(InputPF_{CBR})\big)$ part of equation (6) integrates the CBR output into the NFIS module. For this study, the output of the CBR component serves as input to the NFIS module. Finally, the $InputPF_{NFIS}$ represents the input packet feature vector for the NFIS module, which is the output of the CBR module.

The reverse phase of the CBR process is represented as follows:

$$CBR_{ReversePhase}: C \rightarrow NFIS \rightarrow Y \quad (7)$$

Equation (7) describes the reverse phase of the CBR process which integrates the NFIS, leading to the determination of the desired output "*Y*" from the NFIS ranging from 0-1 representing either to allow or block the traffic.

To retain the output of the NFIS *Y* in case base, the model will update the case representation to include the outcome produced by the NFIS. This can be represented using the following:

$$C = (X_C, Y_C) \quad (8)$$

Where C is the set of cases, and each case c will be represented as a tuple (X_C, Y_C). X_C represents the features associated with the case c, and Y_C represents the output produced by the NFIS module for case C.

The combination (X_C, Y_C) forms a complete case C, which pairs the input data X_C with the corresponding output solution Y_C. This pairing allows the system to learn from past cases and leverage the solutions or outputs associated with similar problem descriptions when encountering new traffic situations that appears malicious.

In this paper, the CBR module's case base CB is denotes a set or collection of individual cases. This is expressed as follows:

$$CB = (C_1(X_1, Y_1), C_2(X_2, Y_2), \dots\dots, C_n(X_n, Y_n)) \quad (9)$$

Where $C_1(X_1, Y_1)$ is represented as a tuple of the first case C_1, X_1 represents the problem description or input features of the case C_1, and Y_1 represents the output or solution associated with case C_1.

$C_2(X_2, Y_2)$ is represented as a tuple of the second case C_2, X_2 represents the problem description or input features of the case C_2, and Y_2 represents the output or solution associated with case C_2.

$C_n(X_n, Y_n)$ is represented as a tuple of the nth case C_n, while X_n represents the problem description or input features of the case C_n, and Y_n represents the output or solution associated with case C_n.

Finally, equation (10) and (11) is used for appending the outcome produced by the NFIS to the new network intrusion case, and for updating and the information on the CBR module case base respectively.

$$C_{New} = (X_{New}, Y_{New}) \quad (10)$$

$$CB_{New} = CB \cup \{C_{New}\} \quad (11)$$

Where C_{New} represents a new case that is added to the system's case base CB. X_{New} represents the problem description or input features of a new malicious situation or problem that the system has encountered and solved successfully. Finally, Y_{New} represents the solution (allow or block) that the system has derived or learned for the new problem described by X_{New}.

In this paper, the updated case base is represented as CB_{New}, which is the union of the previous case base CB and the new case C_{New}.

The output of the NFIS module will be served to refine the decision-making process of the CBR module. This update will enable the system to adapt and respond effectively to both known and unknown cyber-threats tailored to the network.

Sequence Diagram of the proposed system

We used a sequence diagram to illustrate the interactions between the actor (user or system) and the components of the proposed system. The process begins by the actor loading the dataset (packet features) to the system though the user interface. These features are used for formulation of the new problem which is fed into the CBR component of the proposed HCBNFS. The CBR module retrieves network intrusion cases in its case based, based on their similarities to the problem. If a retrieved case is exactly like then the new case, the CBR module will adopt the solution of the old case to solve the new intrusion problem. Otherwise, all unknown traffics will be investigated using the NFIS module of the HCBNFS module. The outcome produced by the NFIS module is then retained in the CBR module case base as the solution for the new problem. The sequence diagram for this study is capture in Figure 8.

