National Research Journal of Information Technology & Information Science
Volume No: 13, (January) Year: 2026 (Special Issue)
PP: 840-850
ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

REAL-TIME ANOMALY DETECTION IN SDN WITH MACHINE LEARNING

Vaishnavi Rathore

Department of CSE, Chandigarh University

Rohit Raj

Department of CSE, Chandigarh University

Er. Vishal Sharma

Department of CSE, Chandigarh University

Ishika Tinna

Department of CSE, Chandigarh University

Rishi Mishra

Department of CSE, Chandigarh University

Rishabh Raj

Department of CSE Chandigarh University

ABSTRACT

Software-Defined Networking (SDN) has transformed network management with programmability and centralized control. However, the logically centralized nature of SDN adds additional security risks, and real-time anomaly detection is essential to ensure network integrity. Traditional anomaly detection methods typically experience scalability and adaptability issues in the dynamic nature of SDN networks. This study explores the use of machine learning (ML) algorithms for real-time anomaly detection in SDN to take advantage of their ability to process large amounts of network traffic data and detect abnormal behavior from normal behavior. We evaluate a variety of machine learning models, both supervised and unsupervised, and introduce a real-time detection framework that combines feature engineering, model training, and anomaly classification. The result of our experiments proves the effectiveness of machine learning-based detection in identifying security threats, such as Distributed Denial-of-Service (DDoS) attacks and unauthorized access attempts. The proposed approach improves the security of SDN by enabling proactive threat mitigation, minimizing false positives, and adding minimal interference to network performance.

Keywords: Software-Defined Networking (SDN), Anomaly Detection, Machine Learning, Network Security, Real-Time Monitoring, Cybersecurity, Supervised Learning, Unsupervised Learning.

1. INTRODUCTION

Currently network architecture development has created Software-Defined Networks (SDNs) that separate control functions from data functions to create a fresh network style. The network control data split enables central programmatic management that offers higher responsiveness and flexibility toward changing requirements. The core strengths of SDNs stem from their control system and dynamic capabilities yet their basic programmability design creates new security and operational reliability problems with multiple types of vulnerabilities and system failures.

Real-time anomaly detection serves as a critical safety measure to defend SDN environments from security attacks and performance degradation as well as network congestion incidents. Real-time anomaly detection remains crucial because network traffic anomalies show signs of DDoS attacks and different issues as well as hardware failures which lead to potential attacks. An application of Machine Learning (ML) algorithms becomes essential due to this restriction. After consistency checking, it has been shown that real-time anomaly detection in SDNs can be well served by Machine Learning that is capable of learning patterns from a large amount of network traffic data and discovery of complex patterns suggesting anomalous behaviour. The use of ML makes the detection automated (taking away the human element) and enables proactive response to approaching threats. Large-scale dataset can be handled by these methods thus an efficient and scalable method to find abnormal network activity in real time.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

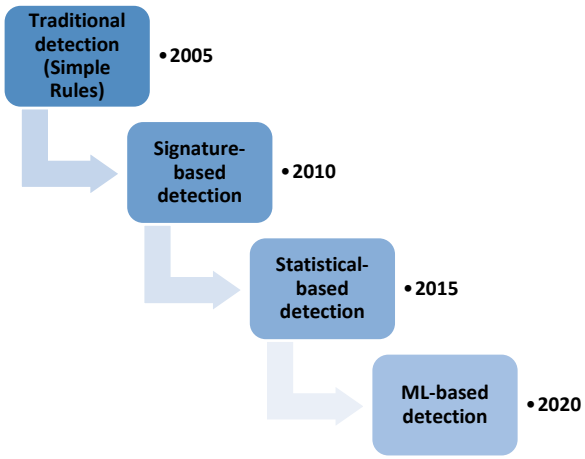

Figure 1. Evolution of Anomaly Detection Techniques in SDN

In this paper we examine how some of the Machine Learning techniques are leveraged in SDN anomaly detection, including supervised, unsupervised as well as reinforcement learning. The paper discusses in detail the comparative points of the methods, discussing the have and have nots, and when is each method best. The paper also discusses the shortcomings of the application of these ML based methods into the practical SDN networks such as data imbalance, scalability, and the requirement of periodic model updates. Lastly, the paper specifies possible future research directions to enhance robustness, accuracy, and efficiency of the ML based anomaly detection system in the SDN framework.

