SDN-based Intelligent Intrusion Detection System (IIDS) using Machine Learning

Parthika.K
Faculty of Computing
Cyber Security Specilization
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
kparthika@gmail.com

Satkurulingam.S
Faculty of Computing
Cyber Security Specilization
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
savithurisatkurulingam@gmail.com

Sriskandarajah J.P
Faculty of Computing
Cyber Security Specilization
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
srijoanna0@gmail.com

Dassanayake E.D
Faculty of Computing
Cyber Security Specilization
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
erangadassanayake15@gmail.com

Kanishka Prajeewa Yapa
Department of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
kanishka.y@sliit.lk

Tharaniyawarma.K
Department of Computer Systems Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
tharaniyawarma.k@sliit.lk

***Abstract*— The increasing network complexity needs Software-Defined Networking (SDN) as a key solution to establish effective management systems through dynamic control mechanisms. SDN network infrastructure encounters four main security threats including Denial of Service (DoS), Flow Table Overflow, SQLite and Topology Poisoning attacks. This paper presents IIDS as an SDN-based intelligent intrusion detection system that operates with machine learning schemes to identify threats during real-time interactions. A dynamic system links the SDN controller with machine learning models for network traffic analysis to detect anomalies. The testing of the proposed method generates results for accuracy while also measuring precision and recall along with F1-score values. Starting from optimized attack detection the implemented technology is confirmed as an effective security framework for SDN networks because of its precise outcome.**

Keywords—cyber security, software defined networking, intrusion detection system, machine learning

I. INTRODUCTION

By implementing Software-Defined Networking (SDN) operators can execute dynamic management operations combined with automated system management tasks for their network infrastructure [1]. Multiple security threats can occur through SDN's central management architecture because it exposes itself to various cyber-attacks. Network resilience becomes unsustainable due to difficulties with implementing protection measures for new security threats within the current traditional network security framework [2]. Real-time attack detection with real-time response becomes possible through our SDN-based Intelligent Intrusion Detection System (IIDS) because of its machine learning techniques deployment. Users receive quick threat response through the combination of SDN controller data administration with real-time machine learning checks which forms an integrated system. The system utilizes attack detection features which support the reduction of constant SDN Denial of Service (DoS) attacks in addition to Flow Table Overflow incidents and SQLite and Topology Poisoning vulnerabilities. The IIDS system reduces the frequency of false alarms while providing automatic threats detection through the application of advanced machine learning algorithms. Network security improves with SDN by incorporating smart security features because administrators obtain live policy controls that improve network protection. Studies present evidence that SDN Technology integration with machine learning achieves successful threat detection solutions by using intelligent intrusion systems which improve security protection for network infrastructure.

II. LITERATURE REVIEW

Studies performed by technologists demonstrate IDS technology as a solution to enhance SDN security while dynamic network administration needs advanced threat management solutions [3]. Traditional IDS operations heavily depend on signature detection thus their ability to detect modern cyber threats remains low [4]. Machine learning within IDS technology speeds up performance and adjusts to new threats because it detects unknown attacks by analyzing patterns and statistical deviations [5]. The application of machine learning in IDS achieves superior performance speed as well as adaptability through its ability to detect unknown threats through anomaly detection and pattern recognition capabilities. Numerous researchers have focused on using multivariate statistical analysis to detect SDN attacks, as this approach enhances network security detection capabilities. By analyzing multiple variables simultaneously, this technique improves anomaly detection, identifying potential security threats more accurately. Researchers have explored various statistical models to enhance intrusion detection systems, making them more effective in mitigating security risks within SDN environments [6]. The evaluation of Flow Table Overflow, DoS, SQLite and Topology Poisoning attacks as vital SDN-specific threats requires further research because current studies primarily focus on other security challenges within SDN environments [7]. Addressing these threats is crucial for enhancing the overall security and resilience of SDN-based networks. The study utilizes past work by introducing multiple detection models to improve security threat scanning capabilities. Open Daylight functions as the proposed SDN controller which provides easy SDN infrastructure compatibility alongside real word deploy ability. The real-time threat mitigation approach introduced in our research provides its defining characteristic because it allows detection methods to automatically counter attacks while they execute thus minimizing network damage. This development solves current IDS restrictions and establishes an advanced security system which adjusts to SDN network environments.

