International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
Article Received: *25 December 2022* ***Revised:*** *12 January 2023* ***Accepted:*** *20 February 2023*

An Efficient Approach to Detect the Attacks in SDN based 5G Network Using fuzzy based XG Boost Algorithm

Kishore Malladi[1]
Ph.D Scholar
Acharya Nagarjuna University,Guntur
Email:kishoremalladhi1@gmail.com

Priyanka Mekala[2]
Asst.Professor
C R Rao AIMSCS,University of Hyderabad
Email:mekalapriyanka@gmail.com

Supriya Goel[3]
Asst.Professor
C R Rao AIMSCS,University of Hyderabad
Email:goelsupriya03@gmail.com

Neelima Guntupalli[4]
Asst.Professor
Acharya Nagarjuna University,Guntur
Email:neelima.guntupalli80@gmail.com

Abstract: The fifth-generation (5G) network has emerged as a next-generation wireless network in recent times. Software-Defined Networking (SDN) is among of the most rapidly expanding network designs that enables an intelligent and programmed control of network configuration to improve the performance of 5G networks. With the increase in the diversity of 5G networks, there is a growing concern about the security of these networks. When unauthorized users introduce malicious attacks, the performance of the SDN based 5G networks is adversely affected. In view of the growing adaptability of SDN based 5G networks various researchers have investigated the need for an effective technology which can identify security attacks in 5G networks. Despite the availability of different security frameworks for preventing malicious attacks in SDN, it is highly challenging to secure the SDN controller. This research intends to propose an efficient security approach to identify the SDN based attacks in 5G networks using machine learning (ML) algorithms. This study proposes the application of a Fuzzy based XGBoost algorithm to strengthen and simplify the network management process and to secure SDN based 5G networks. The findings indicate that the integration of SDN with the ML algorithm classified different network traffic patterns with high accuracy and improved the attack detection performance in 5G communication systems.

Keywords: *Software-Defined Networking Architecture, Machine Learning, XGBoost Algorithm, NSL-KDD dataset, UNSW-NB15 dataset, Fuzzy c-means clustering Algorithm*

1. Introduction

The next-generation wireless networks are gaining huge significance in terms of enhancing the connectivity between various advanced multi-communication systems. Traditional network architectures are not capable of addressing different network requirements such as better accessibility, virtualization, high bandwidth, cloud computing, dynamic management, reliable connectivity and network stability [1]. Software-defined networking (SDN) systems have dynamic, controlled, flexible and programmable architectures which

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
Article Received: *25 December 2022* ***Revised:*** *12 January 2023* ***Accepted:*** *20 February 2023*

makes them one of the potential alternatives to traditional architectures [2]. In general, the architecture of SDN incorporates "data planes, control planes and application planes". The data plane comprises physical devices such as routers and switches which are programmed and controlled using the control plane [3] [4]. The control plane controls the operation of communication devices positioned on the data plane. The controller serves as the SDN's brain, which is also located on the data plane. The communication devices transmit the data packers which depend upon the controller-specified protocols. Correspondingly, the communication of the devices placed on the network architecture is also controlled by the controller. SDN allows the execution of diversified applications such as virtualization, autonomous vehicles, traffic engineering, fine-grained access control etc [5]. SDN is considered advantageous in various commercial applications [6] [7] [8]. Due to its superior attributes, SDN is extensively implemented in 5G communication systems [9]. Security is one of the prominent aspects of SDN which affects its performance and effectiveness [10] 11]. However, it is highly challenging to deal with security issues in SDN-based 5G networks due to the diversity and vibrant nature of the network architecture. In addition, the centralized infrastructure of SDN makes it susceptible to different types of network attacks. Some of the prominent security attacks that affect SDN systems are Denial of Service (DoS), Distributed DoS (DDoS) [12], botnet, worm propagation [13] etc. Several researchers have discussed the implementation of different security frameworks for securing SDN controllers. The existing security frameworks periodically collect the statistics of network data from the forwarding plane of the SDN system in a structured manner i.e., using OpenFlow. The statistics of network information are classified using classification algorithms to detect malicious entities in the system architecture. If any malicious entity or anomaly is detected in the network, the security technique demands that the controller modify the data plane to prevent the attacks [14]. Another prominent type of security attack or threat used by the attacker is the adoption of moving target defense (MTD) algorithms [15]. MTD algorithms make security attacks on the SDN-based 5G system more complex by modifying or tampering with the important attributes of the system [16]. In conventional network architectures, it is not crucial to implement MTD algorithms since it is a complicated process to develop a centralized authority to protect each device in the system. In SDN systems, owing to the SDN controller's centrality, it is significantly easy to eliminate these assaults.

In recent years machine learning (ML) based techniques for securing SDN architecture for 5G networks have gained significant attention among researchers [17] [18]. The main idea behind increasing the implementation of ML algorithms for SDN security is its superior attack detection and classification ability, performance efficiency and high accuracy. Combining SDN with machine learning not only allows programmability but also improves the attack detection accuracy in 5G communication networks. ML algorithms along with SDN play a prominent role in classifying different network traffic patterns with high accuracy. The combined approach can operate with a huge number of resources and may offer solutions for security issues in SDN-based 5G networks.

Considering the advantages of the SDN and ML algorithms, this paper presents an ML-based XGboost (Extreme Gradient Boosting) algorithm for detecting different security attacks in SDN-based 5G networks.

The following are the study's primary contributions:

- This study provides a hybrid approach which combines ML based XGboost algorithm with a Fuzzy c-means clustering approach for securing SDN based 5G network system.
- The presented approach is designed based on a feature selection and classification mechanism to enhance the accuracy of attack detection.
- The suggested attack detection model monitors the network traffic continuously and identifies suspicious activities with the traffic at early stages of the attack.

The remaining paper is structured as follows: Section II elaborates the existing works related to the security of SDN based 5G network systems. Section III illustrates the network architecture of SDN based 5G systems and briefs the security problems in these systems. Section IV discusses the implementation of the proposed framework and briefs the research methodology. Section V provides the findings of the simulation analysis and Section VI finishes the study with significant research findings and future directions.