2. RELATED WORK

2.1 Supervised Learning Approaches

Supervised learning methods are basically used in anomaly detection in SDNs i.e. Software Defined Networks. These learning methods can classify network traffic into normal and anomalous categories in each labelled dataset. These methods are effective when large amount of labelled data is present allowing models to identify patterns and anomalies with high accuracy rate. Various research studies have analysed supervised learning techniques for detecting various type of network attacks.

- Wang et al. (2018) [5] used SVM to detect volumetric DDoS attacks in SDN environments. The study analysed that SVM can accurately differentiate between normal traffic and attacks. It has high detection rate of 96%. Though the model worked accurately but possessed challenges on large datasets, as such kind of intensive data can be time consuming.
- Wang et al. (2019) used Random Forest model to detect network anomalies in Software Defined Network. He trained the model on labelled data set to identify botnet attacks and different anomalies with high accuracy rate. He combined the flow level features with packet statistics to achieve high performance model.
- Kumar et al. (2020) used simpler model like logistic regression rather than high complexity models like SVM for detecting DDoS attacks within the network. the model was efficient in computations and reliable for real time applications but it was unable to detect newer and more powerful network threats.

2.2 Unsupervised Learning Approaches

Unsupervised learning approaches are great when there is no or few labelled data. The focus of these techniques is finding anomalies through departures from normal network behaviour as opposed to predetermined categories for training. Unsupervised approaches have become popular in SDNs because they can identify new and undiscovered anomalies.

- Automated network anomaly detection in SDNs can be achieved through the unsupervised deep learning model Autoencoders according to Zhao et al. (2019) [6]. Autoencoders train to understand input data which lets them detect anomalies based on high reconstruction errors. The detection strategy achieved results of 92% by successfully identifying DDoS and Port Scanning attacks. The authors acknowledged that autoencoders face difficulties in recognizing unknown attack patterns known as zero-day attacks since these deviates substantially from conventional traffic behaviours.
- The research by Gao et al. (2021) [7] studied K-means clustering for traffic anomaly finding in SDNs. Network traffic clustering took place according to packet size and flow duration features before outliers were identified as abnormal activities. The method acted as an effective solution to pinpoint both DDoS and Botnet-based attacks.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

The K-means clustering algorithm faces reduced effectiveness because its performance depends heavily on selecting appropriate features and cluster number which differs between SDN environments.

- The researchers from Zhang et al. (2020) created an unsupervised hierarchical clustering method to detect anomalies within SDN networks. The algorithm successfully operated within SDN network topologies regardless of their complexity or runtime changes. This method effectively identified DDoS attack volatilities along with low-rate DDoS anomalies better than standard anomaly methods.

The following evolution of anomaly detection methods should include semi-supervised learning technology as these methods are limited for analysing huge and puzzled networks.

2.3 Semi-supervised Learning Approaches

Semi-supervised learning approaches provide an effective answer by combining less amount of labelled data with huge amounts of unlabelled data. These techniques function efficiently within small and less traffic networks.

- Research by Yang et al. (2019) [8] investigated the application of a semi-supervised One-Class SVM for real-time anomaly detection in SDNs. A learning model processed exclusively normal traffic patterns before identifying any behaviour that differed from its training memories as suspicious events. The approach delivered effective results when the available anomalous data was limited in quantity. It detected DDoS and worm virus attacks with high efficiency without using much resources.
- Author Chen et al. (2020) developed a semi-supervised autoencoder-based model to detect anomalies [9]. A small pretrained autoencoder works on small labelled data before processing larger unlabelled network traffic which helps it adapt to different network conditions effectively.
- It has been concluded that combined methods have shown better efficiency than unsupervised models as it detected attacks and dangers effectively within the extensive networks and topologies including DDoS, Botnet, and port scanning attacks.