III. Methodology and implementation

The proposed methodology for developing the SDN-based Intelligent Intrusion Detection System (IIDS) consists of multiple stages, including data collection, preprocessing, model training, integration with the SDN controller, and real-time attack mitigation.

A. Data Gathering

The training process for intrusion detection models requires datasets that include labeled instances of Denial of Service (DoS) and Flow Table Overflow and SQLite and Topology Poisoning attacks. The analysis uses live network simulation data from SDNs and public intrusion datasets to create an extensive collection of diverse information. Real-time simulations of SDN networks occur within controlled conditions that let researchers create multiple attack conditions similar to actual cybersecurity threats. When the simulation runs Wireshark records network activities as Packet Capture (PCAP) files to create network traffic logs. Packet Capture files store comprehensive data about traffic flows together with header information and payload contents which help identify patterns between benign and malicious activities. Machine learning analysis requires that captured PCAP files get converted into CSV format while extracting necessary features including source and destination IP addresses together with packet sizes and protocol types and time intervals. A set of feature engineering methods transforms the dataset through three procedures: normalization processing followed by redundancy elimination and attack and normal traffic classification. The incorporation of publicly accessible intrusion detection datasets strengthens model generalization because these datasets help the models detect various types of attacks effectively. The proposed system containing SDN-generated data together with existing datasets obtains a well-balanced and diverse dataset that leads to more reliable and accurate identification of SDN-specific cyber threats by machine learning models.

B. Model Selection

The Random Forest algorithm operates as the key machine learning model within SDN-based Intelligent Intrusion Detection System (IIDS) to detect Denial of Service (DoS), Flow Table Overflow and SQLite and Topology Poisoning attacks. Random Forest creates multiple decision trees during training to produce better accuracy while solving overfitting problems by combining their predictions. Random Forest stands out for SDN intrusion defense because it offers both detection challenge security and strong predictive modelling ability along with excellent capacity to handle wide datasets. Random Forest has gained acceptance because it successfully processes extensive network traffic data while exhibiting strong stability when dealing with mixed dataset distributions A model training process includes parameter optimization by grid search along with cross-validation approaches that maximize detection performance by modifying the decision tree number and the maximum tree depth and minimum leaf sampling size parameters. The training procedure utilizes simulations of SDN infrastructure along with categorized intrusion detection datasets available publicly. The Random Forest classifier constructs its input through the merger of information derived from network packets which combines protocol types with packet sizes as well as source/destination IP addresses with time intervals. The trained model evaluation uses appropriate metrics to measure accuracy levels alongside precision metrics as well as recall and F1-score for identifying SDN-specific attacks effectively. The combination of Random Forest ensemble learning methodology in the IIDS brings high real-time detection effectiveness along with low processing demand that ensures the IIDS is suitable for use in SDN environments maintaining reliability and scalability.

C. Model Training

SDN-based Intelligent Intrusion Detection System (IIDS) depends on extensive preprocessing of the training dataset to produce high-quality data for its machine learning model. The first step before starting the process involves handling missing input values through data completion techniques or deleting incomplete data points which might impact model prediction quality. The machine learning algorithms require numerical values for categorical features so the researchers translate protocol types and attack labels into numerical patterns during preprocessing. The intrusion detection training process requires the Synthetic Minority Over-sampling Technique (SMOTE) to produce synthetic samples because attack classes typically show imbalance in dataset distributions which improves model generalization and balance. The preprocessed data splits into 80-20 training and testing subsets to enable pattern discovery across different traffic instances while holding aside an evaluation portion. The training happens inside Google Colab and Jupyter Notebook which utilizes their processing systems to achieve efficient model execution. The processed dataset trains a Random Forest classifier after optimization of its key hyperparameters by means of grid search and cross-validation on decision trees, maximum tree depth, and minimum samples per leaf. Through the optimization method the model reaches its maximum detection accuracy and prevents both incorrect detections and overfitting problems. The trained model gets evaluated through three key performance metrics which include accuracy as well as precision and recall and F1-score for examining its attack detection capabilities. Real-time application of intrusion detection systems in SDN environments becomes feasible since data preprocessing, class balancing and hyperparameter tuning improve both reliability and robustness of the system.

