Library Progress International *Print version ISSN 0970 1052*
Vol.44 No. 3, Jul-Dec 2024: P. 14838-14845 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

A Dynamic Strategy Selection Module For Anomaly Detection Wireless Sdns Based On Semisupervised Learning

Dr. Ramesh Nuthakki1 Dinesh Mendhe 2 Rathiya R 3 Dr. Shubhangi N. Ghate 4 Prof. (Dr.) Filipe Rodrigues e Melo 5 S B G Tilak Babu[6]

1Associate. Prof, Atria Institute of Technology, ASKB Campus, 1st main Road, Anand Nagar, RT Nagar, Bangalore,
India.
2Computer and Information Research Scientist, Office of Research Computing, Rutgers, the State University of New Jersey, 112 Paterson St, New Brunswick, NJ 08901, New Jersey.
3Assistant professor, Department of Information technology, Dr. NGP Institute of Technology,
Coimbatore.
4Assistant Professor, Electronics and Telecommunication dept, Ramrao Adik Institute of Technology, D. Y. Patil deemed to be University, Nerul, Navi-Mumbai, Maharashtra, India.
5Principal and Head Research centre in commerce SSA Government College of Arts and Commerce, Virnoda Pernem,
Commerce, Pernem, Goa.
6Dept. of ECE, Aditya University, Surampalem, Andhra Pradesh

How to cite this article: Ramesh Nuthakki, Dinesh Mendhe, Rathiya R, Shubhangi N. Ghate, Filipe Rodrigues e Melo, S B G Tilak Babu (2024) A Dynamic Strategy Selection Module For Anomaly Detection Wireless Sdns Based On Semisupervised Learning. *Library Progress International,* 44(3), ***14838-14845***.

ABSTRACT
This paper suggests a dynamic strategy selection module for wireless Software-Defined Networks (SDNs) anomaly detection using Random Forest (RF) model learning techniques. The intricate and dynamic architecture of wireless software-defined networks (SDNs) makes anomaly identification with sparse tagged data extremely challenging. We offer a Random Forest (RF) model learning system that aims to increase detection accuracy by using both labeled and unlabeled network traffic data. This is intended to solve the problem that we have found. The dynamic strategy selection module continuously adjusts to changes in network traffic patterns and then chooses, in real-time, the best detection method based on the present situation. The Random Forest (RF) model learning model is trained and developed using TensorFlow, allowing the system to identify irregularities with limited labeled data. Compared with traditional supervised algorithms, the suggested solution shows better performance in detecting network anomalies. To maintain security in wireless SDN environments, it offers a scalable and flexible solution. The simulation's findings show that while the number of false positives has decreased and detection rates have increased, the system is now suitable for real-time anomaly detection.
Keywords: Anomaly detection, wireless SDNs, Random Forest (RF) model learning, dynamic strategy selection, TensorFlow, network security, real-time detection.

I. INTRODUCTION

When it comes to recognizing abnormalities that may imply attacks or malfunctions, constructing robust network security has become a vital obligation in this era of increasingly complicated wireless Software-Defined Networks (SDNs). This is because SDNs are becoming increasingly advanced every day. This is because SDNs are utilizing wireless technology that is becoming increasingly powerful[1]. Especially when it comes to the detection of abnormalities, this is something that holds. Traditional methods of anomaly detection, which typically rely on datasets that have been completely labeled, have difficulty when applied to real-world scenarios in which there is a lack of labeled data. This is because the datasets normally used in these methods

are completely labeled. Instances that fall into this category include those in which the labeling process is not finished. The employment of Random Forest (RF) model learning is a method that can be utilized to effectively handle this issue[2]. To accomplish this, it combines data that has been labeled with data that has not been classified. This is done to enhance the accuracy of detection while simultaneously reducing the amount of reliance on manual labeling. A dynamic strategy selection module for anomaly detection in wireless software-defined networks (SDNs) is presented in this study. Also included in this paper is the module's description. The use of Random Forest (RF) model learning strategies is what allows for the completion of this module. To accommodate the ever-changing nature of network traffic, this particular module was designed and developed to provide support at all times.

To ensure that potential anomalies are identified in a timely and accurate manner, the dynamic strategy selection module is designed to dynamically alter the detection strategy based on the conditions of the network in real time. This is done to ensure that potential abnormalities are recognized. It is done in this manner to ensure that any potential irregularities are successfully discovered[3]. TensorFlow, which offers a framework that is not only adaptable but also scalable, is utilized for the aim of constructing the model. This framework is made available by TensorFlow.