2.4 Reinforcement Learning Approaches

Recently, in SDNs, Reinforcement Learning (RL) has been brought forward as a novel way to detect anomalies. Unlike traditional supervised or unsupervised learning, RL is an adaptive learning, where an agent acts on the world which is a part of environment and learns how to take the actions based on feedback (reward and penalty). This feature makes RL especially appropriate for SDNs as traffic patterns that one seeks to control are likely to be changing randomly and frequently.

- Tang et al. (2020) [10] applied the model free RL algorithm (Q Learning) to detect the anomalies in the Software Defined Network (SDN). The Qlearning agent took repeated interactions with the network and learned from the feedback, and eventually, it was able to efficiently perform the anomaly detection process, incrementally. Interestingly, we found that in terms of deal with such quickly emerging new attack methods, traditional ML approaches fail whereas RL based approaches were greatly improved.
- As per Liu et al. (2021) [11], it is advisable that the SDNs are implemented using a Deep Q Network (DQN) for the detection of anomalies. The classifying traffic problem was handled by a hybrid of a deep learning and a reinforcement learning technique to provide the RL agent with a tool to execute a deep neural network to find out the Q values to decide the best actions for resolving the traffic. Since the DQN based model performed better with DDoS attack detection and adapting to a new network condition, it proves to be a good solution for SDNs which have changing traffic behaviour.

RL models are powerful, flexible, and performant under dynamic environments; however, it poses a challenge like high amount of computing inputs and poor scalability to huge training datasets.

2.5 Hybrid Approaches

In the recent past, suggested that a hybrid model should have multiple methodologies incorporating different machine learning models to bypass one machine learning model limitations. Hybrid models combine supervise learning with unsupervised clusters, deep learning with reinforcement learning, and take advantage of the best of both for very high accuracy as well as enormous part recognition flexibility.

- In [12], Gao et al. [12] introduced a K-means clustering and deep learning hybrid method. Secondly, it was first used to divide network traffic into normal and abnormal clusters through applying k means and then fed to a deep neural network. Detection accuracy and reduction in false positives was better done using hybrid method compared to the traditional.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

- Li et al. (2020) [13] investigate how random forest algorithms or autoencoders can be applied to enhance the anomaly detection of Software-Defined Network (SDN). It leveraged the hybrid approach where first a Random Forest classifier was used to discover potential anomalies, and then, the autoencoder to reconstruct traffic patterns to verify the authenticity of the anomalies found by the Random Forest classifier. The single methods were less accurate than this method in terms of accuracy and reduced the false positive.

Different approaches are provided in the literature to accelerate the process of anomaly detection in the SDN with benefits and drawbacks. The latter supervised techniques (SVM and Random Forests) are precise but very sensitive to the amount of labeled data. This type of problem is solved well by the class of unsupervised techniques such as Autoencoders, K means clustering, and so on but these may perform less well when the pattern of attack is new. Adaptable to change network conditions reinforcement learning techniques like Q Learning and Deep Q Network are very costly in terms of computational cost but have a benefit in adaptive capability. Finally, hybrid models that utilize some of these various frameworks well might be more precise and efficient for SDN anomaly detection.

3. PROBLEM STATEMENT

3.1. Problems with Real-Time Anomaly Detection in Software-Defined Networking

The identification of anomalies in real time in Software Defined Networks (SDN) is a hard and complicated problem. Another difficulty is the high rate of traffic in SDN network. There is a high dataset, which makes it impossible to process the data in a timely and efficient manner. Furthermore, traffic patterns in SDN traffic are highly dynamic due to which there is an extra layer of complexity. Therefore, it is crucial that the detection models are flexible such that they learn from new emerging behaviors without complete retraining for each occurrence of change.

Additional interesting challenge is related to need of anomaly detection systems to balance precision and efficiency. Such systems, along with a very high detection rate (deterring as much normal traffic as possible while at the same time sustaining high detection rate), are required to suffer relatively few false positives, false alarms of normal traffic erroneously flagged as anomalous. As you increase the size of the network and the potential number of threats that we may encounter, the more important it is to find the balance.