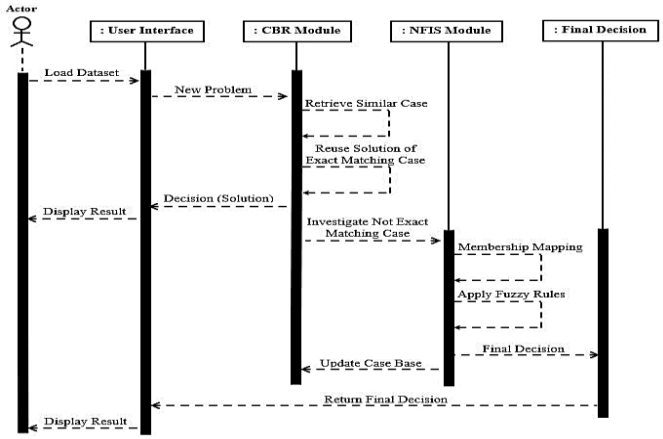

Figure 8. Sequence Diagram of the proposed system.

6. Results and Discussion

Presentation of Results

The results of the developed system when tested using the CIC-IoT2022 dataset are captured in Table 2.

Table 2. Results

SrcIP	DesIP	Protocol	SrcPort	DesPort	Output
192.168.0.16	192.168.0.13	TCP	49784	9020	0.2
104.118.134.215	192.168.0.24	TCP	433	56373	0.9
192.168.0.13	222.169.172.174	TCP	554	2760	0.7
172.217.25.99	192.168.0.14	UDP	433	54028	0.2
193.168.0.13	163.152.127.148	UDP	56361	10101	0.9
192.168.0.24	163.152.1.1	DNS	59784	53	0.9
192.168.0.23	224.0.0.251	DNS	5353	5353	0.6
192.168.0.16	192.168.0.13	ICMP	53182	9020	0.4
46.51.222.63	192.168.0.16	TCP	443	61865	0.4
192.168.0.24	139.150.252.50	HTTP	51586	80	0.9
163.152.1.1	192.168.0.24	UDP	53	59310	0.9
111.238.226.14	192.168.0.13	TCP	6588	554	0.7

3					
97.186.0.16	192.168.0.13	UDP	49784	9020	0.2
192.168.0.16	18.136.162.179	TCP	57217	433	0.2
108.177.97.189	192.168.0.14	UDP	443	50510	0.2
163.152.1.1	192.168.0.16	DNS	53	50788	0.9

The results of the system using the CIC-IoT2022 dataset are captured in Table 2. It captures list of ScrIP and DesIP, which represent the source and destination IP addresses, as well as the network protocol. It also includes the ScrPort and DesPort, which represent the source and destination ports, and finally the output membership values ranging from 0 to 1, with specific ranges corresponding to different types of traffic or attacks. (see equation3).

Table 3. captures the interpretations of the results together with the recommended action.

Table 3. Interpretation of the Results

SrcIP	DesIP	Protocol	Src Port	Des Port	Output	Interpretation	Recommended
192.168.0.16	192.168.0.13	TCP	49784	9020	0.2	Normal	Allow
104.118.134.215	192.168.0.24	TCP	433	56373	0.9	MiraiBotnet	Block
192.168.0.13	222.169.172.174	TCP	554	2760	0.7	DoS	Block
172.217.25.99	192.168.0.14	UDP	433	54028	0.2	Normal	Allow
193.168.0.13	163.152.127.148	UDP	563	10101	0.9	MiraiBotnet	Block
192.168.0.24	163.152.1.1	DNS	59784	53	0.9	MiraiBotnet	Block
192.168.0.23	224.0.0.251	DNS	5353	5353	0.6	Scanning	Block
192.168.0.16	192.168.0.13	ICMP	53182	9020	0.4	MiTM	Block
46.51.222.63	192.168.0.16	TCP	443	61865	0.4	MiTM	Block
192.168.0.24	139.150.252.50	HTTP	51586	80	0.9	MiraiBotnet	Block
163.152.1.1	192.168.0.24	UDP	53	59310	0.9	MiraiBotnet	Block
111.238.226.143	192.168.0.13	TCP	6588	554	0.7	DoS	Block
97.186.0.16	192.168.0.13	UDP	49784	9020	0.2	Normal	Allow
192.168.0.16	18.136.162.179	TCP	57217	433	0.2	Normal	Allow
108.177.97.189	192.168.0.14	UDP	443	50510	0.2	Normal	Allow
163.152.1.1	192.168.0.16	DNS	53	50788	0.9	MiraiBotnet	Block
192.168.0.16	31.13.76.16	TCP	56358	443	0.9	MiraiBotnet	Block
3.0.72.236	192.168.0.16	TCP	443	61854	0.9	MiraiBotnet	Block
210.124.177.97	192.168.0.23	TCP	9600	36762	0.6	Scanning	Block
192.168.0.15	192.168.0.13	TCP	33306	8194	0.6	Scanning	Block
192.168.0.13	192.168.0.15	TCP	44443	33306	0.6	Scanning	Block
192.168.0.15	192.168.0.24	TCP	35878	700	0.6	Scanning	Block

