International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 16, No.1 (Aug-24)

http://dx.doi.org/10.12785/ijcds/160169

Exploring Honeypot as a Deception and Trigger Mechanism for Real-Time Attack Detection in Software-Defined Networking

Harman Yousif Khalid[1] and Najla Badie Aldabagh[2]

[1]*Department of Computer Science, College of Science, University of Duhok, Duhok, Iraq*
[2]*Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq*

Received 30 Mar. 2024, Revised 7 May 2024, Accepted 9 May 2024, Published 10 Aug. 2024

Abstract: Cyberattacks are becoming more frequent and sophisticated, making their detection harder. Probe attacks in Software Defined Networking (SDN) not given much attention by the research community, which represents the starting phase for other attacks. The attacker scans the network to get the necessary details about hosts and services running on it to launch successful attacks exploiting vulnerabilities in the system. The issue with probe attacks is that they occur passively, and the target system is not aware of them. On one hand, an additional mechanism is required to check the network traffic continuously by embedding switches with independent agents, which is against the OpenFlow standard. On the other hand, using statistics provided by OpenFlow switches to the controller, which overloads the controller with the extra task of continuously checking traffic statistics. This work proposes a lightweight detection mechanism that employs machine learning to detect probe attacks in real-time. The detection mechanism integrates a honeypot to detect passive probe attacks, luring attackers with fake services and acting as a trigger mechanism to activate the detection mechanism when needed. The experimental results show that the proposed mechanism successfully detects probe attacks in real-time, achieving accuracy (94.73%) with the minimum CPU load.

Keywords: Intrusion Detection System (IDS), Software Defined Networking (SDN), Probe, Reconnaissance, Honeypot, Machine Learning (ML)

1. Introduction

Probe attacks are often considered a preliminary step in various cyberattacks, particularly network reconnaissance. Network reconnaissance is the first phase of the cyber kill chain, which involves gathering information about a target system or network to identify vulnerabilities and potential entry points for further exploitation [1]. Probe attacks involve port scanning, ping sweeps, network mapping, etc. to discover active hosts, open ports, services, the type of operating system of the target host, etc. Since such information could be useful during attack planning to exploit vulnerabilities associated with a specific OS or understand how different systems are connected. After obtaining the necessary information, attackers may perform vulnerability scanning to find specific weaknesses or vulnerabilities in the target systems and identify potential entry points and weaknesses in the target's network. Upon completion of the initial reconnaissance phase, attackers can proceed with more targeted and specific attacks, such as denial of service (DoS), man-in-the-middle (MitM), and brute force, based on the gathered information.

The issue with probe attacks is that they operate in passive mode and are very hard to detect. Several methods are used by the research community to detect probe attacks, such as deploying Snort Intrusion Detection System (IDS), which continuously monitors the network traffic to detect malicious traffic. However, Snort is not effective to deal with today's attacks because they are signature-based, and any simple deviation in the attack signature could easily be bypassed by the attacker [2] [3]. Moreover, embedding an OpenFlow switch with additional duties is against the OpenFlow protocol since it clearly states that the switch must be as simple as possible and perform only packet forwarding.

SDN's features, such as full control and view over the whole network, as well as programmability, open the door for the deployment of various security applications through the existing open API. Some researchers exploit the power of SDN controllers to deploy artificial intelligence (AI)-based IDS to detect various types of attacks. AI-based IDS methods require feature extraction from networks to check current flows, whether they are malicious or normal. The OpenFlow protocol provides a method to gather various statistics details from switches through OpenFlow statistics message requests and replies. The controller initiates the request message to all switches, and the switches provide

E-mail address: Harman.khalid@uod.ac, Najlabadie@uomosul.edu.iq

the controller with the required details. AI-based IDS take advantage of those details for feature extraction to detect malicious flow. However, periodic inquiries about those features and corresponding replies lead to huge problems for the controller.

This work is extended from a survey [4], where a deep investigation was conducted on the works performed by other researchers regarding the deployment of AI-based IDS in SDN environments; some open issues and challenges were highlighted. The survey shows that, due to the difficulty of traffic feature extractions in SDN and mapping them to the features in the dataset they used, almost all previous studies on deploying IDS neglected real-time attack detection. Their main focus has been on achieving high accuracy by using new machine learning (ML) and deep learning (DL) algorithms or by using feature selection techniques in offline training. The literature's accuracy failed to reflect reality due to its lack of real-time implementation detection work. Furthermore, the survey revealed that the majority of recent work has focused exclusively on DDoS attacks, ignoring other types of attacks because SDN's centralized controllers are a prime target for DDoS attacks. Moreover, the majority of works deployed IDS, checks the network for malicious traffic over a fixed time interval. A longer time interval may make it more difficult to identify the attack in the early stages and may even give the attacker more time to cause more serious damage to the network [2]. If interval is very short, it has many consequences, such as increasing the controller's CPU load, especially in large-scale networks [5].