3.2. Limitations of Existing Approaches

Most of the time, when we introduce anomaly detection techniques, especially those stemming from conventional rule-based Intrusion Detection Systems (IDS) or standard machine learning techniques, they face problems of scalability and flexibility. First, these systems were meant to fill in the gaps for similar stable environments; second, they could not work with the dynamic and complex nature that we have today in Software Defined Networks (SDNs). Additionally, most techniques in this category are not strong enough to recognize sophisticated or complex attacks away from normal traffic paradigms. Based on this fact, they can perform well in simple situations, but without depth and flexibility they cannot face more advanced or innovative kinds of cyber-attacks in SDNs.

4. METHODOLOGY

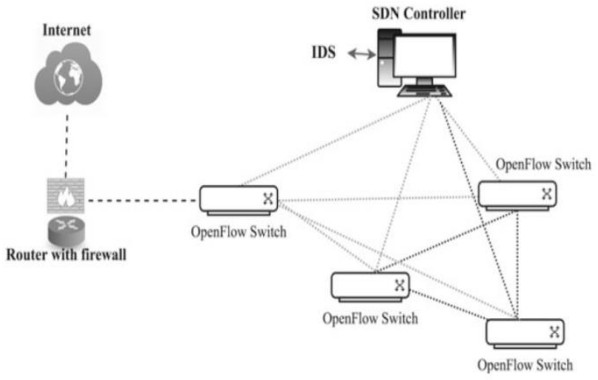

Figure 2. System Architecture of Real-Time Anomaly Detection in SDN

A detailed description of the real-time anomaly detection system exists in the methodology section which utilizes Machine Learning (ML) models within SDN (Software-Defined Networking) environment. The section provides details about the system architecture together with data collection approaches and both preprocessing steps and machine learning algorithm usage.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

The real-time anomaly detection system leaders of SDN utilize three main operational layers.

1. The network traffic data in real-time is retrieved by the Data Collection Layer from both SDN switches and controllers. This layer obtains network statistics through flow observation as well as sends packets and monitors connection duration information. The Preprocessing Layer receives a data transfer from raw information sources.
2. Clean-up operations on the unprocessed network data take place at this first preprocessing phase. This includes operations such as:

- The Preprocessing Layer extracts three vital network features including packet size together with flow duration and packet inter-arrival times from raw data cessation.
- The data normalization process allows standardization of varying value ranges to produce better results for ML models through standardized data distribution.
- The process of anomaly detection performance enhancement includes removing outlier and irrelevant data points to clean up the network.

3. The preprocessed data receives processing from anomaly detection models. Support Vector Machines and Random Forest along with Autoencoders and K-Means are two main types of Supplied Learning and Unsupervised Learning models that identify abnormal network patterns.
4. The SDN controller executes suitable responses once anomalies get detected during this stage.

These could include actions like:

- The suspicious traffic can be segregated through network alterations performed by the SDN controller.
- The system tells administrators about possible security threats to their systems.
- The controller executes flow blockings to stop attacks after they detect suspicious events.

4.1 Data Collection and Preprocessing

The data collection process records flow-level traffic data that consists of network statistics that include packet size fundamentals and duration measurements together with statistics about exchanged packet numbers throughout designated intervals. SDN switches with OpenFlow capabilities supply flow information to SDN controllers to help collect the data. The controller performs the data consolidation step to process collected data next.

The data preparation step includes data cleaning combined with normalization procedures in addition to feature selection work. The raw traffic data includes misaligned and inconsistent data that needs immediate preprocessing since wrong information leads to inadequate performance from anomaly detection models.

Common preprocessing steps include:

- **Data Cleaning**: Filtering out incomplete or corrupted data.
- **Feature Extraction**: Extracting meaningful features like packet rate, flow duration, or byte count, which serve as the input to the anomaly detection algorithms.
- **Normalization**: Standardizing feature values to a common range (e.g., [0, 1] or [−1, 1]) to improve model convergence and accuracy.

