Machine Learning based Security Solutions against DDoS Attacks in Software Defined Networks

*Sanjay Vidhani[1] and Amarsinh Vidhate[2]

[1]Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India
[1]Faculty of K. J. Somaiya College of Engineering, Somaiya Vidyavihar University, Mumbai, India
[2]Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai, India
[1]sanjayvidhani@somaiya.edu [2]amar.vidhate@rait.ac.in

Abstract. Software-Defined Networking (SDN) is a revolutionary network paradigm that enhances the flexibility, scalability, and management of networks. However, this centralized approach to network control also creates potential vulnerabilities that attackers can exploit. Traditional security measures are often ineffective in detecting novel and sophisticated attacks. This paper examines the application of various Machine Learning (ML) techniques, including Support Vector Machines (SVM), Random Forest (RF), Naive Bayes (NB), K-Nearest Neighbors (KNN), Decision Trees (DT), K-Means, DBSCAN, and Agglomerative Clustering, for detecting and mitigating security vulnerabilities in SDN environments. The effectiveness of these models is assessed using network traffic datasets, demonstrating their capability in real-time SDN defense. In future work, Deep Learning (DL) techniques will be explored due to their significant potential to improve the detection and mitigation of Distributed Denial of Service (DDoS) attacks in SDN systems.

Keywords: Software-Defined Networking, Machine Learning, DDoS attacks, Support Vector Machine, Random Forest, Naïve Bayes.

1 Introduction

Software-Defined Networking (SDN), which allows for centralized management of the entire network via a programmable interface, divides the control plane from the data plane [1,9]. While SDN brings numerous advantages, such as scalability, flexibility, and efficient resource management, it also introduces significant security challenges [1]. Since the SDN controller turns into a single point of control, a breach could have disastrous effects on the network as a whole [3]. Moreover, SDN's flexible and programmable nature necessitates advanced techniques for detecting and mitigating known and emerging attacks. Traditional security mechanisms like Intrusion Detection Systems (IDS) [10] and firewalls have limitations in SDN environments. Signature-based IDS solutions are particularly ineffective against zero-day attacks or previously unseen attack vectors [1,9]. Consequently, there is increasing interest in

N. Vyas et al. (eds.), *Proceedings of the International Conference on Advancements in Computing Technologies and Artificial Intelligence (COMPUTATIA 2025)*, Advances in Intelligent Systems Research 189,
https://doi.org/10.2991/978-94-6463-700-7_36

leveraging Machine Learning (ML) models for enhancing SDN security. These models can continuously adapt to emerging threats by analyzing real-time traffic patterns and behaviors. SDN networks face a diverse range of security challenges. Among the most common assaults in SDN are Distributed Denial-of-Service (DDoS) and Denial-of-Service (DoS) attacks, which overload SDN controllers or network resources, disrupting and stopping them from functioning [1,3]. Man-in-the-Middle (MitM) attacks compromise data security and integrity by intercepting and changing communications between SDN components [4]. Flow Table Overload exploits SDN controllers to flood flow tables in switches, blocking legitimate traffic [9].

The aim is to create a system that can precisely identify threats and differentiate them from regular traffic in an SDN environment [13]. Machine learning algorithms [11,13] enable traffic classification by analyzing behavior patterns, effectively identifying both known and emerging attacks. In SDN networks, the OpenFlow protocol offers a variety of mitigation strategies [15], including deep packet inspection, network reconfiguration, redirection, block ports, drop packets, and topology changes; each approach has pros and cons of its own. Block ports and drop packets are the quickest and easiest mitigating strategies. The legitimate traffic may be dropped if they totally block the attack sources in the case of a false alarm or if the legitimate site is compromised. Therefore, in this work, a timeout is introduced to the blocking rule to guarantee that a legal user will not be indefinitely blocked [15]. On the other hand, deep learning algorithms, especially those based on neural networks, can capture more complex, high-dimensional patterns in network traffic data, making them ideal for identifying sophisticated and evolving threats in SDN networks [5,16] . The rest of the paper is organized as follows: Section II presents a comprehensive literature review on machine learning techniques for DDoS detection and mitigation. Section III will explain all the supervised and unsupervised techniques along with advantages, drawbacks and examples. Section IV will describe the methodology of datasets which contain labeled data for both normal traffic and attacks such as DDoS. Section V implements the supervised learning techniques. Section VI implements the unsupervised learning techniques. Section VII presents the findings and analysis of machine learning methods for assessing the security and performance of 5G networks. Section VIII presents the conclusion and future directions.