2. Related Works

SDN effectively manages and controls the network configuration and data transmission in a dynamic and centralized manner. Due to their potential ability to transform wireless communication in the 5G era, SDN aims to enhance the quality of service and enable reliable data transmissions in 5G systems [19]. In this setting, the integration of SDN with 5G networks has garnered a lot of attention among researchers. The security of SDN-based 5G architecture is one of the prominent aspects that need to be addressed to

increase the standard and reliability of service [20]. SDN plays a prominent part in strengthening the reliability and stability of security approaches. It enhances the controllability, and security of the features of the SDN architecture such as centralized control, programmability, and effective network traffic management. As discussed in the works [21] [22] [23], machine learning algorithms when adopted for attack detection allow the analysis of SDN controllers where different types of security attacks are increasing in SDN-based 5G networks.

A Support Vector Machine (SVM) based attack detection approach is discussed [24]. A discrete scalable memory (DSM) along with SVM not only detects the threats in the SDN architecture but also mitigates the attacks using a feature extraction technique. The essential features are extracted using a semantic multilinear component analysis method. The DSM-SVM classifier can predict the target and accurately identify the attack. Results show that the classifier achieves a phenomenal accuracy of 99.7% for the training and tested data collected from the KDD dataset. Implemented different ML classifiers such as "Naive Bayes, k-nearest neighbor (KNN), SVM, Random Forest (RF) and Logistic Regression (LR)" for identified DDoS attacks in SDN enabled Internet of Things (IoT) systems [25]. The ML classifiers were accompanied by the first layer of the multilayered feed forwarding (MLFF) network. The proposed approach employs a remote SDN controller for mitigating the identified DDoS attacks that affect the OpenFlow switches modify the configuration of the network and prevent legitimate hosts from being affected. Results of the experimental analysis show that the MLFF network with different ML algorithms achieves superior accuracy of attack detection with a low false alarm rate (FAR). Intends to identify appropriate ML algorithms for strengthening the security of SDN controllers affected by DDoS attacks [26]. The proposed approach integrates three main aspects namely a dynamic threshold technique, adaptive bandwidth and ML model and focuses mainly on mitigating the DDoS attacks in SDN controllers. An effective ML algorithm known as Extreme Gradient Boosting (XGBoost) was used to achieve a higher attack detection accuracy and enhance the performance of the overall network. ML techniques can accurately classify the data patterns based on the network traffic and predict the threats and attacks affecting the performance of SDN systems. AnML classifier for identifying the anomalies in the SDN environment was discussed [27]. The ML classifier was tested on a larger dataset wherein the data was modeled and the attacks were identified using a signature-based intrusion detection system (IDS). Each feature vector was created against the data pattern which was given as input to the ML classifiers for predicting anomalies. The effectiveness of the ML classifier was tested using a cross-validation approach. Despite the availability of several techniques, there is a great scope for research in this domain since the existing models do not consider the varying data patterns and dynamic features of heterogeneous 5G networks. The heterogeneity and dynamic attributes of SDN-based 5G networks can hurt their detection and classification accuracy. Hence, there is a strong demand for an efficient model which can achieve accurate attack detection performance.

3. SDN Based 5G Network Architecture

The SDN architecture enhances the computational efficiency of the network systems in terms of scalability, and provides a communication interface that connects different applications with smart devices. In conventional network infrastructure, the application has to program the operation of hardware devices to communicate with other devices. If the configuration of the hardware devices are changed externally (by manufacturer), then the actual applications might not be able to communicate with these devices due to different Application Programming Interface (APIs). The change in APIs makes it complicated for the network to program the hardware devices. The SDN framework is configured to overcome this problem [14].

The 5G network design is based on SDN as illustrated in figure 3.1. As observed from the figure, the SDN architecture consists of three important components namely "data plane, control plane and application plane". Compared to conventional network infrastructure that combines data and control into a single application plane, the SDN architecture distinguishes the data and control into two different planes. As the control plane handles the hardware devices automatically using SDN controllers, the separation between the data plane and the control plane helps the system to effectively control the network configuration using a unit interface known as smart gateways (SGW). This makes the system robust against the effect of hardware changes (change due to APIs) which occur in conventional network infrastructure. Similarly, when the devices in the network architecture want to communicate with the applications, the unit interface in the SDN architecture will act as a medium to transfer the information. This overcomes the problem of scalability caused due to the difference in the hardware devices. Due to these advantages, SDN architecture is considered to be the most prominent design aspect in the 5G network systems.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
Article Received: *25 December 2022* ***Revised:*** *12 January 2023* ***Accepted:*** *20 February 2023*