From Table 3, the membership value ranges are interpreted as the degree to which the system classifies the network traffic or activity by analysing them and assigning membership values or scores to the different types of traffic or attacks either to allow or block the traffic. Normal traffic is fuzzified between 0 and 0.3, MiTM attacks are fuzzified between 0.3 and 0.5, scanning is fuzzified between 0.5 and 0.7, DoS is fuzzified between 0.7 and 0.9, and Mirai Botnet is fuzzified between 0.9 and 1.

Performance Evaluation Metric

We employed various performance metrics to determine the efficiency of our model. These metrices includes the accuracy, precision, recall, and F1-score. The CIC-IoT2022 dataset was utilized for this study. However, only few percentages of this dataset were used for the purpose of testing of the system. The formula for calculating these metrics is given in the following equations:

$$IDPS_{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (12)$$

Where, $IDPS_{Accuracy}$ is the accuracy rate, TP represents the value of true positives, TN represents the value of true negatives, FP represents the value of false positives, and FN represents the value of false negatives.

Regarding to our developed IDPS for this study, the system's precision is determined by how many of the predicted intrusions cases were correctly predicted. It measures the number of true positive instances divided by the sum of true positive and false positives. This however represents the accuracy of the positive predictions made by our IDPS. The formula for precision is given as.

$$IDPS_{Precision} = \frac{TP}{(TP+FP)} \quad (13)$$

Where $IDPS_{Precision}$ is the precision of the IDPS TP and FP represent the values for true positives and false positives, respectively.

For this study, recall is regarded as the baseline for truth. That is, given all intrusion truth samples, how many of these intrusions were correctly captured. It calculates the ratio of correctly identified positive instances to the total of correctly identified positive instances and incorrectly identified negative instances. This will be used to represent the ability of the IDPS to correctly identify all relevant instances. The formula for Recall is given as follows.

$$IDPS_{Recall} = \frac{TP}{(TP+FN)} \quad (14)$$

Where $IDPS_{Recall}$ is the recall value of the developed IDPS, TP and FN represent the values for true positive and false negative respectively. The F1-score is a metric commonly used in classification tasks. It measures the developed IDPS accuracy, considering both precision and recall. It is commonly considered as a balance between precision and recall based on the system's goals and priorities. For this study, F1-score is computed using the following:

$$IDPS_{F1Score} = 2 \times \frac{IDPS_{Precision} \times IDPS_{Recall}}{IDPS_{Precision} + IDPS_{Recall}} \quad (15)$$

The computed values of the accuracy, precision, recall, and F-score for the developed IDPS is captured in Table 4.

Table 4. Computed values of the various performance metric.

Performance Metric	Value
$IDPS_{Accuracy}$	0.99
$IDPS_{Precision}$	0.995
$IDPS_{Recall}$	0.995
$IDPS_{F1score}$	0.995

Performance Evaluation of the System

We employed various standard metrics to evaluate the performance of our IDPS. The computed values of these metrics were captured in Table 4. Our IDPS achieved an accuracy rate of 99%, precision, recall, and F1-score of 99.5% respectively. Hence, it actively demonstrated that it could serve as an efficient security system for IoT smart home network in terms of classification of network traffic, and the detection and prevention of anomalies on the smart home network. The line graph in Figure 8 is used to capture the level of performance of the various metrics.