2 Literature Survey

Identifying DDoS attacks, machine learning (ML) and deep learning (DL) models have proven to be more successful than conventional statistical or policy-based techniques [1]. According to [2], recent studies have investigated the use of ML and DL techniques for detecting DDoS attacks in SDN networks. The proposal in [3] included a distinct scalable memory based support vector machine solution for DDoS threats as well as mitigation architecture for attack detection in SDN. SDN networks are under risk due to the growing use of networking devices. The centralized control and dumb forwarding features of switches make SDN a clear target for DDoS attackers. Accord-

ing to [5], DDoS attacks are an issue that has led to more research on techniques to detect and neutralize them. We found publications that focused on SDN's DDoS attack detection and defense strategies after reviewing earlier research [4,6,7] . As per [8], the network performance of SCTP protocols were examined and analyzed. According to [9], SDN can assist reduce DDoS attacks and offers a thorough assessment of DDoS defense strategies that use SDN in cloud computing. Network security is essential for the seamless formation of SDN-based clouds without having to worry about DDoS attacks. The properties of SDN and its ability to detect DDoS assaults in comparison to traditional networks were provided in [10]. For SDN, researchers have investigated a number of DDoS detection methods such as machine learning, entropy-based analysis, connection rate monitoring, traffic pattern analysis and integrating SNORT with OpenFlow [11]. The application of machine learning to SDN security has been examined in [11]. KNN and logical regression (LR) techniques were proposed in [12] as ways to identify DDoS attacks. The use of the Random Forest classifier in conjunction with the K-Means clustering technique for attack detection in SDN is covered in [13]. According to [14], DDoS detection with new SDN features were covered. Random Forest (RF) and K-Nearest Neighbors (KNN) learning models are used to explores the performance of SDN based DDoS attack discovery and mitigation in [15]. Reference [17] discusses entropy based strategies for DDoS recovery and mitigation in SDN. This paper will provide a comprehensive review of machine learning based security solutions for mitigating DDoS attacks in SDN.

3 Machine Learning Techniques

Following are the supervised and unsupervised machine learning algorithms [1,2,3,16] for DDoS detection [5,6] with their advantages, limitations along with example.

3.1 Support Vector Machine (SVM)

Support Vector Machines (SVM) differentiate between benign and malicious traffic by identifying the best possible hyperplane within a high-dimensional space [4,11]. They excel in binary classification and high-dimensional data but are computationally expensive and not scalable for large datasets. An example use is classifying SDN traffic based on packet size, inter-arrival time, and flow duration [11].

3.2 Random Forest (RF)

Random Forest (RF) [13] detects attack patterns using an ensemble of decision trees, offering robustness to overfitting and effective handling of large datasets. It provides feature importance for interpretability but has high computational overhead during training. An example use is real-time identification and blocking of suspicious flows by analyzing traffic characteristics.

3.3 Naïve Bayes (NB)

Naive Bayes (NB) [11,12] predicts the possibility of traffic being malicious using Bayes' theorem. It is simple, fast, and effective for large datasets and categorical data but assumes feature independence, which may not always apply. An example use is classifying packets as normal or malicious based on historical probabilities [11,12].

3.4 K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) [12,15] classifies traffic by comparing its similarity to nearest neighbors in feature space. It is easy to implement and requires no explicit training but is computationally intensive during prediction for large datasets. An example use is identifying DDoS traffic by comparing new flows to known malicious patterns.

3.5 Decision Trees (DT)

Decision Trees (DTs) [3,13,15] classify traffic by creating a tree structure based on flow features. They are easy to use, implement and manage for both categorical and continuous data but can overfit if not properly pruned. An example use is detecting anomalies in SDN traffic through complex decision rules.

3.6 K-Means Clustering

K-Means [13] groups traffic flows into clusters and flags outliers as potential DDoS attacks. It is simple, fast, and effective for detecting volumetric DDoS attacks but requires prior knowledge of the number of clusters. An example use is identifying potential attacks by clustering traffic based on flow characteristics.