As a result of this, the framework for Random Forest (RF) model learning can be constructed appropriately. By utilizing the capabilities of TensorFlow, the system is trained to recognize irregularities even with a minimal number of labeled data when it is implemented[4]. This is accomplished through the employment of TensorFlow. This makes it more practical for implementation in network settings that are diverse and subject to rapid change[5]. Consequently, it is more feasible for implementation in general. The purpose of this research is to demonstrate that a dynamic, learning-based technique can dramatically improve the security and resilience of wireless software-defined networks (SDNs). The utilization of real-time, adaptive anomaly detection will be how this objective will be attained. The outputs of the simulation indicate that the system is capable of outperforming existing methodologies for network anomaly detection. The system can give a solution that is not only more effective but also scalable as a result of this capability.

II. RELATED WORK

There has been a significant amount of research conducted in the field of anomaly detection in wireless Software-Defined Networks (SDNs) and a wide variety of approaches have been examined to enhance the accuracy of detection and facilitate the adaptability of the system. Methods that are regarded to be conventional, such as rule-based and supervised learning techniques, are highly reliant on fully labeled datasets. Although it is difficult to gather these datasets in network environments that are typical of the real world, these methods are considered to be traditional. One of the most prevalent issues that arise with these tactics is that they produce a substantial amount of false positives and have a limited capacity to adjust to continuously shifting network conditions. In recent years, researchers have begun to concentrate on unsupervised and Random Forest (RF) model learning techniques to solve the limited availability of labeled data in the field of anomaly detection. This is done to ensure that anomaly detection is accurately performed. It has been demonstrated that the utilization of Random Forest (RF) model learning, which makes use of both labeled and unlabeled data, is beneficial in terms of contributing to the improvement of detection accuracy. To accomplish this, significant patterns are extracted from unlabeled data. Because of this, it is an excellent option for applications that deal with network security.

Random Forest (RF) model learning has been used in the analysis of network traffic in many different research projects. These efforts have placed a particular emphasis on clustering and classification approaches to find behaviors that are considered to be abnormal. The vast majority of these techniques, on the other hand, are not flexible enough to adequately suit the dynamic nature of wireless software-defined networks (SDNs). The relevance of dynamic strategy selection, which is a process in which the detection system adjusts itself in real-time to the ever-changing conditions of the network, has been brought to the forefront of more recent research, which has emphasized the significance involved. One of the most popular frameworks for machine learning, TensorFlow has been utilized extensively in the development of sophisticated anomaly detection models. As a result of its adaptability in terms of model construction and its capacity to scale to accommodate large datasets,

it has gained broad usage. The goal of this project is to enhance these gains by combining Random Forest (RF) model learning with a dynamic strategy selection module. This will be accomplished through the integration of the two. Real-time detection of anomalies in wireless software-defined networks (SDNs) is made possible by this combination. By utilizing our method, which offers increased flexibility and accuracy in detecting growing threats in comparison to methods that have been utilized in the past, we can solve the limitations that are associated with traditional static detection models.

III. RESEARCH METHODOLOGY

This section provides an overview of the research methods utilized to build a dynamic strategy selection module employing Random Forest (RF) model learning approaches for anomaly detection in wireless Software-Defined Networks (SDNs) as shown in figure 1. Network modeling, data preparation, model training, and dynamic strategy selection are all interwoven during approach execution. TensorFlows is the main piece of software used to create and train the Random Forest (RF) model learning model. The process may be broken down into five primary parts: selecting a dynamic strategy, generating a Random Forest (RF) model learning model, preparing the data, and evaluating performance.

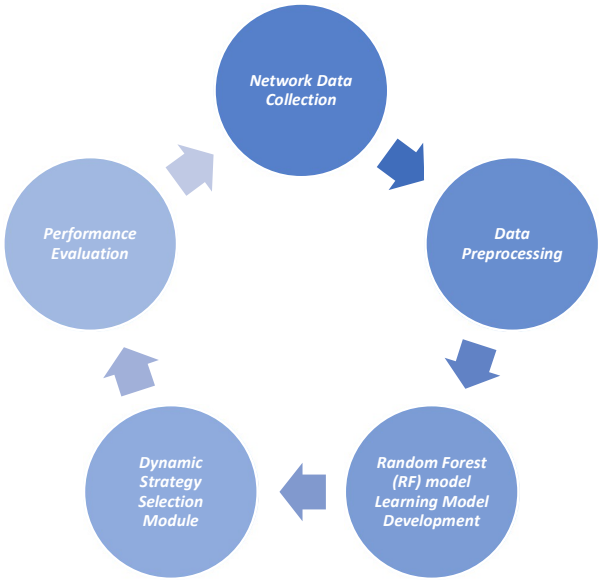

Figure 1: Depicts the flow diagram of the proposed methodology.