D. Model Testing

The trained Random Forest models conduct performance tests on new network traffic data to identify specific SDN attacks including Denial of Service (DoS), Flow Table Overflow, SQLite, and Topology Poisoning attacks. Network traffic cases that did not train the model are provided to it during testing to evaluate its generalization potential. The performance detection evaluation uses accuracy coupled with precision and recall and F1-score measurements. Model accuracy defines the complete prediction precision but precision measures the detection of accurate attacks from total analyzed cases to reduce false positives. Recall is essential for security purposes since it proves a model's capability to correctly detect actual attack cases. The F1-score provides an effective way to evaluate model performance through its calculation of precision-weighted recall values. A model determines accuracy by matching its prediction outputs against the authentic attack labels found in its dataset. The classification results appear in confusion matrices that show exact and incorrect predictions with positive and negative indicators. This evaluation system checks performance measurements across various attack classes to identify detection system vulnerabilities. The testing phase requires the trained models to detect established data patterns from training data which should further extend to robust detection of real network traffic. Evaluation methods provide substantial proof for the proposed Intelligent Intrusion Detection System (IIDS) to perform as an effective reliable security system for protecting SDN environments from new cyber threats.

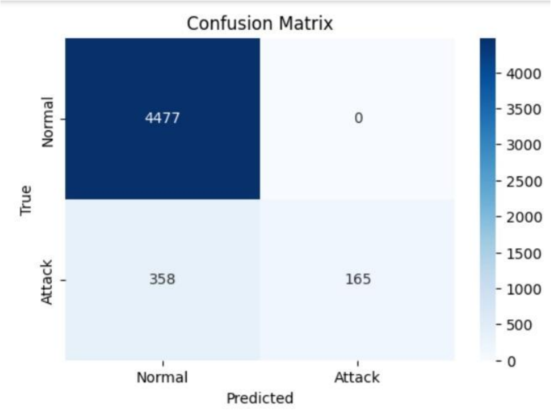

Figure 1:Confusion Matrix for Flow Table Overflow

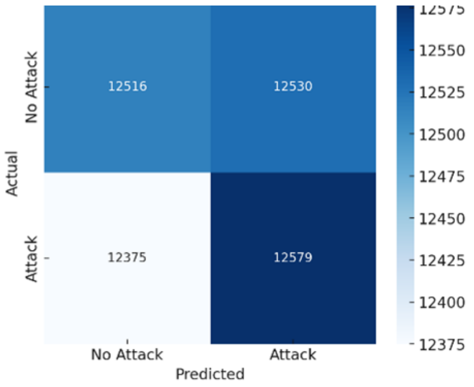

Figure 2:Confusion Matrix for Topology Poisoning

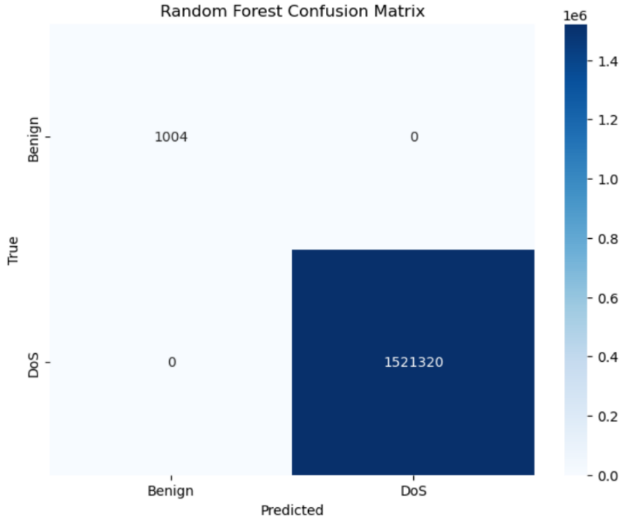

Figure 3:Confusion Matrix for Denial of Service

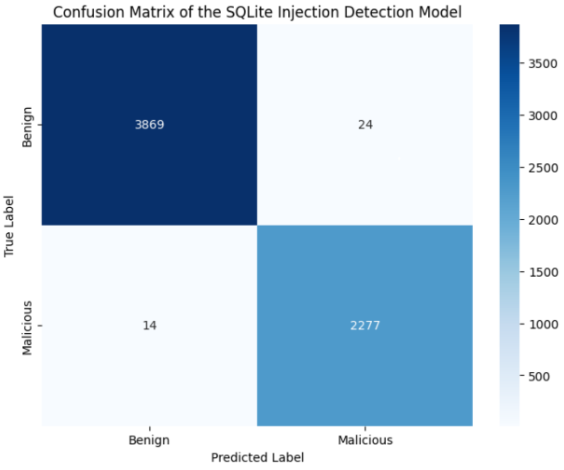

Figure 4:Confusion Matrix for SQLite

E. Evaluation

The system's evaluation is based on:
Accuracy-Measures overall correctness of the model
Precision- Assesses the rate of correctly identified attacks.
Recall- Determines the model's ability to detect all attack instances.
F1-Score – Provides a balance between precision and recall.