3.7 DBSCAN (Density-Based Spatial Clustering of Applications with Noise)

DBSCAN [13,14] groups dense regions of data points while considering sparse areas as anomalies. It does not require a predefined number of groups and efficiently detects sound and outliers. However, its performance is highly dependent on parameter selection, such as epsilon and the minimum number of points. An example use is detecting sparse malicious flows in SDN traffic.

3.8 Agglomerative Clustering

Agglomerative clustering [13,14,17] creates a hierarchy of clusters using a bottom-up approach. It provides a dendrogram for visualization and works well with small datasets but is computationally expensive for large ones. An example use is identifying hierarchical relationships in traffic data to separate normal traffic from DDoS flows.

4 Methodology

4.1 Data Collection

The initial step in developing an intelligent defense system [7,9] involves gathering network traffic data. While we reference datasets such as CICIDS 2018 and NSL-KDD [1], we have also created our own dataset, SDN-Per [1], which includes labeled data for both normal traffic and various attacks, including DoS, DDoS, and botnets. Traffic features such as Latency during attack, Latency Post Mitigation, Throughput Downlink, Throughput Uplink, PDR During Attack, PDR Post Mitigation, Signal Strength, Jitter, Bandwidth Utilization, Error Rate, and Energy Efficiency are extracted for training the models [1,7,9] . Sample datasets are displayed in Table 1.

Table 1. Rows of Dataset

Latency_During_Attack(ms)	Latency_Post_Mitigation(ms)	Throughput_Downlink(Mbps)	Throughput_Uplink(Mbps)	PDR_During_Attack(%)	PDR_Post_Mitigation(%)	Signal_Strength(dBm)	Jitter(ms)	Bandwidth_Utilization(%)	Error_Rate(%)	Energy_Efficiency(bits/J)	Connected_Devices
172.7	28.80	756.4	402.8	83.33	99.92	-90.8	4.2	26.9	5.3	52.25	351
250.0	9.332	489.3	190.0	88.20	95.61	-47.6	69.1	30.2	0.7	76.49	162
136.7	29.36	492.53	121.9	80.41	98.19	-38.7	36.3	39.2	3.2	81.32	164
483.8	16.23	833.99	59.26	86.43	99.09	-51.2	90.3	52.5	9.4	32.89	316
368.6	34.26	990.72	371.6	76.91	95.12	-61.5	19.9	20.7	8.2	16.57	78

4.2 Feature Engineering

Feature extraction is critical for converting raw network traffic data into a usable format for machine learning algorithms [1,3,11] . The following features are considered:
Packet Size: The total amount of bytes transmitted within a network flow.
Flow Duration: The total time taken for data transmission in a specific flow.
Source IP & Destination IP: The network addresses of the sender and receiver involved in a data flow.
Protocol: Type of protocol used in the communication.
Flow statistics: Derived metrics such as average packet inter-arrival times and flow rate.

5 Implementation of Supervised Learning Models

5.1 Support Vector Machine (SVM)

SVM [3,4] is used to create a hyperplane that separates normal traffic from malicious traffic. Labelled traffic data is used to train the model, with kernel functions employed to handle non-linear separations. Its grouping report is shown in Table 2.

Table 2. SVM Grouping Report

Grouping Report				
	precision	recall	f1-score	support
Moderate Class	0.95	1.00	0.97	190
Poor Class	0.00	0.00	0.00	10
Overall Accuracy	0.95		200 samples	
Overall Mean	0.48	0.50	0.49	
Adjusted Mean	0.90	0.95	0.92	
Regression Results				
Mean Squared Error:	78582.75			
R-squared Value:	-0.0105			

5.2 Random Forest (RF)

Random Forest [13,15] builds multiple decision trees and aggregates their results to classify traffic. This ensemble method is robust and can handle high dimensional data effectively, making it suitable for SDN traffic classification. Its grouping report is shown in Table 3.

Table 3. RF Grouping Report

Grouping Report				
	precision	recall	f1-score	support
Moderate Class	1.00	1.00	1.00	190
Poor Class	1.00	1.00	1.00	10
Overall Accuracy	1.00		200 samples	
Overall Mean	1.00	1.00	1.00	
Adjusted Mean	1.00	1.00	1.00	
Regression Results				
Mean Squared Error:	85476.92			
R-squared Value:	-0.097			

5.3 Naïve Bayes (NB)

Naive Bayes [12,14] applies Bayes' theorem to predict the likelihood of a class using observable features. Despite the "naive" independence assumptions, it performs well in classifying SDN traffic. Its grouping report is shown in Table 4.