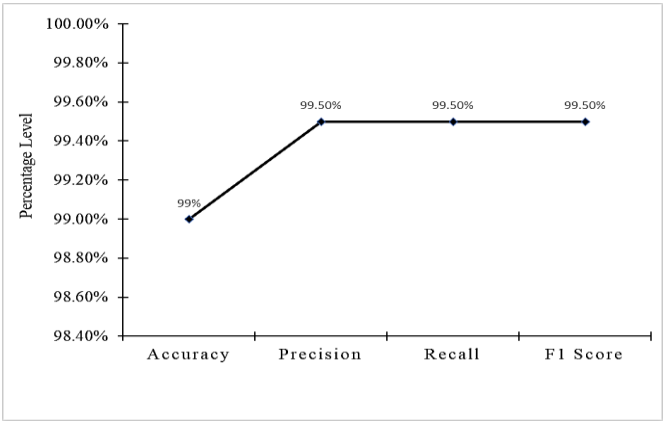

Figure 9. Line Graph Representation of the Performance Evaluation of the system.

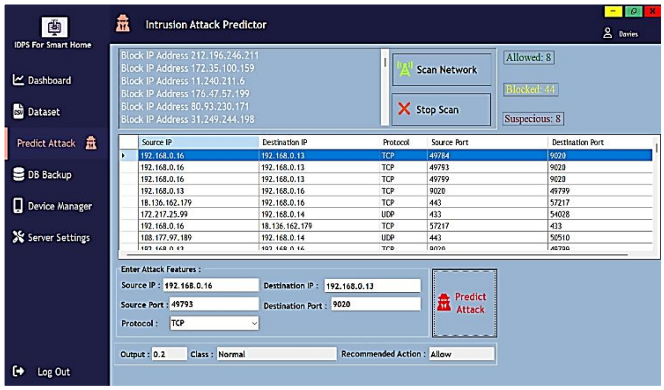

Fig. A: Attack Predictor Interface

7. Conclusion and Future Scope

In this study, a Hybrid Case-Based Neuro-Fuzzy System (HCBNFS) was presented for the design and development of a robust Intrusion Detection and Prevention System (IDPS) for IoT-based smart home network. The study combined the strengths of Case-Based Reasoning (CBR) and Neuro-Fuzzy Inference System (NFIS) techniques. The CBR component served as the primary intrusion detection engine, capable of detecting known traffic patterns, while the NFIS component was employed as a tuning factor to the reverse phase of the CBR, further investigating unknown network traffic patterns. This integration of CBR and NFIS was proposed as a novel approach to enhance the system's network classification and anomaly identification capabilities within the dynamic and complex IoT network environment.

The performance of the HCBNFS-based IDPS was evaluated using the comprehensive CIC-IoT2022 dataset, which provided both training and testing data. The developed IDPS achieved a detection accuracy rate of 99%, with precision, recall, and F1-score of 99.5%. Additionally, the system achieved a high prevention rate of 99%, successfully preventing 1,980 out of 1,990 detected intrusions. Furthermore, the system had a False Prevention Rate of 0%, indicating that no normal instances were incorrectly subjected to prevention measures. These prevention performance metrics, combined with the high detection performance metrics, suggest that the developed IDPS will be highly effective in both detecting and preventing intrusions or anomalies in the IoT smart home network ecosystem.

The proposed IDPS exhibited adaptability to changing patterns of behavior and the ability to detect anomalies with a high degree of accuracy, highlighting its potential for real-world deployment in IoT network environments. By combining the strengths of CBR and NFIS, the HCBNFS-based IDPS offers a robust and flexible solution for detecting and mitigating a wide range of intrusions and cyber threats in the rapidly evolving IoT landscape.