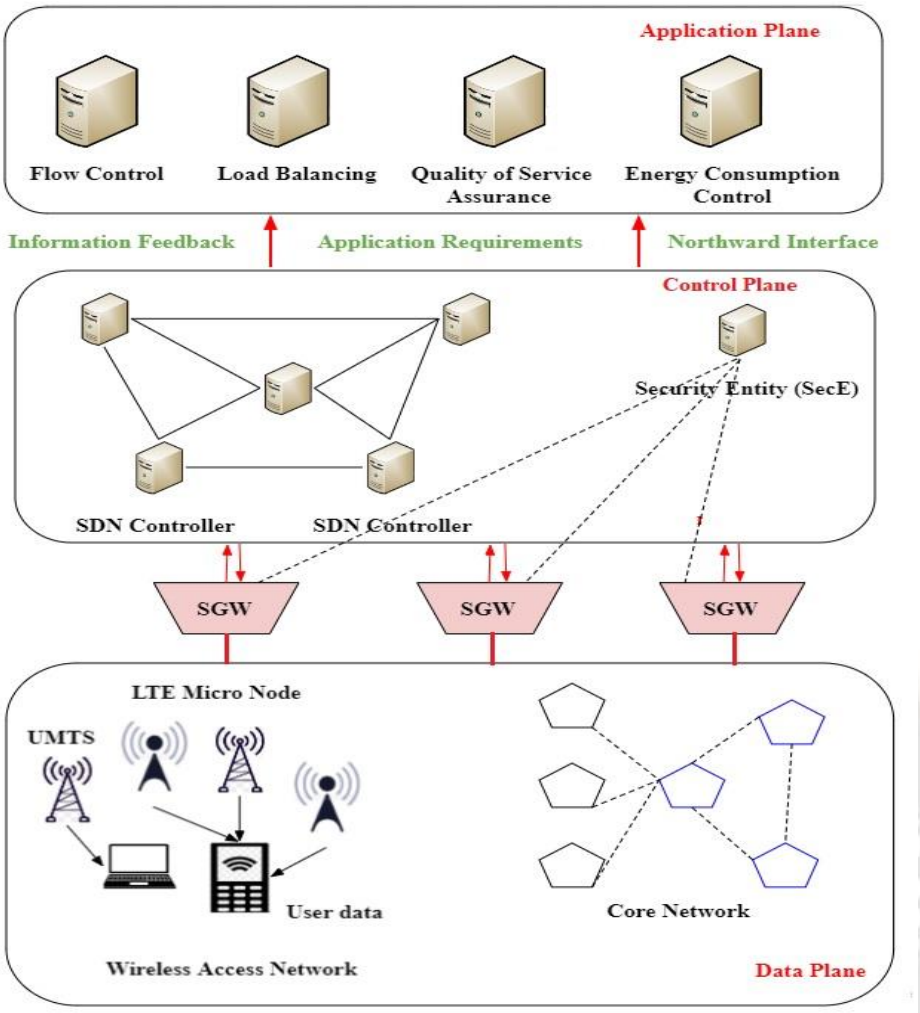

Figure 3.1 SDN architecture

Though SDN architecture can significantly enhance the scalability of conventional 5G network systems, the centralized control approach and programming of hardware devices using unit interface introduces security problems in 5G systems. The communicating medium of SDN is highly susceptible to security attacks. Some of the prominent reasons for the vulnerability of SDN towards security attacks are as follows [28] [29] [30].

- Lack of trust and poor authentication of the controller and the applications which increases the threat of spoofing attacks by spoofing the API messages.
- Regular revisions of data in 5G networks raise the danger of security risks owing to increased mobility and optimum utilisation of resources based on periodic traffic conditions.
- Ineffective authorization can result in the introduction of malicious entities into the device and permits improper entry to gain illegal access to the application. The user must be authorized before requesting a service to the application in order to ensure the safety of the SDN based 5G systems.
- Lack of encryption of network traffic data among the controller and the devices which can lead to eavesdropping. During eavesdropping, the attacker eavesdrops on the data exchanged between the users and access the authenticated information from the sender.

In addition to these issues, potential attacks such as “Denial of Service (DoS), Distributed DoS (DDoS) attacks, backdoor attacks, exploits, fuzzers and worms” affect the security of SDN based 5G systems. This research aims to address these problems using a ML based attack detection approach.

4. ML based Attack Detection Approach in SDN Based 5G Network

This research implements a Fuzzy based XGboost algorithm for attack detection in SDN based 5G network systems. Here,

the XGboost algorithm is integrated with a Fuzzy C-means clustering algorithm. The proposed attack detection approach will be designed based on feature selection and classification mechanism. The selection of appropriate numbers of features reduces the execution time of the ML algorithm and improves the attack detection accuracy. The inclusion of redundant and inappropriate features affects the computational performance and increases detection time. In addition, the irrelevant features will have a negative impact on the attack detection accuracy. Hence, it is important to select only relevant features to maximize the computational performance of the ML approach. Here, the features are selected for each specific type of attack and the particular characteristics are employed to teach the fuzzy based XGboost classifier for attack detection.

4.1 Data Collection

The information needed for the attack identification experimental analysis approach is obtained from the "UNSW-NB15 attack detection dataset". The "UNSW-NB15 dataset" is a publicly available dataset which consists of different attack related data [31]. The dataset consists of 2.5 million records which are distributed in .CSV format and are split into 49 features and consists of both flow and packet-based features. The data in this dataset are categorized into four sets namely content, basic, flow, and time [32]. The data in the dataset is labeled as normal and attack type. Compared to "NSL-KDD dataset", the "UNSW-NB15" is a new dataset which contains novel and advanced low footprint attack types and hence it is used extensively for testing the classifiers. Also, the amount of redundant attacks and unbalanced traffic data is less in UNSW-NB15 compared to other conventional datasets.

4.2 Data Preprocessing

Preprocessing is an important stage in the attack detection and classification process. In this stage, the data is processed to eliminate redundant information and uncertainties from the input data. Uncertainties such as additive noise, missing values, null values etc affect the detection accuracy. Hence preprocessing plays an important role in achieving better performance. During the preprocessing stage, the raw data is prepared and is made suitable for classification purposes. The phases in the preprocessing process are as follows:

- ***Finding missing data:*** Missing data occurs when there is no information in one or more data columns. It is also referred to as Not Available (NA).
- ***Encoding categorical data:*** The categorical data which is composed of alphabets and string formats is converted into numeric format using a different encoding process. In this research a label encoding process is used for encoding the categorical data.
- ***Removal of highly correlated features:*** Correlated features must be removed from the data since it raises the storage concerns and increases the execution speed of the process. In general, an effective property is one that is associated with the class and is not repetitive with any other features. The use of correlation as a metric to find relevant characteristics will aid in the selection of appropriate features from the feature collection. If two features are highly correlated, then one of the features must be removed in order to overcome the problem of multicollinearity.
- ***Splitting dataset into training and testing dataset:*** The data is split into two subsets namely training data and testing data. The training subset is used to train and fit the ML algorithm and the testing subset is used to evaluate and validate the performance of the ML algorithm.
- ***Feature Scaling:*** This step is applied to independent features of the dataset. Scaling of features helps in data normalization and fits the data into a particular range.

4.3 Feature Importance using Random Forest:

Several techniques are used to identify important features from the input data. Determining the importance of the features is one the crucial steps in the implementation of ML model for different purposes. In this stage, the importance of features is determined and based on this the features are selected accordingly. The selection of key features minimizes the computational complexity and reduces the generalization error as a result of noise which is introduced due to least important features. In this research, the feature importance is measured using a ML based Random forest (RF) algorithm. The RF algorithm computes the importance of features as the average impurity from all the decision trees in the forest. The RF model is trained to measure the feature importance value and then visualize to select important features.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
Article Received: *25 December 2022* ***Revised:*** *12 January 2023* ***Accepted:*** *20 February 2023*