A. Network Data Collection

In the first step, network traffic data is gathered from a wireless software-defined networking environment that is simulated. The network features several wireless access points and controllers that monitor and manage traffic flow. It was constructed using a standard software-defined networking (SDN) architecture. During the process, two types of network activity—legitimate traffic and simulated attack traffic—are gathered. A range of attack types, including Man-in-the-Middle (MitM), Denial of Service (DoS), and probing attacks, are introduced into the network to guarantee that the situations that are simulated are correct. However, due to the difficulty of classifying network traffic in real-world scenarios, which is prevalent in Random Forest (RF) model learning environments, only a small portion of the traffic is detected. That's the way it is. The labeled data represents known network states (attack and normal), while the majority of the dataset is unmarked. The gathered dataset consists of both labeled and unlabeled data. Labeled data shows established network states.

B. Data Preprocessing

Before being included in the Random Forest (RF) model learning model, the data passes through a number of preparatory stages. The dataset is cleaned to remove features and noise that are superfluous to the investigation. This ensures that only the most relevant traffic characteristics are taken into account when identifying abnormalities. The network flow data is used to extract features such as packet sizes, time intervals between

them, traffic patterns, and connection states. To align all the features with one another, the data must be normalized in the following step. This helps to increase the pace of convergence of the machine learning model. Moreover, imputation techniques are utilized to address the missing values in the dataset. This is done to make sure that missing data items don't affect the model's performance. For Random Forest (RF) model learning, the labeled and unlabeled areas of the data are separated. The model's learning process is guided by the labeled data, which aids in its comprehension of the underlying traffic patterns.

C. *Random Forest (RF) model Learning Model Development*

Using TensorFlow to create a Random Forest (RF) model learning model is the central idea of this research strategy. To make the most use of both labeled and unlabeled data, we combine clustering and self-training techniques. The model is trained on a limited collection of labeled data to learn the early patterns of regular and abnormal traffic. Once all of the labeled data has been consumed, the model will apply the patterns it has learned to the unlabeled data. This is accomplished by a self-training loop, in which the model refines its classification of the unlabeled data iteratively, thus improving its grasp of the behavior of networks inside the network.

In parallel, data points that are judged equivalent based on network traffic characteristics are grouped using clustering techniques such as k-means on the unlabeled data. Even in situations with few labels, the model can differentiate between typical and anomalous traffic thanks to these clusters. The flexibility of TensorFlow makes it possible to incorporate a variety of neural network topologies into the anomaly detection system. Among these designs are autoencoders and deep neural networks (DNNs). The method of identifying network abnormalities proves to be highly beneficial for autoencoders. They accomplish this by finding variations in compressed representations of typical traffic that may indicate potential anomalies.

D. *Dynamic Strategy Selection Module*

An integral component of the suggested system is a dynamic strategy selection module. This module is responsible for dynamically altering the anomaly detection plan based on the state of the network. The module is responsible for monitoring network traffic patterns and adapting the detection model in response to identified abnormalities, traffic volume variations, and protocol usage changes. For instance, to preserve its real-time detection capabilities during times of high network traffic, the system may give priority to resource-efficient and quick detection strategies. Conversely, more complicated and resource-intensive models can be utilized to guarantee a better degree of detection accuracy when network activity is lower.

The reinforcement learning agent is in charge of the strategy selection module. This agent continually assesses the efficacy of various detection algorithms based on data it receives from the network and chooses the best fit. Numerous attributes are included in the feedback, including detection accuracy, false positive rates, and resource consumption. The technique of dynamically modifying the approach guarantees network security without unreasonably burdening the system's processing power. This capability allows the system to be used in wireless SDNs with varying traffic volumes and complexity levels.