Table 4. NB Grouping Report

Grouping Report				
	precision	recall	f1-score	support
Moderate	0.99	1.00	1.00	190
Poor	1.00	0.90	0.97	10
Overall Accuracy	0.99		200 Samples	
Overall Mean	1.00	0.95	0.98	
Adjusted Mean	1.00	0.99	0.99	

5.4 K-Nearest Neighbours (KNN)

KNN [12,15] uses proximity to classify traffic. By identifying the K nearest neighbors of a flow and using their labels, the model classifies the traffic as benign or malicious. Its grouping report is shown in Table 5.

Table 5. KNN Grouping Report

Grouping Report				
	precision	recall	f1-score	support
Moderate	0.93	1.00	0.96	190
Poor	0.00	0.00	0.00	10
Overall Accuracy	0.93		200 samples	
Overall Mean	0.47	0.50	0.48	
Adjusted mean	0.86	0.93	0.89	

5.5 Decision Trees (DT)

The feature space is recursively divided using splits that maximize class purity to create decision trees [13]. This model is perfect for classifying SDN traffic because it is clear and easy to understand. Its grouping report is shown in Table 6.

Table 6. DT Grouping Report

Grouping Report				
	precision	recall	f1-score	support
Moderate	1.00	1.00	1.00	190
Poor	1.00	1.00	1.00	10
Overall Accuracy	1.00		200 samples	
Overall Mean	1.00	1.00	1.00	
Adjusted Mean	1.00	1.00	1.00	

6 Implementation of Unsupervised Learning Models

6.1 K-Means Clustering

K-Means [13] partitions the traffic data into K clusters, grouping similar traffic together. Outliers or traffic that does not fit into any cluster are flagged as potential threats.

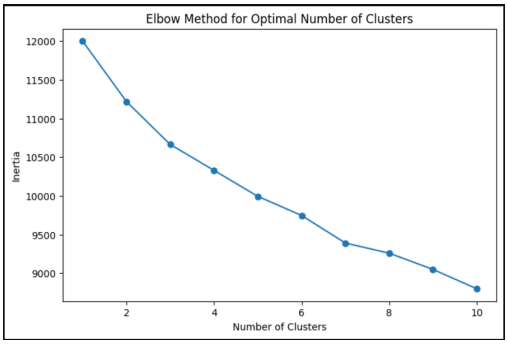

Fig. 1. K-Means Graph

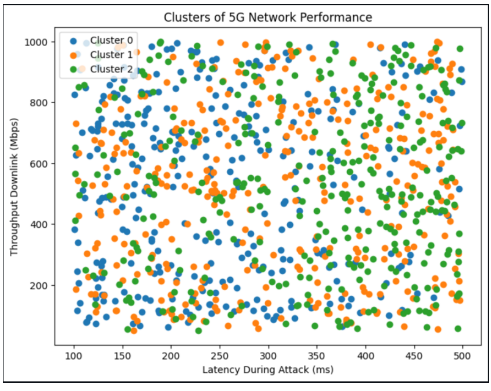

Fig. 2. K-Means Clustering

6.2 DBSCAN

DBSCAN [13,14] detects clusters based on density, making it highly effective for detecting DDoS attacks, which often generate high-density traffic in specific network segments.