In this work, we aim to solve the mentioned open issues, and we propose a lightweight machine learning mechanism to detect passive attacks such as probe attacks in an SDN environment using features provided by OpenFlow switches. The mechanism comprises of a triggering module based on honeypots and detection and mitigation modules. The following is a summary of this work's primary contributions:

- Early detection and prevention of attacks since we consider probe attacks, which are regarded as the initial stage of nearly all other kinds of attacks, including DDoS, botnet, MitM, etc. In other words, we break the first phase of the attack kill chain before gaining control of the target system.
- To identify passive attacks, we proposed a lightweight mechanism that exploits the honeypot's capability, which acts as a trap by luring attackers by providing fake services.
- We used the honeypot as an alerting mechanism to minimize the CPU overhead of the SDN controller in a large-scale network by triggering the detection module when needed instead of continuously checking the traffic. Moreover, the honeypot contributes to filtering the flow of traffic by providing additional useful details about the attacker.
- We conducted a simulation scenario to verify the proposed mechanism in real-time, considering attack detection in its early stages as well as attack mitigation.

2. Related Works

The novel SDN architecture allows the research community to take advantage of its features of programmability, flexibility, and ease of deployment. Implementing a security application, which resides in the controller, exploits the power of ML and DL to detect malicious traffic in a network. This section briefly introduces some recent and popular approaches that have been suggested for attack detection in an SDN environment using AI capabilities.

In [6], the Grey Wolf Optimization (GWO) feature selection technique was put into practice in an effort to boost IDS's ability to more precisely identify probe attacks. They discussed how feature selection improves the detection model as a whole. They emphasized that choosing the right features is crucial to cutting down on computation time, which will increase the classifier's accuracy when the best features are chosen and reduce the amount of data needed for training and testing. Furthermore, it is simpler to extract fewer characteristics for real-time detection, which reduces detection time. They demonstrated that accuracy increased to 99.8% when the Light Gradient Boosting Machine (LightGBM) classifier was used to select a subset of 8 characteristics from the InSDN dataset. Accuracy was only 77.3% when all features were used. They did not, however, use real-time detection, and their topology was identical to the dataset's author.

In another direction, some studies have developed hybrid IDS that merge flow-based IDS with signature-based IDS to provide a more robust detection mechanism. The author in [7] implemented two approaches for detecting DDoS attacks in SDN. First, they use signature-based Snort IDS alongside SDN to analyze the network for malicious traffic. Second, implementing a machine learning Support Vector Machine (SVM) model trained with NSL-KDD dataset to detect unknown attacks. The motive behind using both methods is that the unknown attacks are detectable by machine learning, and their signatures will be stored in the Snort database to make Snort able to detect them next time they occur again. The drawback of this method is that they implemented Snort as an independent hardware module connected to a switch, which requires additional resources and whose performance degrades when the network is larger.

In the same context, the author in [8] used Snort IDS, which connects to an Open vSwitch and monitors the network through port mirroring. They also provide flow-based IDS in the controller to overcome the shortcoming of Snort being unable to detect novel attacks. The OpenFlow statistics message was used to extract features for machine learning over a certain time interval. They selected seven features that can be easily obtained by SDN nature. However, using a combination of both Snort and machine learning leads to issues in large networks and creates a load.

The detection of DDoS attacks in SDN environments has been covered in [9]. The author proposed two lines of security. First, they used Snort to detect known attacks in signature-based databases. The second defense line was using ML and DL to detect anomaly-based attacks. They used SVM and Deep Neural Network (DNN) models, which trained on the NSL-KDD dataset, and the accuracy was 74.3% and 92.3%, respectively. Snort is used in this work for detecting known attacks in signature databases and as a data collector for ML and DL models. However, they periodically monitor the network for anomalies, which makes their detection mechanism active even when there is no traffic overload on the controller.