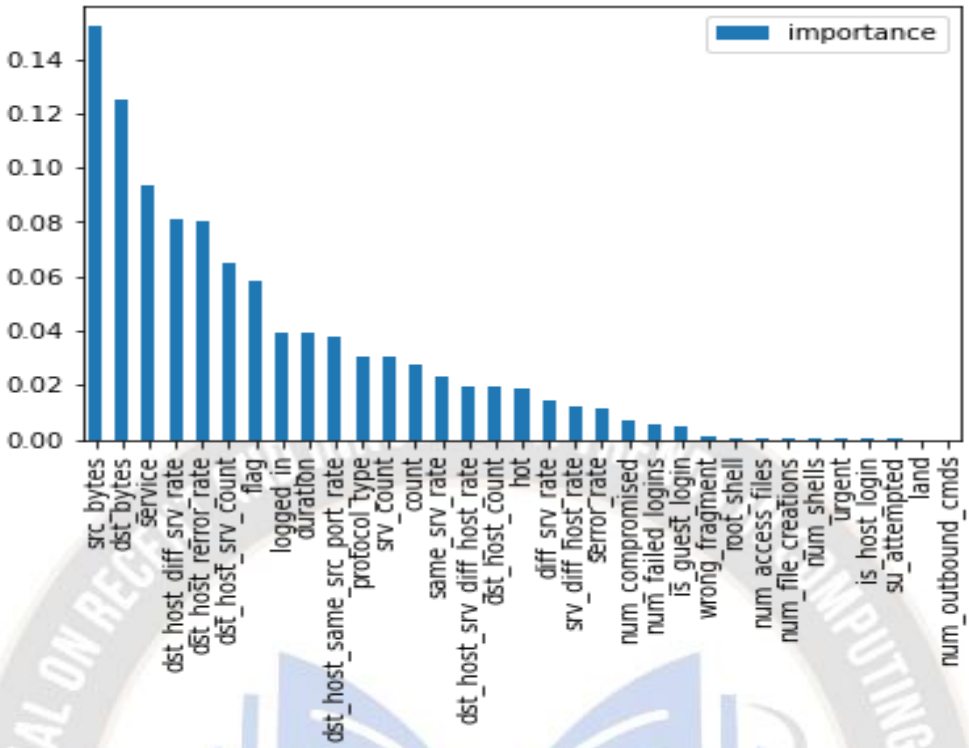

Figure 4.1. Feature Importance of UNSW-NB15

Fig. 4.1 illustrates the feature importance of the UNSW-NB15 dataset. It's excellent to recognize the main contributors because it lets you order characteristics according to their significance score, which indicates how much of an impact they have on prediction performance. Our feature of interest may be chosen, relationships can be found, and significant features can be found with the help of this analysis. Models get better as a result of feature relevance driving predictions.

4.4 Feature selection:

In this stage, the key features are selected using "Principal Component Analysis (PCA)". It is known for its cpability to reduce the issue of data dimensionality. PCA attempts to determine the lower-dimensional surface to show the high-dimensional data. It reduces the computational complexity of the models which makes machine learning algorithms run faster [33].

The steps involved in PCA for feature selection process are as follows:

- Standardization of the data. (with mean =0 and variance = 1)
- Computing the Covariance matrix of dimensions.
- Obtain the Eigenvectors and Eigenvalues from the covariance matrix.
- Arrange the obtained eigenvalues in descending order and select the top Eigenvectors that correspond to the largest eigenvalues.
- Create the projection matrix using the Eigenvectors you've chosen.
- Modify the initial data set using the new matrix to obtain the new dimensional feature subspace.
- Select the appropriate features from the newly created feature subspace.

In general, the processing speed of the machine learning algorithms will be reduced due to high data dimensionality in the collected data and hence it is essential to overcome this problem. PCA decreases model difficulty, allowing machine learning algorithms to run quicker.

4.4 Attack Detection using Fuzzy based XGBoost Algorithm

The security attacks will be classified and detected using a machine learning approach. A Fuzzy based XGBoost algorithm is used for detecting security threats in the SDN architecture. The process flow involved in the proposed attack detection is illustrated in figure 4.2.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
Article Received: *25 December 2022* ***Revised:*** *12 January 2023* ***Accepted:*** *20 February 2023*

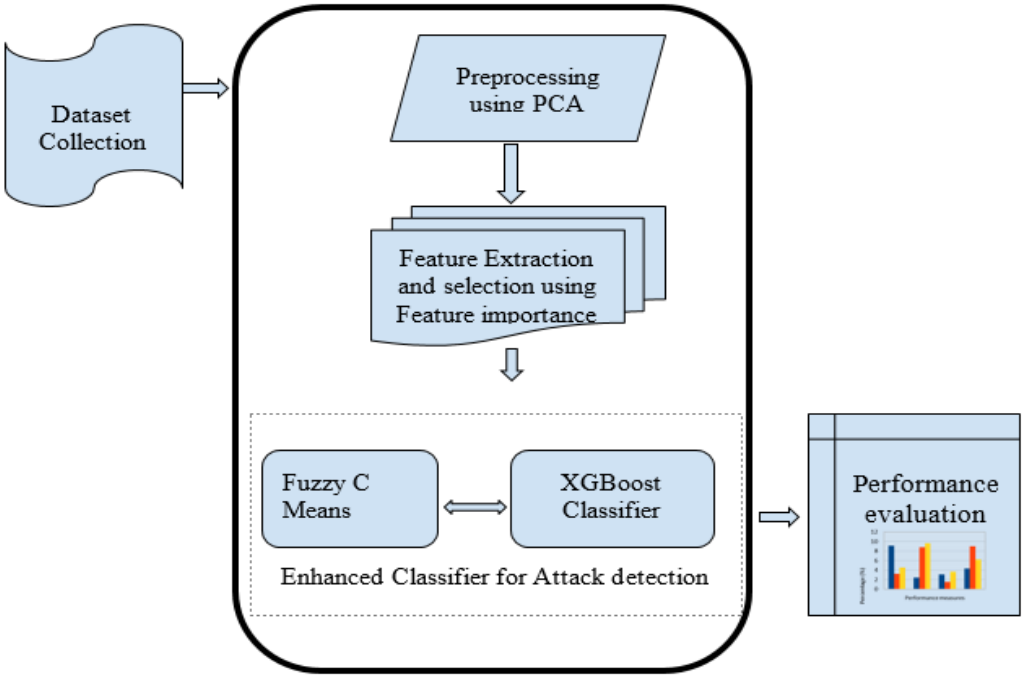

Figure 4.2 Attack detection process using Fuzzy based XGBoost classifier

4.4.1 Fuzzy based XGboost Classification Algorithm

Extreme Gradient Boosting (XGBoost) is a supervised and most popular ML algorithm introduced by Chen in 2014 [34]. XGBoost is a scalable algorithm which uses a tree-based model for performing classification tasks. The XGboost implements gradient boosted decision trees for achieving high speed and better performance. The proposed XGboost algorithm is an To assign and forecast a target label, a group of separate decision trees is used. The XGboost algorithm can handle large scale data and can make predictions with high accuracy.