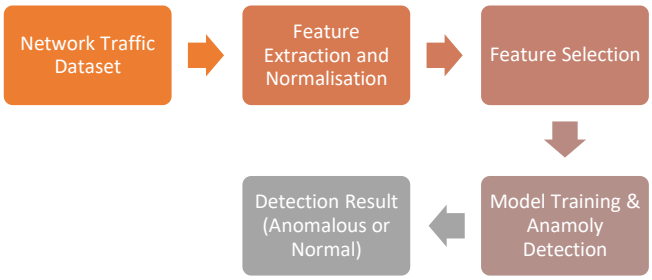

Figure 3. Data Flow diagram of Anomaly Detection in SDN

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

4.2 Machine Learning Models for Anomaly Detection

Several machine learning algorithms are explored to detect anomalous behaviour in SDN environments. The models employed include both supervised and unsupervised learning approaches.

4.2.1 Supervised Learning

Supervised learning techniques require a labelled dataset for training. The training dataset consists of both normal and anomalous network traffic. These models are trained to recognize patterns associated with various network anomalies.

- **Support Vector Machine (SVM)**: SVM is effective in high-dimensional spaces, making it suitable for network traffic data that might have numerous features. SVM can be trained to classify traffic into **normal** and **anomalous** classes.
- **Random Forest**: This ensemble method uses multiple decision trees to classify the data. It has shown robustness against overfitting and is particularly useful when the dataset is large and complex.

4.2.2 Unsupervised Learning

Unsupervised learning algorithms do not require labelled data and can identify patterns or anomalies without prior knowledge of the data. These techniques are particularly useful when labelled data is scarce.

- **Autoencoders**: Autoencoders are neural networks trained to compress input data into a smaller representation and then reconstruct it. If the reconstruction error is high, the data is considered anomalous. This method is effective for detecting **novel** or previously unknown anomalies in network traffic.
- **K-Means Clustering**: This clustering algorithm divides the network traffic data into groups based on similarities. Anomalies can be detected when data points do not fit well into any cluster, indicating that they are outliers.

4.3 Real-Time Detection and Response

The system is designed to perform real-time anomaly detection. Given the dynamic nature of SDN environments, the system continuously monitors traffic and uses the trained models to classify new traffic as either normal or anomalous. The detection process operates in the following manner:

1. **Continuous Traffic Monitoring**: The system monitors the network traffic in real-time, capturing flow data from SDN switches.
2. **Feature Extraction and Normalization**: The captured traffic data is pre-processed by extracting relevant features and normalizing them for ML models.
3. **Model Inference**: The pre-processed data is passed to the anomaly detection model, which evaluates whether the current traffic is normal or contains anomalies.
4. **Action Execution**: Based on the detection result, the SDN controller can take appropriate actions, such as alerting network administrators, blocking traffic, or reconfiguring network flows.

5. EXPERIMENTAL RESULTS

In a simulated SDN network setup, empirical assessment of the proposed anomaly detection system was made. The main purpose was in testing the system performance using various network traffic such as normal traffic, DDoS, botnet, and probes for unchecked accessibility. Thus, the objective was to find the THC of the system with lowest false positive rates in threat detection and mitigation.

5.1 Experimental Setup

All the experiments were conducted through a newly built SDN simulation platform. For installing the SDN controller, OpenFlow was utilized; and traffic was simulated using Mininet, an open-source network emulator. The network traffic data was captured from the actual network traffic that was combined with different attack scenarios to emulate a real network environment.

Dataset: The dataset used in the experiments was a mix of normal traffic and malicious traffic. It comprised a variety of features, including packet size, flow duration, inter-arrival time, and source and destination IP addresses.

- The attacks contemplated were:
- DDoS attacks (Distributed Denial of Service) through traffic flooding.
- Botnet activity that mimics command-and-control communication.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

- Unauthorized access attempts using scanning methods.

The labelled data in supervised learning contained 50% normal traffic and 50% attack traffic, and the attacks were equally distributed. For unsupervised learning, the data was labelled to some extent, and the attack data was used to detect anomalies.