In [10], a comparative analysis of different feature selection algorithms for detecting DDoS attacks based on various machine learning models is presented. The experiment was conducted using feature selection algorithms such as Information Gain (IG), Correlation Coefficient, Chi-square, Forward Feature Selection (FFS), Backward Feature Selection (BFS) and Recursive Feature Elimination (RFE). Machine learning classifiers such as SVM, Decision Tree (DT), Random Forest (RF), Naïve Bayes (NB), and K-Nearest Neighbor (KNN) are used for binary classification. The optimal model was RF with an accuracy of 99.97% using a feature subset of 28 selected by the RFE. They mentioned that the detection model used OpenFlow statistics messages to get 41 features of NSL-KDD. However, it is very difficult to get those numbers in real time, as highlighted before.

In order to categorize DDoS attacks in an SDN context, an attack detection and mitigation mechanism was presented in [11]. It used a hybrid model of Convolutional Neural Network (CNN) and Extreme Learning Machine (CNN-ELM). Their model used data from the SDN environment to detect DDoS attacks. Both packet-in messages and statistics messages supplied to the controller by OpenFlow switches included this information. Twelve features were mapped from the InSDN dataset to the extracted features of the OpenFlow switch. They also created additional features. The results of the experiments demonstrated that utilizing a selection of 12 attributes reduced test times while simultaneously improving accuracy. But since each packet-in message needs to be examined, this mechanism adds overhead to the controller and is ineffective in the event of a DDoS attack. Furthermore, their methodology was not validated, and it was unclear how features would be taken out of packet-in. Plus, the four features that were manually constructed were also unverified.

Similar to this, the Deep Convolutional Neural Network (DCNN) was suggested by the authors in [12] as a way to identify DDoS in SDN. They offer identical mitigation and detection methods as previous studies. With the exception of sending periodic OpenFlow statistics messages to the controller for anomaly identification, they utilized just the features offered by the flow table. They stated that they only gathered data from the OpenFlow switch using 12 of the InSDN's mapped features. They claim that the current system has issues with using a lot of features for deep learning or machine learning and that more functions are required to extract them, which increases latency and congestion on the network. On the other hand, relying solely on a small set of features is insufficient for accurate assault detection. However, during actual usage, they employed 78 features—rather than just 12—for training. Mapping the 78 properties of the InSDN dataset to the fundamental features offered by OpenFlow switches is challenging. Furthermore, the controller incurs expenses due to their technique, which necessitates the controller inspecting each packet-in message and frequently seeking statistics from the switch. As with earlier studies, features were not verified and the process of extracting them from packet-in was not clearly described.

To prevent DDoS attacks, a lightweight machine learning model was proposed in [13] that uses flow fluctuation as a single feature, which represents the number of packet-in messages repeatedly sent to the controller over a fixed time slice. To implement the proposed methodology, they created their own dataset. However, the InSDN dataset was utilized exclusively for model training and testing. Concentrating on a single feature is justified on the grounds that it will be more straightforward to acquire and necessitate reduced time and resources for real-time prediction and training. Concentrating on a single feature is advantageous because it simplifies acquisition and reduces the time and resources needed for real-time detection. Seven features from the InSDN dataset were utilized, and multiple machine learning models were implemented. The findings of the research suggest that Binary Tree (BT) and KNN exhibited the highest levels of effectiveness with regard to accuracy. However, KNN demonstrated superior performance in terms of training time, CPU utilization, and decision time as well. They assessed the proposed work using their own dataset and obtained a 99.4% accuracy rate with BT using a single feature. As stated by the author, an excessive number of features may either improve the performance of certain models or cause them to become overfitted. However, the authors neglected to specify the feature selection methods that were employed. Furthermore, a number of the selected features, including timestamps and flow-ID, were deemed irrelevant and potentially hindered the learning process throughout model training. In addition, the decision time is prolonged due to the time slice methodology and the controller's continuous verification of the packet-in count. In conclusion, the utilization of a single attribute for training the intrusion system produces minimal possibilities for success.