The dataset must be split before training the ML model. Randomly splitting the dataset minimizes the consistency of training and testing the model. In this research, the dataset is split using a Fuzzy C-means clustering (FCM) algorithm. The membership degree for each data sample is determined and the cluster to which the data belongs to is calculated using the FCM membership degree. The degree of FCM membership and the cluster matching can be minimized as shown in equation 1 and 2 [35].

$$v_i = \frac{\sum_{k=1}^{n} u^m{}_{ik} x_k}{\sum_{k=1}^{n} u^m{}_{ik}}; 1 \le i \le c \quad \text{.....(1)}$$

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}{}^2 A}{d_{jk}{}^2 A} \right)^{\frac{2}{m-1}}} \quad \text{....(2)}$$

The main objective of dividing the dataset into clusters to achieve an optimal loss function and make this function reach the minimum value.

$$J(U,V) = \sum_{j=1}^{N} \sum_{i=1}^{c} (u_{ij})^2 (d_{ij})^2 \quad \text{....(3)}$$

Here, u_{ij} denotes the degree of membership of j.

For achieving better classification performance using the Fuzzy based XGBoost classifier, multiple individual decision trees (DT) are integrated. For each DT, the value of the loss function defined in equation 3 is reduced. Hence, the DTs that are added in the classifier are used for increasing the classification performance and accuracy. Weights of the data samples perform a significant influence in enhancing the performance of XGBoost algorithm. Here, the weights are assigned to all independent features or variables which are given as input to the DT for predicting the outcome. The values of the DT-predicted variables are raised, and these variables are subsequently put into the second decision tree. In this way, the individual classifiers/predictors are ensemble to obtain a robust and more precise model for classification.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
Article Received: *25 December 2022* ***Revised:*** *12 January 2023* ***Accepted:*** *20 February 2023*

The XGBoost algorithm aggregates the output of each individual DT using gradient boosted trees. After aggregating the output of each tree, XGboost will assign a unique label to each network attack based on which the attacks are classified as normal or malicious.

5. Results and Discussion

The Fuzzy based XGBoost classifier is designed to enhance the detection process. As discussed in the research methodology section, the proposed model can detect attacks over the heterogeneous 5G network with controlled performance in the presence of attacks.

5.1 Dataset Description

The UNSW-NB15 dataset was used for empirical and simulation analyses. The performance of the classifier for identifying security attacks is analyzed in terms of its ability to improve the execution of the classifier for identifying abnormal activities in a network. In this research, the attack detection model is constructed using both the training and testing data. Since this work uses a supervised ML model, it is essential to validate that the model performs well on the new data and provides accurate results. The presented model is simulated utilizing the training and testing data. The data split is performed by randomly selecting the data quantity. The data consists of both Malicious (51.9%) and Normal (48.1%) data as shown in Figure 5.1.

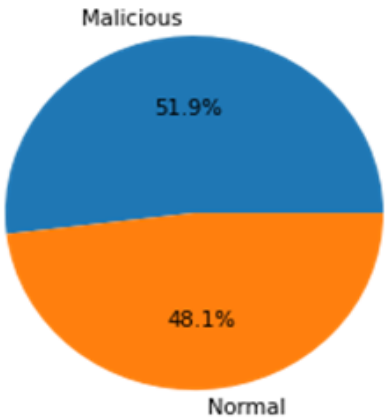

Figure 5.1 Percentage of data split as malicious and normal data

5.2 Performance Evaluation

The performance of the presented Fuzzy based XGBoost classifier is evaluated using different evaluation metrics such as "accuracy, precision, recall, F1 score, Specificity, and False alarm rate (FAR)". In this research, the accuracy is measured using four various classification parts namely: "True positives (TP), True negatives (TN), False positives (FP), False negatives (FN)". These terms are used for constructing a confusion matrix. The confusion matrix is primarily utilised to solve classification accuracy issues when the output might be multiple categories. Confusion matrix is a table with four different combinations of predicted and actual values as shown in below:

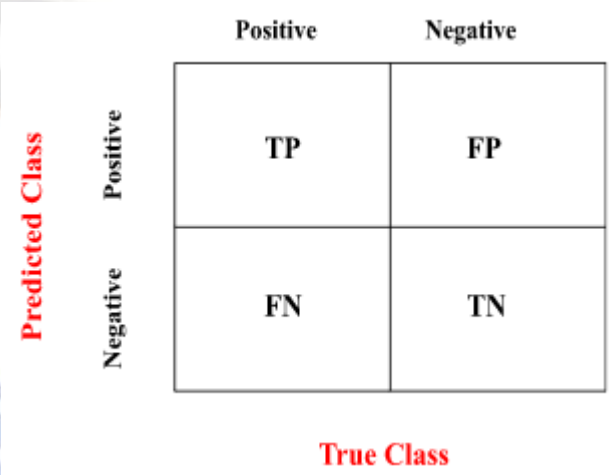

Figure 5.2 Confusion matrix

Accuracy is defined as the proportion of accurately recognised assaults, as shown in the equation below:

$$\textbf{\textit{Accuracy}} = \frac{TP+TN}{TP+TN+FP+FN} \ldots. (4)$$

Recall for a function is calculated as the ratio of correctly classified security attacks and is provided as:

$$\textbf{\textit{Recall}} = \frac{TP}{TP+FN} \ldots.. (5)$$

The F1 score, which can range between 1 and 0, is used to assess the system's accuracy. Where 1 denotes the best value and 0 denotes the poorest value. Similarly, the F1 score is defined as:

$$\textbf{\textit{F1 score}} = \frac{2 * Precision * Recall}{Precision + Recall} \ldots. (6)$$

Precision is defined as the number of accurate positive predictions. It is calculated as the fraction of accurately classified network attacks to all other attacks. It is defined as:

$$\textbf{\textit{Precision}} = \frac{TP}{TP+FP} \ldots. (7)$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
Article Received: *25 December 2022* ***Revised:*** *12 January 2023* ***Accepted:*** *20 February 2023*