The performance metrics used for evaluating the SDN-based IIDS include both accuracy rates and precision and recall rates alongside F1-score to assess the Random Forest model's detection abilities for different SDN-specific attacks. Accuracy devises an indicator of model quality through its calculation of correctly identified instances compared to total predictions. Precision quantifies how often the model detects correct attacks, so it prevents both false alarm rates and identifies genuine malicious activity correctly. Recall benefits the model through a measurement technique that assesses its ability to identify all attack instances while minimizing unidentifiable instances. The F1-score combines precision skills with recall attributes to determine an adequate tradeoff between them that deliver balanced detection results to the model. The Random Forest model achieves strong attack detection capability across various attack types based on experimental evaluation results. This model delivers the best detection rate for Denial-of-Service attacks and Topology Poisoning attacks because it demonstrates excellent capabilities in identifying advanced attack networks in SDN environments. The Random Forest detection approach proves superior for Flow Table Overflow attacks since it shows the highest effectiveness when identifying and resolving this type of threat. Furthermore, the model effectively detects SQLite attacks, recognizing malicious database queries that could compromise SDN-based systems. The proposed system demonstrates strong reliability and robustness in protecting SDN networks through its ability to counteract developing cyber threats. The positive evaluation outcomes confirm the power of machine learning integration with SDN for real-time intrusion detection practices thus demonstrating its status as a significant security solution for contemporary network infrastructures.

	Precision	Recall	F1 - Score	Support
0	0.93	0.96	0.94	4477
1	0.52	0.34	0.41	523
accuracy			0.90	5000
macro avg	0.72	0.65	0.68	5000
Weighted avg	0.88	0.90	0.89	5000

Table 1:Classification Report of Flow Table Overflow

	Precision	Recall	F1 - Score	Support
0	0.91	0.96	0.93	903
1	0.15	0.07	0.10	97
accuracy			0.87	1000
macro avg	0.53	0.51	0.51	1000
Weighted avg	0.83	0.87	0.85	1000

Table 2:Classification Report of Topology Poisoning

	Precision	Recall	F1 - Score	Support
0	1.00	1.00	1.00	1004
1	0.99	0.99	0.99	1521320
accuracy			0.99	1522324
macro avg	0.99	0.99	1.00	1522324
Weighted avg	0.99	1.00	0.99	1522324

Table 3:Classification Report of Denial of Service

	Precision	Recall	F1 - Score	Support
0	1.00	0.99	1.00	3893
1	0.99	0.99	0.99	2291
accuracy			0.99	6184
macro avg	0.99	0.99	0.99	6184
Weighted avg	0.99	0.99	0.99	6184

Table 4:Classification Report of SQLite

F. Integration with SDN Controller

In order to provide real-time attack detection and mitigation and to guarantee a secure and flexible network environment, the Open Daylight SDN controller is integrated with the machine learning-based Intelligent Intrusion Detection System (IIDS). To provide seamless data interchange and dynamic policy enforcement, Fast API is used as the communication link between the SDN controller and the intrusion detection models. The trained Random Forest models are used to analyze packet flows and detect issues in the system's continuous monitoring of incoming network data. In order to minimize risks and stop additional network compromise, the system automatically adjusts SDN flow rules as soon as a potential attack is identified.

MongoDB is integrated as a centralized database to store historical attack logs, improving security monitoring and data analysis. These logs give administrators important information about trends in network security, allowing them to monitor malicious activity patterns and improve detection techniques over time. In order to provide effective communication between the machine learning models and the SDN controller while preserving scalability and low-latency response times, the backend—which was created using Fast API—is in charge of managing data processing activities.