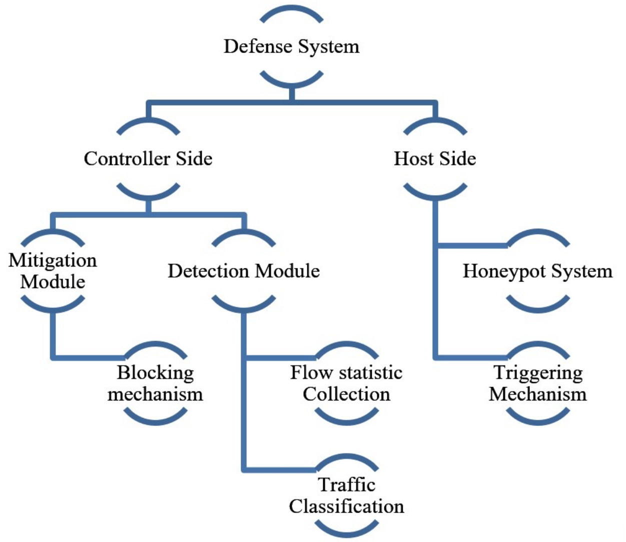

Figure 1. Defense System Modules Diagram.

In summary, the majority of the existing work in the literature targets DDoS attack detection only, neglecting other attacks. The focus of the previous works was on the analysis of the proposed models to achieve high accuracy or implementing some studies on feature selection algorithms, and the real-time detection process was not adequately explained or neglected. Moreover, some works monitor packet-in messages as well as statistics messages to construct their features, but none of them considers the overload that is created on the controller when it periodically checks the network traffic or extracts traffic features. In this work, we consider the probe attack, which is considered as the first step of other attacks, and extract a limited number of features through the statistics provided by OpenFlow switches to ensure a fast and lightweight detection mechanism. We consider the tradeoff between high accuracy and low controller overhead by using a triggering mechanism for checking malicious traffic when needed instead of periodically. We implemented our own dataset, which contains feature dimensions that are easily obtained from OpenFlow switches, and evaluated the efficiency of the dataset using many common supervised learning algorithms such as DT, RF, Adaptive Boosting (AdaBoost), NB, XGBoost, and KNN.

3. The Proposed Detection Mechanism

In this section, we describe the methodological stages followed to build the proposed probe detection mechanism. The SDN features make SDN operation easier and offer a number of benefits [14]. This motivates the deployment of light, effective, and attack detection in real-time. Our proposed SDN defense system is distributed on two sides, as shown in Figure 1. On the controller side, where the detection and mitigation modules reside, as well as on the host side, where honeypot resides as a trap to lure attackers and notify the controller when possible malicious traffic is detected. Honeypot is a deception mechanism used to lure attackers by providing fake services. Any contact with the honeypot is considered a possible attack. This contact will trigger the detection module in the controller to start checking the current flows in the network instead of continuously checking the network periodically. In addition, the honeypot is programmed to send some useful details to the controller with a triggered message about the attacker traffic to filter out the possible malicious flows and reduce the load on the CPU.

Machine learning techniques have recently dominated IDS research because they produce more accurate predictions than other techniques [15]. Machine learning models can overcome the drawbacks of other methods by classifying abnormal traffic as an anomaly with self-learning capabilities [16]. ML has been deployed in a wide area, from medical analysis and image processing to data mining. The concept of machine learning is to make machines learn automatically from the given training data without human intervention [17]. However, machine learning required input values to feed to create prediction output based on those inputs. These inputs should be extracted from the live attributes of traffic flow to accurately detect malicious traffic. SDN controllers' beneficious features, such as power and storage provided, open programmability, global visibility and control, and statistics features provided by the OpenFlow switch, make them suitable locations for implementing machine learning intrusion detection and mitigation applications. Due to the availability of those features, in the proposed work, machine learning models are implemented in the controller for classification.

As shown in Figure 2, the detection and mitigation applications are inactive in the controller, waiting to be triggered to check traffic flow when necessary. Honeypot is actively providing some fake services to lure attackers along with real services on the network. The defense mechanism can be described in pseudocode in Figure 3. When the attacker first initiates the probe attack to check for active hosts and services in the network, the honeypot is configured to register all events on a dedicated log file for every connection attempt to the services they provide. In general, the log file is used for analyzing the attacker's movements and steps during the attack.