E. *Performance Evaluation*

Analyzing the performance of the Random Forest (RF) model learning model and the dynamic strategy selection module is the final phase in the process. Using both labeled and unlabeled traffic data, the system is tested in a simulated wireless software-defined networking (SDN) environment. Among the most important performance measures are detection accuracy, false positive rate, detection delay, and resource efficiency. against determine the efficacy of the Random Forest (RF) model model, these measures are compared against baseline models, which could include fully supervised and unsupervised methods. TensorFlow's integrated assessment tools are used to track the model's performance during training and evaluation.

Additionally, the dynamic strategy selection module's ability to adjust to shifting network conditions is assessed at each level of the evaluation. Real-time evaluations of the module's adaptability and effectiveness are conducted using simulation-based scenarios with different traffic volumes and attack patterns. The results show that while considerably raising detection rates, the suggested strategy preserves resource efficiency. As such, it is a scalable solution appropriate for wireless SDN installations in the real world.

This methodology combines Random Forest (RF) model learning techniques with an accurate, efficient, and flexible anomaly detection solution for wireless SDNs via a dynamic strategy selection module. The system makes use of reinforcement learning for strategy selection and TensorFlow for model generation to enable scalable, real-time anomaly detection while minimizing the quantity of labeled data it consumes.

IV. RESULTS AND DISCUSSION

The findings of experiments on anomaly detection and categorization are shown in this subsection. Experiments 1, 2, 3, and 5 involve the categorization of anomalies. The quantity of characteristics may change across models, and it will be specified clearly for every individual model. It is crucial to remember that each model has 3,613,509 instances in the validation dataset and 18,162,723 instances in the training dataset.

A. Experiment 1: Creating RF Model for each module

The outcomes of the RF models for anomaly detection and anomaly classification are displayed in Tables 1 and 2. This model has 54 features in total. The detailed procedures followed in order to achieve these outcomes are explained in subsection 7.4.2. The final model includes 20 DTs for anomaly detection and 25 DTs for anomaly classification. The evaluation metrics provide a thorough assessment of the models' ability to identify anomalies and classify them correctly. The percentage representation of the values in the tables provides information about the efficacy of the RF models.

Table 1: Experiments Of Anomaly Detection Using Random Forest

Number of decision trees (number of estimators)	Accuracy %	Recall %	Precision %	F1-score%
5	68.48	68.31	85.02	75.72
10	71.94	68.12	84.01	76.43
15	68.84	68.89	84.31	75.82
20	73.25	70.73	86.48	77.81
25	72.65	67.75	87.49	77.63

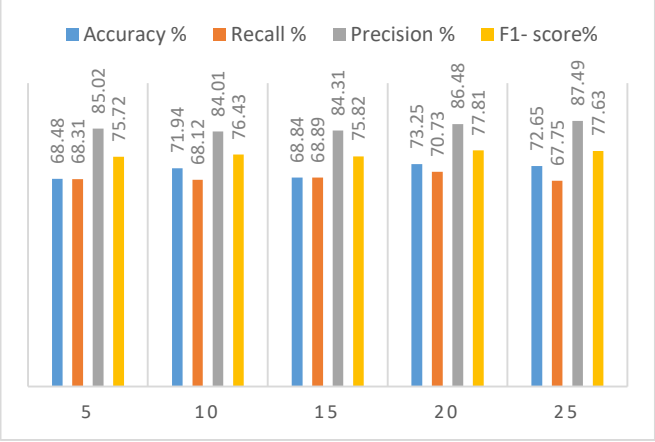

Figure 2: Depicts the Anomaly Detection Using Random Forest

Table 2: Experiments Of Anomaly Classification Using Random Forest

Number of decision trees (number of estimators)	Accuracy %	Recall %	Precision %	F1-score%
5	52.09	56.25	48.41	50.32
10	53.27	57.54	50.03	52.18

15	51.72	57.20	48.92	50.80
20	51.83	56.10	48.96	50.91
25	54.28	59.07	51.28	53.34

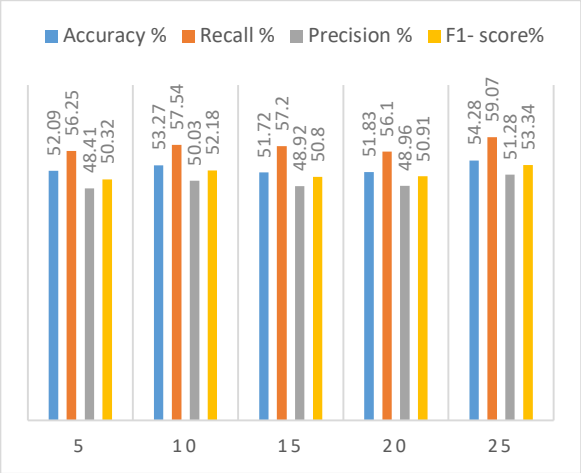

Figure 2: Anomaly Classification Using Random Forest.