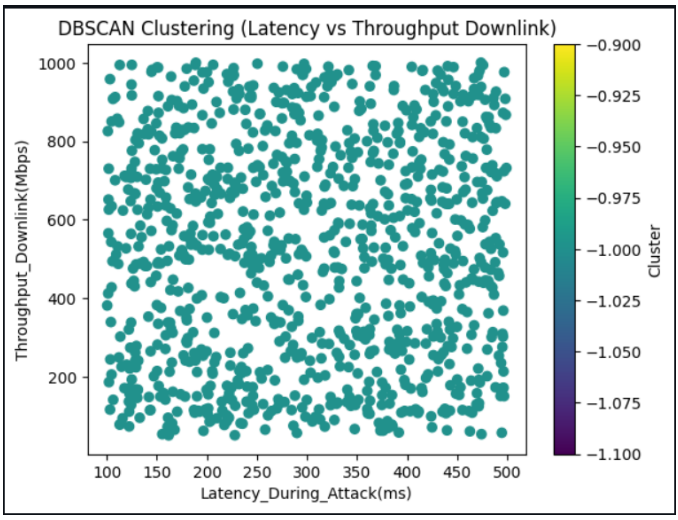

Fig. 3. DBSCAN Clustering

6.3 Hierarchical Clustering Dendogram

Agglomerative clustering [13] creates a hierarchy of clusters and combines them based on similarity. It can be useful for detecting multi-phase attacks that evolve over time.

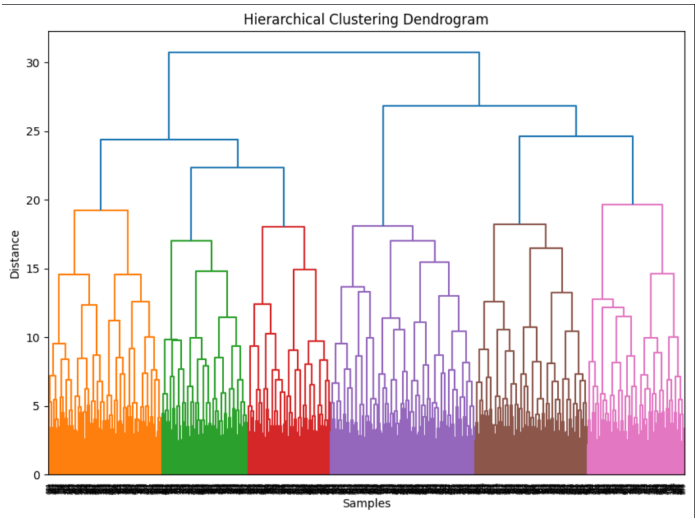

Fig. 4. DT Classification Report

7 Results and Analysis

Machine Learning Techniques for 5G Network Security and Performance Analysis: This paper evaluates several machine learning techniques applied to a dataset involving 5G network performance and security. Each method is analyzed [3,8,11] for its classification, regression, clustering performance, and its potential use for detecting attacks and ensuring security. Table 7 presents the calculated metrics for each supervised machine learning algorithm, while Fig. 5 visually represents these metrics in a graphical format. An optimal algorithm should have low false alarm rates, high recall, accuracy, precision, and F1-measure, as well as detection time to measure efficiency [13,14].

Table 7. Algorithms Calculated Values

Sr. No.	Algorithms	Accuracy	Precision	Recall	F1-score
1	SVM	95	95	100	97
2	RF	100	100	100	100
3	NB	99	99	100	100
4	KNN	93	93	100	96
5	DT	100	100	100	100

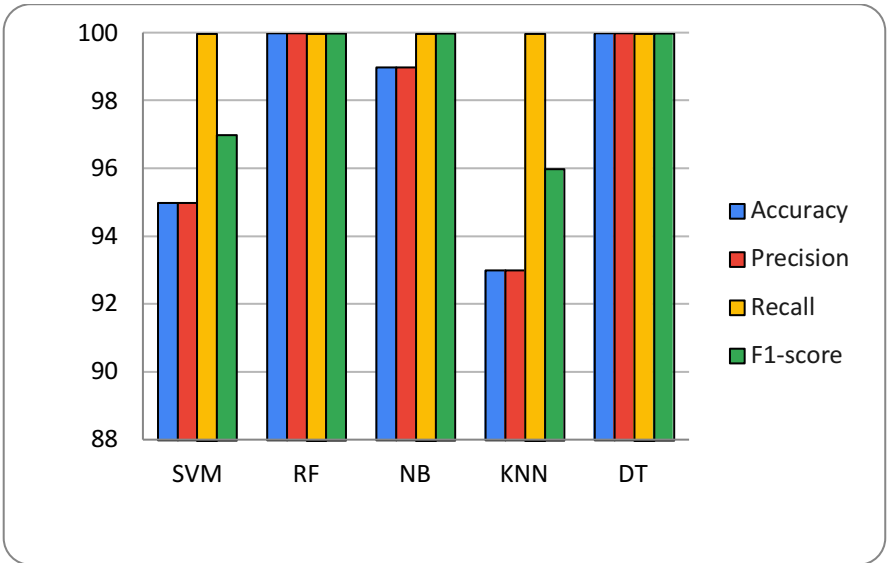

Fig. 5. Comparison of Supervised ML methods

5G performances parameters like latency, throughput, efficiency and packet delivery ratio (PDR) [8] are shown in Fig. 6. and implemented using Python programming. Results are recorded during the attack and after post mitigation attack. The values of latency, throughput, efficiency and packet delivery ratio improved once the DDoS attack mitigated [13,14,15] .