The confusion matrix for the proposed attack detection is shown in table 1

Table 1. Confusion Matrix

Sample Class		**Predicted**	
		Normal	Attack
Real	Normal	TP	FP
	Attack	FN	TN

The findings from experiment are illustrated in the graphs below:

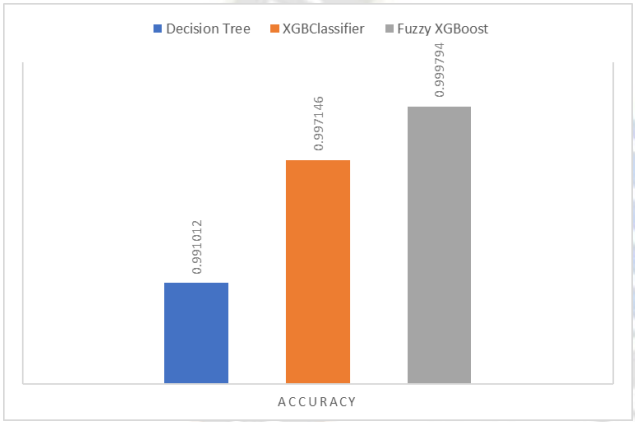

Figure 5.3 Accuracy of the XGBoost classifier

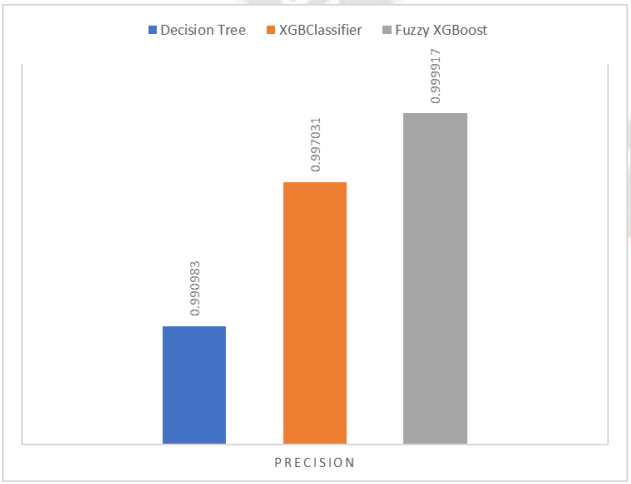

Figure 5.4 Precision of the XGBoost classifier

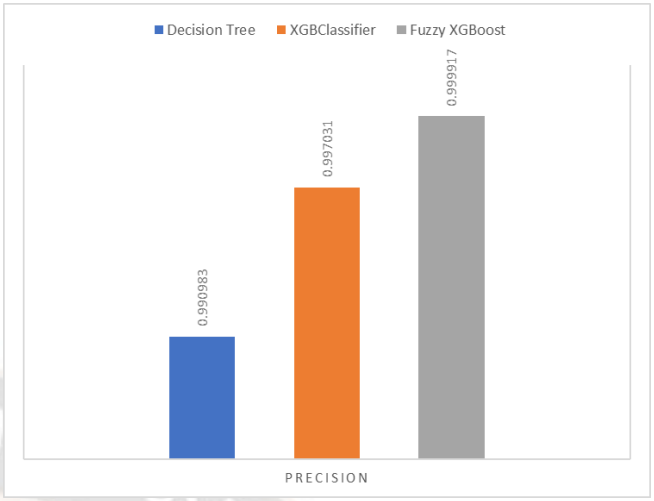

Figure 5.5 Recall score of the XGBoost classifier

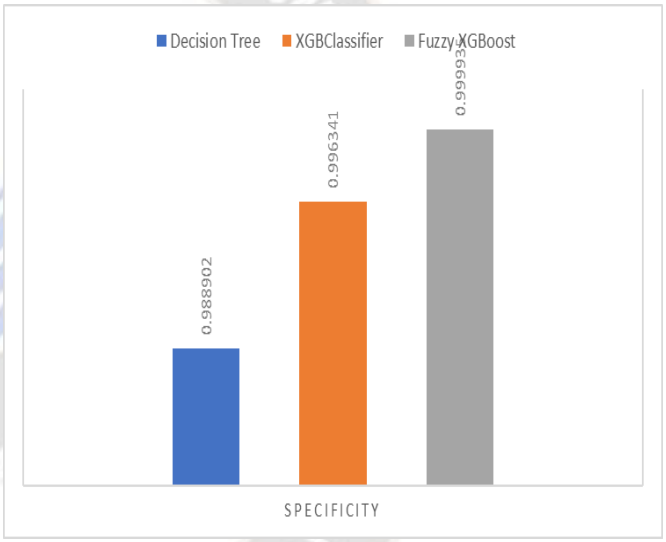

Figure 5.6 Specificity of the XGBoost classifier

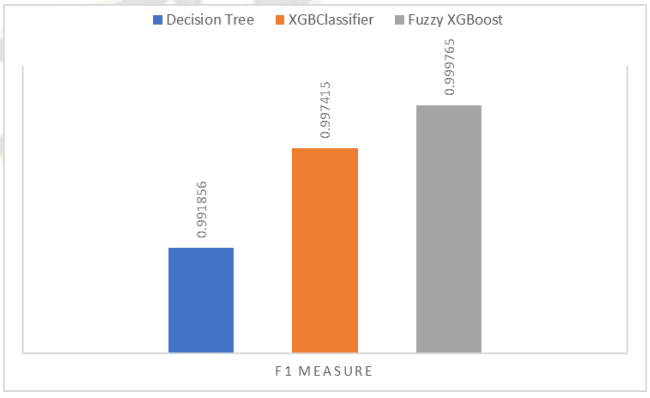

Figure 5.7 F1 score of the XGBoost classifier

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
Article Received: *25 December 2022* ***Revised:*** *12 January 2023* ***Accepted:*** *20 February 2023*