Future research efforts could focus on further refining the HCBNFS model's performance by exploring advanced techniques for feature selection, ensemble methods, or incorporating additional machine learning algorithms. Additionally, investigating the system's scalability and resource efficiency would be valuable in facilitating its deployment across diverse IoT network configurations and resource-constrained devices.

Conflict of Interest

Authors declare that they do not have any conflict of interest.

Funding source

None.

Authors' Contributions

Author-1 (Davies, Isobo Nelson): Conceptualized and developed the initial idea, overseeing the entire research process from inception to completion. Including conducting the implementation of the algorithms, supervised the experimentation process, analysed the results, and drafted the manuscript.

Author-2 (Taylor O. E.): Researched current existing literatures and formulated the research question for the study. He defined the research objectives and formulating the methodology.

Author-3 (Bennett E. O): Surveyed existing algorithms and technologies relevant to the research domain. He collaborated

with the lead author in designing experiments, interpreting results, and refining the theoretical framework.
Author-4 (Anireh V. I. E): Provided his expertise for the development process, enhancing the robustness and efficiency of the implemented solutions.
Finally, all authors reviewed and edited the manuscript.

References

[1] Ahanger, T. A., Tariq, U., Ibrahim, A., Ullah, I., & Bouteraa, Y. "IoT-Inspired Framework of Intruder Detection for Smart Home Security Systems". *Electronics*, Vol. **9**, Issue.**9**, pp.**1361**, **2020**.

[2] Alalade, E. D. "Intrusion detection system in smart home network using artificial immune system and extreme learning machine hybrid approach". *IEEE*, pp.**1-2**, **2020**.

[3] Almseidin, M., Al-Sawwa, J., & Alkasassbeh, M. "Anomaly-based Intrusion Detection System using Fuzzy Logic". *In the Proceedings of the 2021 International Conference on Information Technology (ICIT)*, pp.**290-295**, **2021**.

[4] Alrayes, F. S., Alshuqayran, N., Nour, M. K., Al Duhayyim, M., Mohamed, A., Mohammed, A. A. A., . . . Yaseen, I. "Optimal Fuzzy Logic Enabled Intrusion Detection for Secure IoT-Cloud Environment". *CMC-COMPUTERS MATERIALS & CONTINUA*, Vol.**74**, Issue.**3**, pp.**6737-6753**, 2**023**.

[5] Butt, N., Shahid, A., Qureshi, K. N., Haider, S., Ibrahim, A. O., Binzagr, F., & Arshad, N. (2022). "Intelligent Deep Learning for Anomaly-Based Intrusion Detection in IoT Smart Home Networks". Mathematics, Vol.**10**, Issue.**23**, pp.**4598**, **2022**.

[6] Farhin, F., Sultana, I., Islam, N., Kaiser, M. S., Rahman, M. S., & Mahmud, M. "Attack detection in internet of things using software defined network and fuzzy neural network". *In the Proceedings of the 2020 Joint 9th International Conference on Informatics, Electronics & Vision (ICIEV)*, pp. **1-6**, **2020**.

[7] Imtiaz, S. I., Khan, L. A., Almadhor, A. S., Abbas, S., Alsubai, S., Gregus, M., & Jalil, Z. "Efficient Approach for Anomaly Detection in Internet of Things Traffic Using Deep Learning". *Wireless Communications & Mobile Computing*, pp.**1-15**, **2020.**

[8] Johannesson, P., & Perjons, E. "An Introduction to Design Science" (2nd ed.). *Springer*: **978-3-030-7813**, **2021.**

[9] Kponyo, J. J., Agyemang, J. O., & Klogo, G. S. "Detecting End-Point (EP) Man-In-The-Middle (MITM) attack based on ARP analysis: a machine learning approach". *International Journal of Communication Networks and Information Security*, Vol.**12**, Issue.**3**, pp.**384-388**, **2020.**

[10] Marimuthu, D., Rao, G. R. K., Mehbodniya, A., Mohanasundaram, D., Sundaram, C. K., Maria, A. B., & Mani, D. "Mathematically Modified Adaptive Neuro-Fuzzy Inference System for an Intelligent Cyber Security System". *SN Computer Science*, Vol.**4**, Issue.**5**, pp.**453**, **2023**.