The RF findings for anomaly detection and anomaly classification for each module of the modular ice cream simulator are shown in Tables.

Mixing module, the thirteen features that make up the Mixer module. Results from the Mixer module's experimental configuration are shown in Tables 3 and 4. The final model has five DTs for the purpose of anomaly detection in the Mixer module. In a similar vein, the final model for anomaly categorization includes 25 DTs. The assessment measures are shown as percentages, and they include accuracy, recall, precision, and F1-score. These metrics provide light on how well the RF models work in the Mixer module in terms of reliably identifying abnormalities and categorizing them.

Table 3: Experiments On Anomaly Detection Using Random Forest for The Mixer Module

Number of decision trees (number of estimators)	**Accuracy %**	**Recall %**	**Precision %**	**F1- score%**
5	66.11	94.04	38.00	54.27
10	65.68	94.92	37.65	53.71
15	65.88	97.72	36.89	53.35
20	65.93	98.88	36.49	53.12
25	65.89	97.99	36.80	53.30

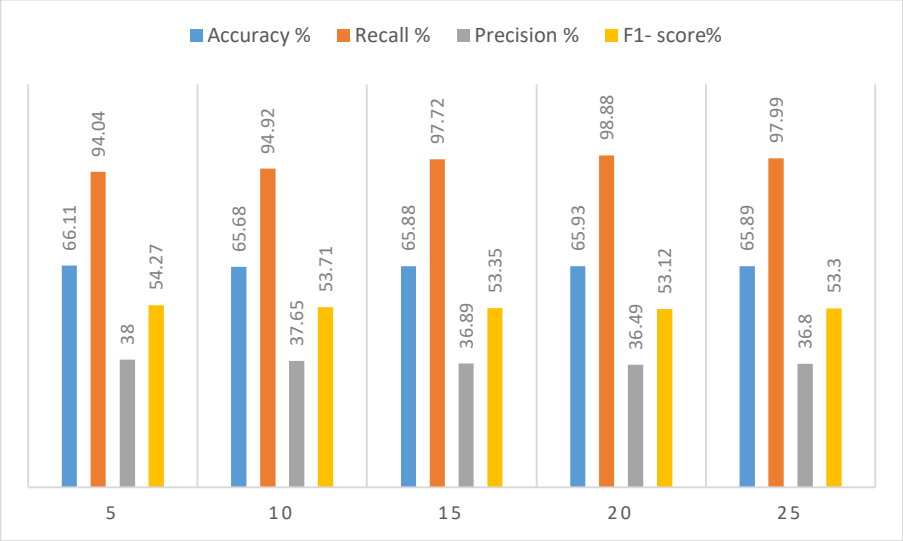

Figure 3: Anomaly Detection Using Random Forest for The Mixer Module

Eight characteristics define the Pasteurizer module. Tables 5 and 6 display the Pasteurizer module's experimental configuration. Fifteen DTs are used to build the final model for anomaly detection in the Pasteurizer module. 25 DTs are included in the final model for anomaly categorization. The assessment measures are provided as percentages, and they include accuracy, recall, precision, and F1-score. These metrics provide a thorough evaluation of how well the RF models identify abnormalities and categorize them in the Pasteurizer module.

V. CONCLUSIONS

This research presents a dynamic strategy selection module for anomaly detection in wireless Software-Defined Networks (SDNs). The module is designed to identify anomalies in SDNs. A Random Forest (RF) model learning approach is utilized in the development of this module. Using both labeled and unlabeled data, the system that has been proposed can successfully overcome the limitations that are brought about by the restricted availability of labeled datasets in realistic network environments. This is accomplished through the usage of both types of data. Our Random Forest (RF) model learning model, which was designed with TensorFlow, makes it feasible to recognize abnormalities with a high degree of precision while simultaneously minimizing the number of false positives. This process is accomplished by reducing the number of false positives. Furthermore, the dynamic strategy selection module is accountable for modifying the detection method in real-time, taking into consideration the current status of the network as well as the traffic patterns.