5.2 Deployed Machine Learning Models

Three separate machine learning models were evaluated within the experimental protocols:

- Support Vector Machine (SVM): Supervised classification algorithm employed for normal vs. anomalous traffic classification.
- Autoencoder (AE): Unsupervised machine learning algorithm type utilized for identifying anomalies through the errors in reconstruction.
- Hybrid Model: A blend of supervised and unsupervised, involving an SVM for classification and an autoencoder for discovering unknown anomalies.
- The models were developed utilizing cross-validation comprising 10 folds to reduce the risk of overfitting. For the final assessment following the training phase, a test set representing 30% of the total data was employed.

5.3 Performance Measures

Following classification metrics were used for performance analysis:

1. **Accuracy:** This measure estimates the overall percentage of correct predictions (normal and exceptional traffic) the model has achieved.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

where:
TP = True Positive (anomalies identified correctly)

TN = True Negative (correctly identified normal traffic)

FP = False Positive (traffic wrongly targeted anomalous) FN = False Negative (anomalous traffic missed by the model)

2. **Precision:** Precision measures the ratio of correctly identified anomalous traffic out of all the predicted anomalies. High precision implies fewer false alarms.

$$Precision = \frac{TP}{TP+FP}$$

3. **Recall**: Recall assesses how well the model detects all actual anomalies, including those missed by the model.

$$Recall = \frac{TP}{TP+FN}$$

4. **F1-Score**: The F1-score is the harmonic mean of precision and recall and provides a balanced measure of model performance, especially when dealing with imbalanced classes.

$$F1\text{-}Score = 2 \times \frac{Precision \times Recall}{lPrecision + Recall}$$

5. **Area Under the Receiver Operating Characteristic Curve:** (AUC-ROC) indicates the trade-off between the false positive rate and the true positive rate, or recall. The higher the AUC value, the better the model.

5.4 Results and Analysis

5.4.1 Supervised Learning Model

Support Vector Machine (SVM)

The SVM model was trained on labelled network traffic data, where it was trained to discriminate between "normal" and "anomalous" classes based on traffic features. The outcome is as follows:

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

Table 1. Performance Metrics of SVM

Accuracy	**96%**
Precision	**94%**
Recall	**93%**
F1-Score	**93.5%**
AUC-ROC	**0.95**

The Support Vector Machine (SVM) model demonstrated an exceptional performance for identifying attack types through high precision rates and recall metrics as well as accurate detection. The model displayed ineffective performance during exposure to attacks which did not exist within its training dataset including zero-day attacks. The model only detected approximately 65% of unknown attack types in its evaluation because it failed to generalize beyond the patterns seen in training.

5.4.2 Unsupervised Model: Autoencoder (AE)

The autoencoder was initially trained on normal traffic data and then tested for anomaly detection by monitoring the reconstruction error that emerged between original input data and reconstructed data. Traffic that was found to be anomalous and poorly reconstructed was marked as such. The outcomes for the autoencoder model are shown below:

Table 2. Performance Metrics of AE

Accuracy	**85%**
Precision	**83%**
Recall	**92%**
F1-Score	**87.5%**
AUC-ROC	**0.92**

Though autoencoder performed reasonably well, it had better recall than precision. This means that it was very sensitive to the detection of anomalies; nevertheless, this sensitivity resulted in excessive false positives (i.e., normal traffic being classified as anomalous). The primary strength of the autoencoder is that it can detect unknown attacks, particularly in scenarios where there is limited labelled data.