[11] Moudni, H., Er-rouidi, M., Mouncif, H., & El Hadadi, B. "Black Hole attack Detection using Fuzzy based Intrusion Detection Systems in MANET". *Procedia Computer Science*, Issue.**151**, pp.**1176-1181**, **2019**.

[12] Rajput, P. K., & Sikka, G. Multi-agent Architecture "Approach for Self-healing Systems: Run-time Recovery with Case-based Reasoning". *Concurrency and Computation: Practice and Experience*, Vol.**35**, Issue.**1**, **2023.**

[13] Raushan Kashypa."Smart Home Design Using IoT", *International Journal of Computer Sciences and Engineering,* Vol.**8**, Issue.**1**, pp.**146-150**, **2020**.

[14] Richa, M., Sakshi, K., & Shubham, G. "Detecting Various Intrusion Attacks using A Fuzzy Triangular Membership Function". *International Research Journal of Engineering and Technology (IRJET)*, Vol.**9**, Issue.**1**, pp.**932-944**, **2022**.

[15] Sajjad, D., Hassan, M., Priscilla, K. D., Zohourian, A., Kelvin, A. T., & Ali, G. A.. "CIC-IoT-Dataset2022 [Dataset]". *Kaggle*, **2022**.

[16] Suhasini, V. and Avinash, Y. "Artificial Intelligence Powering Internet of Things", *International Journal of Computer Sciences and Engineering*, Vol.**7**, Issue.**6**, pp.**449-456**, **2019**.

[17] Tariq, N., Asim, M., Khan, F. A., Baker, T., Khalid, U., & Derhab, A. "A blockchain-based multi-mobile code-driven trust mechanism for detecting internal attacks in internet of things". *Sensors*, Vol.**21**, Issue.**1**, pp.**23**, **2020**.

AUTHORS PROFILE

Mr. Davies, Isobo Nelson pursued B.Sc. degree in Computer Science at Kwame Nkrumah University of Science and Technology, Kumasi, Ghana in 2013.
M.Sc. & PhD (in View) at the Rivers State University, Port-Harcourt, Rivers State, Nigeria 2019 and 2024 respectively. He is a researcher, a member of the Computer Professionals of Nigeria (CPN), and a PhD student at the said university. He has published four (4) research papers in both local and international journals. His research works focus on Artificial Intelligence, Machine Intelligence Systems, and IoT Networks.

Dr. O. E. Taylor pursued B.Sc. degree in Computer Science at Rivers State University, MSc at the University of Ibadan, & PhD at the University of Port Harcourt. He is currently an Associate Professor & a Lecturer in the Departmentof Computer Science, Rivers State University, Port-Harcourt. He is a member of the Computer Professionals of Nigeria (CPN). He has published over 50 research papers in reputed international journals. His research works focuses on Machine Intelligence Systems, Cyber Security, & IoT.

Dr. V.I.E Anireh pursued B.Sc. degree in Computer Science at the University of Nigeria. MSc, & PhD at the University of at the University of Port Harcourt. He is currently an Associate Professor and a Lecturer in the Department of Computer Science, Rivers State University, Port-Harcourt. He is a researcher fellow and a member of IEEE. He has many scholarly publications in both local and international journals. His main research work focuses on Artificial Neural Networks, Machine Learning, Computer Networks, IoT, & Big Data.

Dr. E. O. Bennett graduated with a B.Sc. degree in Computer Science from Rivers State University, Rivers State, Port Harcourt, Nigeria in 1998, MSc and PhD at the University of Port Harcourt in 2008 and 2014 respectively. Currently he is an Associate Professor & a Lecturer in the Department of Computer Science, Rivers State University, Port-Harcourt. He is a member of the Computer Professionals of Nigeria (CPN). He has published over 50 research papers in reputed international journals. His research works focuses on Algorithms, Parallel, Distributed & Intelligent Computing.