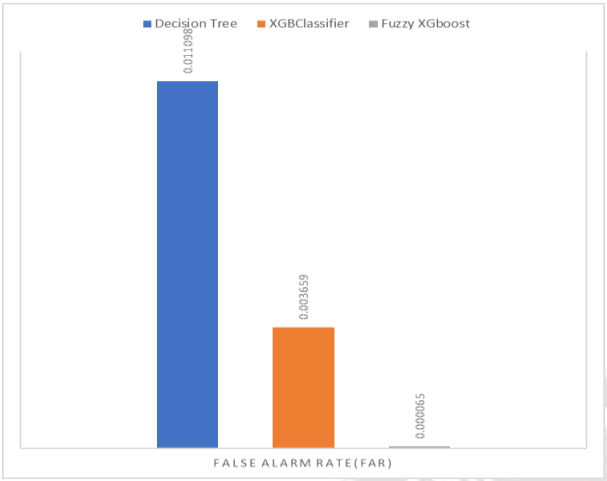

Figure 5.8 FAR of the XGBoost classifier

5.3 Comparative Analysis

The performance of the presented approach was validated by contrasting simulation findings of the proposed approach with other existing classification models. In this research, the performance of the Fuzzy based XGBoost classification model was contrasted with existing Decision tree, and the XG boost classification model. The decision tree model is selected for comparison since it is a flexible, easy to use ML algorithm. In addition, the DT is simple and diverse for performing classification and regression tasks. On the other hand, the XG boost classifier is also a type of decision tree based algorithm. The XG boost classifier incorporates the unstructured data in the form of text, images etc. The XG boost classifier is more appropriate compared to other neural networks since it is more suitable for small to medium structured or tabular data. The classification report and performance of the Fuzzy based XGBoost model and other classifiers are tabulated in table 2.

Table 2. Performance metrics for the classifiers for attack detection

Evaluation Metrics	Decision Tree	XGBoost	Proposed Fuzzy based XGBoost
Accuracy	99.10%	99.71%	99.97%
Precision	99.09%	99.70%	99.99%
Recall	99.27%	99.77%	99.96%
F1-score	99.18%	99.74%	99.97%
Specificity	98.89%	99.63%	99.99%
False Alarm Rate	0.0011098	0.003659	0.000065

It can be inferred from the simulation results and the performance evaluation metrics (from table 1) that the proposed Fuzzy based XGBoost classifier achieves a phenomenal accuracy with respect to different evaluation metrics. The accuracy of the Fuzzy based XGBoost classifier was found to be 99.97% which is superior compared to existing decision trees and XGBoost classifiers. Correspondingly, the False Alarm Rate of the proposed model is the least with a value of 0.000065 which is lesser than the DT and XG-Boost classifier. Results validate the effectiveness of the proposed attack detection approach.

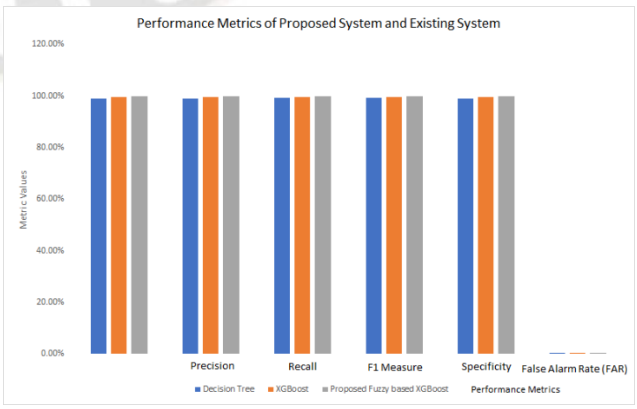

Figure 5.9 Comparison of Existing and Proposed system

The proposed fuzzy-based XGBoost system outperforms both the traditional Decision Tree and XGBoost models. Its outstanding predictive potential is demonstrated by recall, accuracy, precision, and F1-score, all of which are above 99.9%. Fascinatingly, there is a distinguished decline in inaccurate forecasts, with the false alarm rate falling to 0.000065. Conversely, XGBoost and Decision Tree models exhibit worse performance across all metrics, underscoring the benefits of including fuzzy logic in the XGBoost technique. When everything is said and done, the recommended method yields novel results, making it a viable option for applications requiring a high degree of accuracy and dependability as shown in figure 5.9.

6. Conclusion

In this study, a Fuzzy based XGBoost attack detection model is proposed for securing the SDN based 5G network from various kinds of safety attacks. The proposed approach was evaluated using the simulated data collected from the "UNSW-NB15 dataset". The dataset incorporates training samples for detecting different types of attacks in SDN architecture. Further the data was preprocessed and normalized using a labeled encoding process which allows the classifier to process numerical variable data. A feature selection process was employed where the relevant features from the dataset were extracted and were represented as low-dimensional feature vectors in order to simplify the classification process. For attack detection, this research employs a Fuzzy based XGBoost classification model which classified the data type as 'normal' or 'attack'. Experimental analysis was conducted to evaluate the performance of the proposed approach using different performance metrics namely, "accuracy, precision, and recall and F1 measure". The proposed approach was compared with other two machine learning algorithms such as Decision Tree and XGBoost classifier. Results show that the proposed Fuzzy based XGBoost achieves superior attack detection accuracy of 99.97% compared to other two ML algorithms. In future work, the study intends to test the performance of the presented attack detection model with other benchmark datasets and to validate the efficiency of the model using various neural networks.

References

1. Polat, H., Polat, O., & Cetin, A. (2020). Detecting DDoS attacks in software-defined networks through feature selection methods and machine learning models. *Sustainability*, *12*(3), 1035.
2. Rawat, D. B., & Reddy, S. R. (2016). Software defined networking architecture, security and energy efficiency: A survey. *IEEE Communications Surveys & Tutorials*, *19*(1), 325-346.
3. Kreutz, D., Ramos, F. M., Verissimo, P. E., Rothenberg, C. E., Azodolmolky, S., & Uhlig, S. (2014). Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, *103*(1), 14-76.
4. Benzekki, K., El Fergougui, A., & Elbelrhiti Elalaoui, A. (2016). Software-defined networking (SDN): a survey. *Security and communication networks*, *9*(18), 5803-5833.
5. Karakus, M., & Durresi, A. (2017). A survey: Control plane scalability issues and approaches in software-defined networking (SDN). *Computer Networks*, *112*, 279-293.
6. Xie, J., Yu, F. R., Huang, T., Xie, R., Liu, J., Wang, C., & Liu, Y. (2018). A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges. *IEEE Communications Surveys & Tutorials*, *21*(1), 393-430.
7. Zhao, Y., Li, Y., Zhang, X., Geng, G., Zhang, W., & Sun, Y. (2019). A survey of networking applications applying the software defined networking concept based on machine learning. *IEEE Access*, *7*, 95397-95417.
8. Sahay, R., Meng, W., & Jensen, C. D. (2019). The application of Software Defined Networking on securing computer networks: A survey. *Journal of Network and Computer Applications*, *131*, 89-108.
9. Muthanna, A., Ateya, A. A., Al Balushi, M., & Kirichek, R. (2018, May). D2D enabled communication system structure based on software defined networking for 5G network. In *2018 International Symposium on Consumer Technologies (ISCT)* (pp. 41-44). IEEE.
10. Prabakaran, D., Nizar, S. M., & Kumar, K. S. (2021). Software-defined network (SDN) architecture and security considerations for 5G communications. In *Design methodologies and tools for 5G network development and application* (pp. 28-43). IGI Global.
11. Yungaicela-Naula, N. M., Vargas-Rosales, C., Pérez-Díaz, J. A., & Zareei, M. (2022). Towards security automation in software defined networks. *Computer Communications*, *183*, 64-82.
12. Eliyan, L. F., & Di Pietro, R. (2021). DoS and DDoS attacks in Software Defined Networks: A survey of existing solutions and research challenges. *Future Generation Computer Systems*, *122*, 149-171.
13. Jafarian, T., Masdari, M., Ghaffari, A., & Majidzadeh, K. (2021). A survey and classification of the security anomaly detection mechanisms in software defined networks. *Cluster Computing*, *24*(2), 1235-1253.