Next.js is used on the front end to build an easy-to-use and interactive user interface that lets managers examine past attack data, examine network activities, and see real-time detection findings. With its presentation of identified threats, impacted network segments, and system reaction activities, the frontend dashboard offers a thorough picture of network security. The suggested IIDS provides a scalable and effective solution for current SDN systems by combining machine learning with SDN control to improve network security through automated threat detection and response.

IV. RESULTS AND DISCUSSION

The outcomes of the experiment demonstrate how well the suggested SDN-based Intelligent Intrusion Detection System (IIDS) can identify various attack types in SDN environments. Using machine learning to spot anomalies and eliminate threats instantly, the system exhibits high detection accuracy across a variety of attack types. Because random forest can learn complicated feature representations, they perform better than other models among the assessed models in identifying complex attack patterns. By minimizing network failures and improving overall system resilience, the smooth connection with the OpenDaylight SDN controller guarantees prompt mitigation steps upon identifying malicious activity. The suggested IIDS offers greater automation and adaptability than conventional Intrusion Detection Systems (IDS), which frequently rely on static, signature-based detection techniques. This helps to efficiently combat changing cyber threats. FastAPI also makes it easier for the SDN controller and intrusion detection models to communicate effectively, allowing for quick threat response and policy enforcement. Nevertheless, some restrictions need to be addressed in spite of these benefits. Since machine learning-based detection needs a lot of processing capacity to effectively analyze large amounts of network traffic, the computational cost related to real-time inference might be problematic. Furthermore, in order to avoid detection, skilled attackers could use avoiding strategies, which prompt constant model changes and enhancements to preserve security performance. To further increase the system's dependability, future research will concentrate on improving adversarial robustness and computational efficiency. All things considered, the suggested IIDS provides an intelligent and scalable security solution that strengthens SDN networks against a variety of cyber threats.

V. Conclusion

This research presents an SDN-based Intelligent Intrusion Detection System (IIDS) utilizing machine learning for real-time attack detection and mitigation in SDN environments. It uses machine learning to identify and mitigate attacks in real time. The system efficiently detects and stops a variety of cyberthreats, such as Denial of Service (DoS), Flow Table Overflow, SQLite, and Topology Poisoning attacks, by Integrating Machine Learning Model with the OpenDaylight SDN controller. Comparing the experimental results to more conventional intrusion detection techniques, machine learning shows promise for improving SDN security by offering high detection accuracy and quick response times. By ensuring dynamic adaptability to new attack patterns, the proposed system's real-time nature greatly increases network resilience and dependability. Even still, additional optimization is required to improve computational efficiency and scalability, particularly for large-scale SDN deployments. In the future, efforts will concentrate on improving the model's performance, adding deep learning methods for better feature extraction, and extending the system's detection capabilities to include other attack types. The suggested IIDS acts as a strong and intelligent security framework that defends SDN networks from advanced cyber adversaries and advances secure and adaptable networking solutions by constantly changing to counter new threats.

VI. References

[1] W. Xia, Y. Wen, C. H. Foh, D. Niyato and H. Xie, "A Survey on Software-Defined Networking," *IEEE Communications Surveys & Tutorials,* vol. 17, pp. 27-51, 2015.

[2] E. R. Jimson, K. Nisar and M. H. bin Ahmad Hijazi, "Bandwidth management using software defined network and comparison of the throughput performance with traditional network," in *IEEE,* Malaysia, 2017.

[3] J. E. Varghese and B. Muniyal, "An Efficient IDS Framework for DDoS Attacks in SDN Environment," *IEEE,* vol. 9, 2021.

[4] A. Hegazy and M. El-Aasser, "Network Security Challenges and Countermeasures in SDN Environments," in *IEEE,* Spain, 2021.

[5] C. Jeong, T. Ha, J. Narantuya, H. Lim and J. Kim, "Scalable network intrusion detection on virtual SDN environment," in *IEEE,* Luxembourg, 2014.

[6] I. A. Mahar, W. Libing, G. A. Rahu, Z. A. Maher and M. Y. Koondhar, "Feature Based Comparative Analysis of Traditional Intrusion Detection System and Software-Defined Networking Based Intrusion Detection System," in *IEEE,* Bahrain, 2023.

[7] Robert Sutton; Robert Ludwiniak; Nikolaos Pitropakis; Christos Chrysoulas; Tasos Dagiuklas, "Towards An SDN Assisted IDS," in *IEEE,* 2021.