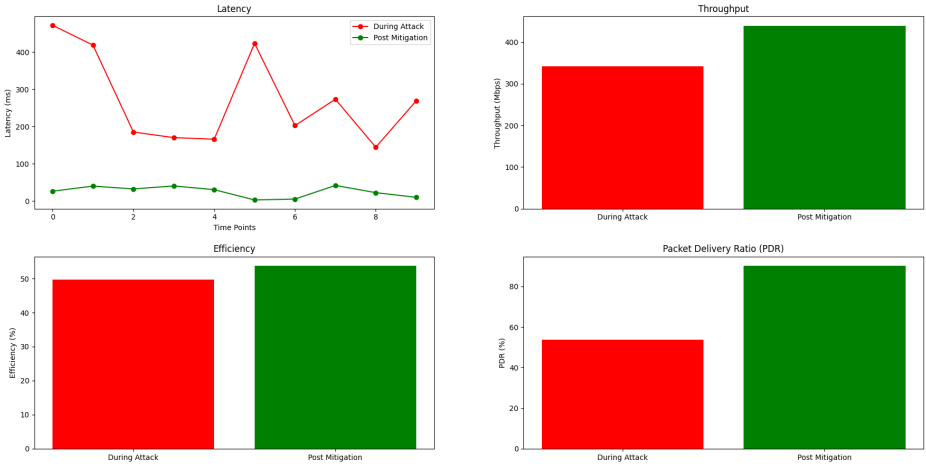

Fig. 6. Latency, Throughput, Efficiency, and Packet delivery ratio Graphs during the attack and post-mitigation of Supervised ML methods

8 Conclusion and Future Work

This paper offers a detailed evaluation of machine learning models for defending SDN networks against various attack scenarios. The analysis demonstrates the effectiveness of supervised models like Random Forest and decision trees in detecting attacks with high accuracy and recall. Unsupervised models such as DBSCAN show exceptional performance in identifying novel or unexpected attack patterns. Comparisons of all models are made on the basis of accuracy, regression suitability, and Security/Attack Detection suitability, and it is found that DBSCAN models have high Security/Attack detection suitability. 5G network performance key metrics such as latency, throughput, efficiency, and packet delivery ratio (PDR) were analyzed during the attack and post-mitigation attack, and results are recorded and displayed. Overall, the results highlight the strength of advanced models in detecting and mitigating network threats while maintaining critical performance metrics. Machine Learning (ML) and Deep Learning (DL) solutions offer efficient and adaptive approaches for enhancing SDN security and management.

Deep Learning (DL) algorithms have significant potential to enhance the detection and mitigation of Distributed Denial of Service (DDoS) attacks in SDN environments. The future work in this area focuses on addressing current limitations, enhancing scalability, improving detection accuracy and creating robust, real-time solutions. By focusing on these areas, DL can transform the detection and mitigation of DDoS attacks in SDN environments, enhancing network security and resilience.

Acknowledgements. We sincerely thank Ramrao Adik Institute of Technology, Nerul, Navi Mumbai and K. J. Somaiya College of Engineering, Somaiya Vidyavihar University for their support and resources in conducting this research. We also appreciate the valuable feedback from reviewers and the conference organizers for providing this platform to share our work.

Disclosure of Interests. We declare that there are no conflicts of interest related to this research work. All findings and conclusions presented are based solely on independent research and analysis.