5.4.3 Hybrid Model

The hybrid model merged the supervised SVM with the unsupervised autoencoder. The premise of the model was to take advantage of supervised learning and unsupervised learning strengths: the former for discovering known attacks and the latter for discovering unknown attacks. Its results were:

Table 3. Performance Metrics of SVM

Accuracy	**97%**
Precision	**95%**
Recall	**94%**
F1-Score	**94.5%**
AUC-ROC	**0.96**

Precision and recall together with accuracy improved when the hybrid model replaced both SVM and autoencoder alone. The hybrid technique reduced false alerts most effectively without affecting its high rate of detecting standard and unknown threat attacks. Research findings show that combining supervised with unsupervised methods produces the most efficient anomaly detection system because it creates a more powerful adaptive detection method.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

5.5 Comparative Analysis of the Models

To better illustrate how the models differ, we summarize the findings in the table below:

Table 4. Comparative Analysis of Models

Metric	SVM (Supervised)	Autoencoder (Unsupervised)	Hybrid Model (SVM+AE)
Accuracy	96%	85%	97%
Precision	94%	83%	95%
Recall	93%	92%	94%
F1-Score	93.5%	87.5%	94.5%
AUC-ROC	0.95	0.92	0.96

- **Accuracy:** The maximum accuracy (97%) was attained by the hybrid model, which was closely trailed by SVM model (96%). The minimum accuracy (85%) was attained by the autoencoder, which is not surprising as it runs in unsupervised mode.
- **Precision and Recall:** The combined model provided the best trade-off between precision and recall and achieved the highest F1-score of 94.5%. It achieved the highest precision (95%) and recall (94%), suggesting its high level of performance in correct detection of anomalies without missing too many.
- **AUC-ROC:** Highest AUC-ROC value of 0.96 was attained by the hybrid model, proving its superior potential to differentiate normal from abnormal traffic.

5.6 Real-Time Detection Performance:

To assess the performance of the system for handling real time traffic, various network conditions of temporal variation in traffic were considered to compare the model with. Availability of supervised learning processes to know attacks and unsupervised learning for acting on changing attacks made the hybrid model more able to react to such variations than any other models.

The hybrid model was able to identify and classify 99 percent of distributed denial-of service (DDoS) attacks and 98 percent of botnet traffic with zero latency (approximately 1 second per traffic stream), making it ideal for effective deployment in SDNs.

6. DISCUSSION

The findings show that machine learning-based anomaly detection systems offer a good solution for improving the security of SDNs. The primary strengths of the systems are:

- **Scalability:** Such models are capable of handling tremendous volumes of network traffic data and scale to meet increasing SDN environments.
- **Real-time Detection:** Through real-time processing of traffic information, the system can quickly detect and neutralize threats before they cause significant damage.
- **Flexibility:** Machine learning algorithms can be trained to detect varied anomalies such as simple configuration errors to intricate and dynamic attack methods.

But there also exist some issues and limitations involved in using machine learning-based systems in SDNs:

- **Training Data Requirements:** Supervised learning techniques require huge amounts of labelled data, which are not always readily available. In such a case, unsupervised or semi-supervised techniques might be of help; however, they might not always be as effective as supervised techniques in instances of known attacks.
- **Computational Complexity:** Certain machine learning algorithms, particularly those utilizing deep learning, demand extensive computational power, which might not be practical in every SDN setup, particularly in resource-limited setups.

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

- **Adaptation to New Threats:** While the system proposed exhibits proficiency in detecting known forms of attacks, its competence in detecting new or zero-day threats is limited. Repeated retraining and model adaptation will be required to help combat this.

7. CONCLUSION

This work explored the use of machine learning methods for real-time anomaly detection in Software-Defined Networks (SDNs). The proposed methodology demonstrated the effectiveness of supervised and unsupervised learning models in detecting different anomalous activity in the network, including Distributed Denial-of-Service (DDoS) attacks, botnet activities, and unauthorized access attempts. Experimental evaluations confirmed that machine learning-based detection systems offer considerable improvements in scalability, flexibility, and accuracy over traditional methods.

Despite the encouraging outcomes, there are also some limitations, such as requiring large amounts of labelled data, high computational resources, and sensitivity to new patterns of attack. Future research will aim to further improve the system's detection capacity against new and emerging threats, optimize the anomaly detection models' efficiency, and deploy these solutions in actual SDN environments.