However, in this work, we utilize the honeypot to be exploited as a triggering mechanism to alert for possible attacks. Once the log file registers an event, a special Python script monitors the log file and extracts the IP of the attacker, the protocol of flow, and the port used by the attacker. Later, the mechanism constructs a special packet-in message to be sent by the OpenFlow switch to the controller with a special ethertype field code '0x88FF' to be recognized by the controller as a trigger message. In the next step, once the controller receives the trigger message, it will forward it to the detection and mitigation application. The application creates a flow statistics request message

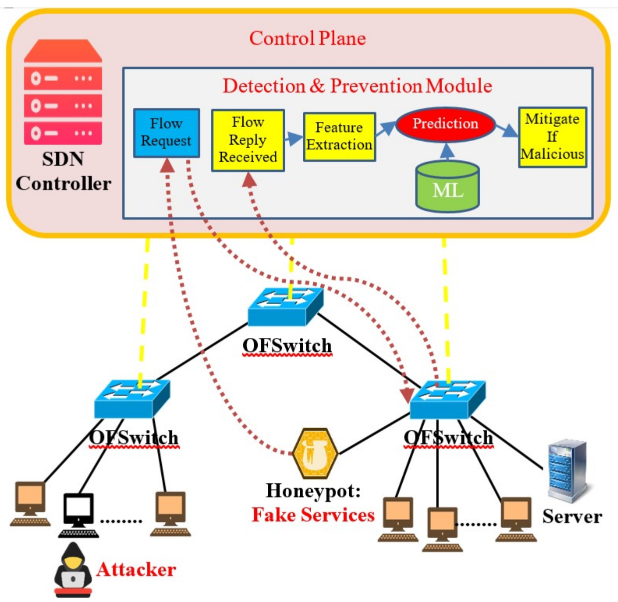

Figure 2. The Framework of The Proposed Mechanism.

```
If packet_in message, then
        If Ethertype = '0x88FF'
        Craft OFP stats request to switch(ID)
if OFP stats reply message
        Filter according to (IP, port, protocol)
        Extract features from filtered flow
        Result= ML_Model (Extracted features)
        If Result ==0 then
                Take no action
        Else: Result ==1 then
                Block the attacker port
```

Figure 3. Defense System Pseudocode.

to be forwarded to the switch, which is connected directly to the host where honeypot exists through the switch ID provided by the packet-in trigger message. After receiving the reply messages from the switch, which contain flow entry statistics, they are filtered by other details provided by Honeypot, such as IP, port, and protocol. Then the necessary features are extracted from the filtered flows and fed to the ML module for testing. Once the malicious flow is detected, the attacker's domain is located, and the controller initiates a packet-out message to block the attacker's source port.

A. Dataset Preparation

The accuracy of the predictions generated by machine or deep learning-based IDS is significantly influenced by the use of high-quality datasets during model training. Comparing to other domains, such as computer vision, there are limited datasets for intrusion detection in general and for probe attacks specifically due to privacy and legal issues [2] [18]. In the past few decades, academics have used a number of well-known public datasets, including the KDD Cup, NSL-KDD, UNSW-NB15, and CICIDS2017 databases. Since SDN is a new paradigm, compared to traditional networks, there are fewer research studies that address the intrusion detection problem in SDN. According to the authors of [15] [19] [20], the majority of published research addresses SDN intrusion detection, which is identical to traditional network intrusion detection.

Scholars employed classical network datasets, like KDD Cup and NSL-KDD, as training datasets for anomaly detection in SDN contexts. Training the SDN-based IDSs using old datasets can cause significant issues, as they only detect attacks, which have similar behavior in both SDN and traditional networks. Because intrusion attack tactics are always evolving, they are getting more complex and difficult to spot [19]. Since these datasets were published so long ago, they are either old or unreliable [15] [21], or they suffer from compatibility issues since they were collected during the traditional networks and the SDN architecture is different. As a result, using them for real-time detection is ineffective [21].

The InSDN dataset [22] was published as an SDN-specific dataset in 2020 for the purpose of assessing and training IDS within SDN environments. The work in [4] conducted regarding this dataset reveals that applying this dataset for IDS deployment in SDN research has mostly focused on achieving high accuracy by using new machine or deep learning algorithms or by training offline while employing feature selection algorithms. The literature's accuracy failed to reflect reality due to its lack of real-time functionality and the absence of a clearly defined implementation of real-time detection work. This is due to the struggle of extracting the necessary features from the traffic flow in order to match them to the features in the dataset.