References

1. Aslam, N., Shrivastava, S., Gore, M.M.: A Comprehensive Analysis of Machine Learning and Deep Learning Based Solutions for DDoS Attack Detection in SDN. Arabian Journal for Science and Engineering 49, 3533–3573 (2024). https://doi.org/10.1007/s13369-023-08075-2
2. Ali, T.E., Chong, Y.-W., Manickam, S.: Machine Learning Techniques to Detect a DDoS Attack in SDN: A Systematic Review. Applied Science MDPI (2023). https://doi.org/10.3390/app13053183
3. Revathi, M., Ramalingam, V.V., Amutha, B.: A Machine Learning-Based Detection and Mitigation of the DDoS Attack Using SDN Controller Framework. Wireless Personal Communications 127, 2417–2441 (2022). https://doi.org/10.1007/s11277-021-09071-1
4. Yungaicela-Naula, N.M., Vargas-Rosales, C., Perez-Diaz, J.A.: SDN-Based Architecture for Transport and Application Layer DDoS Attack Detection Using Machine and Deep Learning. IEEE Access (2021). https://doi.org/10.1109/ACCESS.2021.3101650
5. Alashhab, A.A., Zahid, M.S.: Enhancing DDoS Attack Detection and Mitigation in SDN Using an Ensemble Online Machine Learning Model. IEEE Access (2024). https://doi.org/10.1109/ACCESS.2024.3384398
6. Gupta, S., Grover, D.: A Comprehensive Review on Detection of DDoS Attacks using ML in SDN Environment. International Conference on Artificial Intelligence and Smart Systems (ICAIS-2021) (2021). ISBN: 978-1-7281-9537-7
7. Luong, T.-K., Tran, T.-D., Le, G.-T.: DDoS Attack Detection and Defense in SDN Based on Machine Learning. 7th NAFOSTED Conference on Information and Computer Science (NICS) (2020).
8. Baharudin, S.A., Taufik, N.I.M., Munisamy, S.B.: Network Performance Analysis of SCTP Protocol in Multihoming Transmission Through 5G Network. International Visualization, Informatics and Technology Conference (IVIT) (2024). https://doi.org/10.1109/IVIT62102.2024.10692896

9. Swami, R., Dave, M., Ranga, V.: Software-defined Networking-Based DDoS Defense Mechanisms. ACM Computing Surveys (2019). https://doi.org/10.1145/3301614
10. Bindra, N., Sood, M.: Detecting DDoS Attacks Using Machine Learning Techniques and Contemporary Intrusion Detection Dataset. Automatic Control and Computer Sciences 53(5), 419–428 (2019). https://doi.org/10.3103/S0146411619050043
11. Aslam, M., Ye, D., Hanif, M., Asad, M.: Machine Learning-Based SDN-enabled Distributed Denial of Service Attacks Detection and Mitigation System. Conference Paper (Oct 2020).
12. Priya, G.G., Shriram, S.H., Jeeva, S., Sakthi Priya, G., Balasubadra, K.: Detection of Distributed Denial of Service (DDoS) Attack Using Logistic Regression and K-Nearest Neighbor Algorithms. International Journal of Intelligent Systems and Applications in Engineering (IJISAE). ISSN: 2147-6799
13. Firdaus, D., Munadi, R., Purwanto, Y.: DDoS Attack Detection in Software-Defined Network Using Ensemble K-means++ and Random Forest. Conference Paper (Dec 2020).
14. Gadallah, W.G., Omar, N.M., Ibrahim, H.M.: Machine Learning-Based Distributed Denial of Service Attacks Detection Technique Using New Features in Software-Defined Networks. International Journal of Computer Network and Information Security (IJCNIS) 3, 15–27 (2021). https://doi.org/10.5815/ijcnis.2021.03.02
15. Mohsin, M.A., Hamad, A.H.: Performance Evaluation of SDN DDoS Attack Detection and Mitigation Based on Random Forest and K-Nearest Neighbors Machine Learning Algorithms. International Information and Engineering Technology Association (IIETA) 36(2) (Apr 2022). https://doi.org/10.18280/ria.360207
16. Sahbi, A., Jaidi, F.: Artificial Intelligence for SDN Security: Analysis, Challenges, and Approach Proposal. 15th International Conference on Security of Information and Networks (SIN) (Apr 2022). https://doi.org/10.1109/SIN56466.2022.9970501
17. Vidhani, S.M., Vidhate, A.: Entropy-Based DDoS Detection and Mitigation in Software-Defined Networks. Journal of Systems Engineering and Electronics 34(8), 46–50 (Aug 2024). https://doi.org/20.14118.jsee.2024.V34I8.1854