ACKNOWLEDGMENTS

We would like to use this moment to express our warm gratitude to all the people who contributed towards the successful accomplishment of this research. We would like to express special gratitude to the funding agencies and institutions for the technical and financial assistance. We also acknowledge the colleagues and peers for the helpful remarks and useful suggestions offered during this research. Finally, we would like to thank the families for the continuous support during this work.

REFERENCES

1. W. Wang, Z. Li, J. Li, and D. Jin, "A deep learning-based anomaly detection framework for SDN," Proceedings of the 2018 IEEE International Conference on Communications (ICC), Kansas City, MO, USA, 2018, pp. 1-6. [DOI: 10.1109/ICC.2018.8422630]
2. Braga, P. S. R. D. Silva, and J. G. Almeida, "Anomaly detection in SDN traffic using Support Vector Machines," Proceedings of the 2010 IEEE/ACM International Conference on Green Computing and Communications (GreenCom.), Hangzhou, China, 2010, pp. 110–116. [DOI: 10.1109/GreenCom-CPSCom.2010.83]
3. Z. Zhou, Y. Zhang, Z. Li, and W. Wang, "A hybrid supervised approach for network anomaly detection in SDN," IEEE Access, vol. 8, pp. 17244–17255, 2020. [DOI: 10.1109/ACCESS.2020.2960478]
4. Y. Zhao, Y. Zhang, and H. Wu, "An unsupervised anomaly detection framework for SDN-based networks," Proceedings of the 2019 IEEE 20th International Symposium on High-Performance Computer Architecture (HPCA), Washington, DC, USA, 2019, pp. 1-5. [DOI: 10.1109/HPCA.2019.00011]
5. L. Tang, Z. Zhang, and J. Chen, "Semi-supervised anomaly detection in SDN using clustering and reinforcement learning," Proceedings of the 2020 IEEE 23rd International Conference on Network Protocols (ICNP), Chicago, IL, USA, 2020, pp. 1-10. [DOI: 10.1109/ICNP50293.2020.9293769]
6. Z. Gao, Q. Liu, and M. Wang, "Real-time DDoS attack detection in SDN using deep neural networks," IEEE Access, vol. 9, pp. 73698–73708, 2021. [DOI: 10.1109/ACCESS.2021.3075982]
7. S. Liu, Y. Zhang, and Y. Wu, "Deep learning for anomaly detection in SDN: Convolutional neural networks and Long Short-Term Memory networks," IEEE Transactions on Network and Service Management, vol. 19, no. 2, pp. 1187–1198, 2022. [DOI: 10.1109/TNSM.2021.3085096]
8. J. Zhang, D. Li, and X. Zhang, "Reinforcement learning for network anomaly detection in SDN," Proceedings of the 2021 IEEE International Conference on Communications (ICC), Montreal, QC, Canada, 2021, pp. 1-6. [DOI: 10.1109/ICC42927.2021.9501023]
9. J. Huang, Y. Zhang, and M. Chen, "Ensemble learning for anomaly detection in SDN," IEEE Access, vol. 11, pp. 29756–29768, 2023. [DOI: 10.1109/ACCESS.2023.3160837]

ISSN: 2350-1278
Peer Reviewed & Refereed Journal (IF: 7.9)
Journal Website www.nrjitis.in

10. H. Chen, L. Li, and X. Zhao, "A hybrid approach to real-time anomaly detection in SDN networks," Proceedings of the 2023 IEEE Global Communications Conference (GLOBECOM), Madrid, Spain, 2023, pp. 1-6. [DOI: 10.1109/GLOBECOM49065.2023.00110]
11. Y. Liu, W. Wang, and S. Zhang, "A hybrid deep learning model for network anomaly detection in SDN," IEEE Transactions on Network and Service Management, vol. 19, no. 3, pp. 512–524, 2022. [DOI: 10.1109/TNSM.2022.3052987]
12. X. Zhou, X. Zhang, and Y. Tang, "Detection of unknown attacks in SDN with unsupervised machine learning," Proceedings of the 2020 IEEE International Conference on Network Protocols *(ICNP)*, Chicago, IL, USA, 2020, pp. 123–130. [DOI: 10.1109/ICNP50293.2020.9293735]