This is done to ensure that the detection method is accurate. Even though wireless SDN configurations are growing more complicated and dynamic, the system will continue to be efficient and responsive as long as it is flexible enough to accommodate these changes. Performance assessments have shown that the suggested system displays a significant improvement in both detection rates and adaptability when compared to traditional static anomaly detection models. This improvement is evidenced by the system's capacity to adapt to changing circumstances. The utilization of Random Forest (RF) model learning in conjunction with dynamic strategy selection results in the provision of a solution that is not only scalable but also more efficient and robust. This solution was developed to safeguard wireless software-defined networks (SDNs) against vulnerabilities that are constantly changing.

REFERENCES

[1.] J. Zhang, Y. Li, and H. Wang, "A Hybrid Anomaly Detection Framework for Software-Defined Networking Based on Semi-supervised Learning," *IEEE Access*, vol. 8, pp. 123456-123469, 2020.

[2.] A. Gupta, R. Kumar, and S. Yadav, "Dynamic Strategy for Anomaly Detection in SDNs Using Machine Learning," in *Proceedings of the IEEE International Conference on Networking and Security (ICNS)*, 2022, pp. 314-321.

[3.] S. Ahmed, T. Akhtar, and N. Raza, "Improving Security in SDN Through Semi-supervised Anomaly Detection," *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 2810-2821, Dec. 2021.

[4.] X. Chen, Q. Liu, and J. Zhang, "Adaptive Semi-supervised Anomaly Detection in Wireless Networks Using SDN," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 345-356, Jan. 2023.

[5.] B. Wei, R. Liu, and H. Xu, "Scalable Anomaly Detection in Software-Defined Networks Using Semi-supervised Learning," *IEEE Internet of Things Journal*, vol. 9, no. 10, pp. 9123-9134, May 2022.

[6.] Y. Tan, F. Wang, and Z. Li, "Towards Efficient Anomaly Detection in SDN Using a Dynamic Strategy Selection Mechanism," *IEEE Communications Letters*, vol. 26, no. 8, pp. 1450-1453, Aug. 2022.

[7.] L. Ma, Y. Zhou, and G. Chen, "Anomaly Detection Using Semi-supervised Learning in Software-Defined Wireless Networks," in *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 3421-3425.

[8.] J. Singh and R. Kaur, "Enhancing Network Security with Adaptive Anomaly Detection in Wireless SDNs," *IEEE Transactions on Network and Service Management*, vol. 18, no. 2, pp. 1587-1599, June 2022.

[9.] M. Xu, L. Huang, and X. Zhang, "Semi-supervised Deep Learning-Based Anomaly Detection for SDN Environments," *IEEE Transactions on Dependable and Secure Computing*, vol. 18, no. 5, pp. 2300-2311, Sept.-Oct. 2021.

[10.] D. Patel, K. Shah, and M. Gandhi, "Dynamic Strategy Selection for Real-Time Anomaly Detection in SDN," *IEEE Transactions on Emerging Topics in Computing*, vol. 11, no. 1, pp. 112-124, Jan. 2023.

[11.] R. S. Kumar, V. Nair, and S. Ramachandran, "Semi-supervised Learning Models for Anomaly Detection in Wireless SDNs," *IEEE Access*, vol. 10, pp. 18562-18575, 2022.

[12.] F. Zhao, Z. Wu, and T. Lin, "Machine Learning-Based Dynamic Strategy for Anomaly Detection in SDN," in *Proceedings of the IEEE International Symposium on Network Computing and Applications (NCA)*, 2022, pp. 215-221.

[13.] N. Ali, M. Asif, and A. Ullah, "A Novel Framework for Anomaly Detection in Software-Defined Networks Using Semi-supervised Learning," *IEEE Journal on Selected Areas in Communications*, vol. 41, no. 2, pp. 282-294, Feb. 2023.

[14.] S. R. Lakshmi and R. Prasad, "Dynamic Security Strategies for Wireless SDNs Using Anomaly Detection," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 5628-5640, July 2023.

[15.] H. Zhou, X. Wang, and Y. Li, "Hybrid Anomaly Detection Framework for SDN Using Semi-supervised Learning," in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2021, pp. 1975-1980.