14. Yao, J., Han, Z., Sohail, M., & Wang, L. (2019). A robust security architecture for SDN-based 5G networks. *Future Internet*, *11*(4), 85.
15. Soussi, W., Christopoulou, M., Xilouris, G., & Gür, G. (2021). Moving Target Defense as a Proactive Defense Element for beyond 5G. *IEEE Communications Standards Magazine*, *5*(3), 72-79.
16. Sengupta, S., Chowdhary, A., Sabur, A., Alshamrani, A., Huang, D., & Kambhampati, S. (2020). A survey of moving target defenses for network security. *IEEE Communications Surveys & Tutorials*, *22*(3), 1909-1941.
17. Sedjelmaci, H. (2021). Cooperative attacks detection based on artificial intelligence system for 5G networks. *Computers & Electrical Engineering*, *91*, 107045.
18. Aryal, B., Abbas, R., & Collings, I. B. (2021). SDN Enabled DDoS Attack Detection and Mitigation for. *Journal of Communications*, *16*(7).
19. Zhang, Y., Chen, M., & Lai, R. (2016). Cloudified and software defined 5G networks: Architecture, solutions, and emerging applications. *Mobile Networks and Applications*, *21*(5), 727-728.
20. Singh, S., & Prakash, S. (2019, March). A Survey on Software Defined Network based on Architecture, Issues and Challenges. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 568-573). IEEE.
21. Afaq, A., Haider, N., Baig, M. Z., Khan, K. S., Imran, M., & Razzak, I. (2021). Machine learning for 5G security: Architecture, recent advances, and challenges. *Ad Hoc Networks*, *123*, 102667.
22. Abdulqadder, I. H., Zhou, S., Zou, D., Aziz, I. T., & Akber, S. M. A. (2020). Multi-layered intrusion detection and prevention in the SDN/NFV enabled cloud of 5G networks using AI-based defense mechanisms. *Computer Networks*, *179*, 107364.
23. Lin, B. S. P. (2021). Toward an AI-enabled SDN-based 5G & IoT network. *Netw. Commun. Technol.*, *5*(2), 1-7.
24. Revathi, M., Ramalingam, V. V., & Amutha, B. (2021). A machine learning based detection and mitigation of the DDOS attack by using SDN controller framework. *Wireless Personal Communications*, 1-25.
25. Aslam, M., Ye, D., Tariq, A., Asad, M., Hanif, M., Ndzi, D., ... & Jilani, S. F. (2022). Adaptive Machine Learning Based Distributed Denial-of-Services Attacks Detection and Mitigation System for SDN-Enabled IoT. *Sensors*, *22*(7), 2697.
26. Alamri, H. A., Thayananthan, V., & Yazdani, J. (2021). Machine Learning for Securing SDN based 5G network. *Int. J. Comput. Appl*, *174*(14), 9-16.
27. Latif, Z., Umer, Q., Lee, C., Sharif, K., Li, F., & Biswas, S. (2022). A Machine Learning-Based Anomaly Prediction Service for Software-Defined Networks. *Sensors*, *22*(21), 8434.
28. Feghali, A., Kilany, R., & Chamoun, M. (2015, July). SDN security problems and solutions analysis. In *2015 International Conference on Protocol Engineering (ICPE) and International Conference on New Technologies of Distributed Systems (NTDS)* (pp. 1-5). IEEE.
29. Patil, V., Patil, C., & Awale, R. N. (2017, July). Security challenges in software defined network and their solutions. In *2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-5). IEEE.
30. Aziz, N. A., Mantoro, T., & Khairudin, M. A. (2018, September). Software defined networking (SDN) and its security issues. In *2018 International Conference on Computing, Engineering, and Design (ICCED)* (pp. 40-45). IEEE.
31. Raman, M. G., Somu, N., Kirthivasan, K., Liscano, R., & Sriram, V. S. (2017). An efficient intrusion detection system based on hypergraph-Genetic algorithm for parameter optimization and feature selection in support vector machine. *Knowledge-Based Systems*, *134*, 1-12.
32. Zhiqiang, L., Mohi-Ud-Din, G., Bing, L., Jianchao, L., Ye, Z., & Zhijun, L. (2019, August). Modeling network intrusion detection system using feed-forward neural network using unsw-nb15 dataset. In *2019 IEEE 7th International Conference on Smart Energy Grid Engineering (SEGE)* (pp. 299-303). IEEE.
33. Bhattacharya, S., Maddikunta, P. K. R., Kaluri, R., Singh, S., Gadekallu, T. R., Alazab, M., & Tariq, U. (2020). A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU. *Electronics*, *9*(2), 219.
34. Fan, J., Wang, X., Wu, L., Zhou, H., Zhang, F., Yu, X., ... & Xiang, Y. (2018). Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China. *Energy conversion and management*, *164*, 102-111.
35. Qin, J., Fu, W., Gao, H., & Zheng, W. X. (2016). Distributed $ k $-means algorithm and fuzzy $ c $-means algorithm for sensor networks based on multiagent consensus theory. *IEEE transactions on cybernetics*, *47*(3), 772-783.