In this work, we focus on a binary classification for detecting probe attacks and do not delve further into classifying the various types of attacks. In this work, we created our own dataset specifically for training and evaluating the proposed IDS for detecting probe attacks in the SDN context. The dataset contains 282128 instances in total, with sample sizes of 126685 (45%) for the normal class and 155443 (55%) for the probe class, respectively. The probe traffic was collected by initiating a live host scan, port scan, version scan, service scan, OS detection, etc. For normal traffic, we replicated the proper random inter-departure time and packet size by creating traffic at the packet level as well as multiple sessions of parallel, sequential, and bidirectional streams. Numerous widely used application services, including HTTPS, HTTP, FTP, SFTP, Telnet, DNS, VoIP, and others, are represented in the normal traffic. All hosts connected to the internet and ordinary daily tasks executed by executing several applications, such as browsing different websites, watching YouTube videos, and sending emails, to mimic real internet traffic.

Deploying AI-based IDS requires features from real-time traffic to feed the classifier model for decision-making in real-time. The first task of the IDS is to collect traffic statistics and related information about flows. Many studies have implemented ML and DL models in SDN, but finding the data source that offers the attributes needed for real-time detection is a difficult task. Either they had to use methods used in traditional networks by using dedicated tools in the forwarding plane acting as sensors like Snort and sFlow agents, which degrade performance, or using features provided by SDN. Nowadays, networks are faster due to the requirements of applications and advances in technology. The intrusion detection mechanisms need to consider their performance in terms of detection time. The demand for lightweight IDS is a hot topic in academia nowadays. The trade-off between model accuracy and execution must be addressed carefully. Choosing the best features is essential to improving SDN-based IDS efficiency.

Having a large dataset with a huge number of features will delay training, testing, and model detection. In addition, using a high-dimensional dataset may not necessarily lead to higher accuracy due to overfitting and redundant features [23]. Reducing the number of features or selecting relevant features in general will lead to a decrease in model training, testing, and detection time, as well as model complexity, since real-time feature extraction is not an easy process [2] [23]. Moreover, feature selection reducing high dimensionality reduces the likelihood of overfitting issues in the model [2]. In order to obtain excellent model performance utilizing ML and DL, many researchers have turned their attention to the removal of noisy, duplicate, and useless features [2] [24] [25].

B. Feature Extraction

In the SDN framework, there are some points of feature extraction that can be beneficial for feeding the model, such as:

1) Controller packet-in messages: Upon receiving the initial packet of a particular flow, the OpenFlow switch will examine the packet header and determine whether any flow rules exist in one of its flow tables that correspond with the flow [26]. There is an action connected to every flow rule. The traffic flow will be redirected to the appropriate destination if a match is identified. When there is no match and a miss table appears, the switch uses the packet-in message to convey the packet to the controller, who processes it and determines what to do with it. Some researchers used packet-in behavior to construct some useful features that successfully detect DDoS attacks, as in [11] [12] [13] [27].
2) OpenFlow statistical messages: Basic information about the flows could be obtained by parsing the OFP_Stats_Reply messages of an SDN network [21]. These messages were sent to the controller in response to the OFP_stats_request by the controller. Usually, this happens periodically, depending on the specific event triggered. This process does not require additional effort because it is provided by the controller according to the OpenFlow specification.
3) Using independent agents: using other methods for capturing network traffic, like in [27], which combined an OpenFlow switch and sFlow for effective anomaly detection in an SDN environment [28]. In addition, others in [9] have used Snort as a data collector.

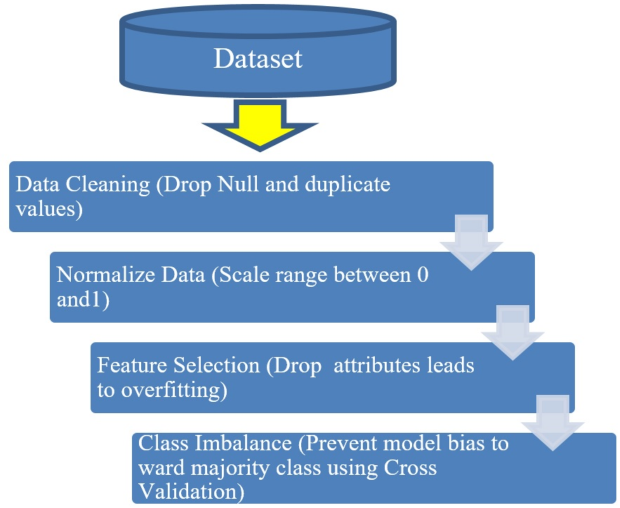

Figure 4. Data Preprocessing Steps.

In our mechanism, we used the OpenFlow statistical messages method to obtain the required features since they are easily obtained in an SDN environment and do not require an additional agent extended to the data plane switch following the OpenFlow protocol specification. The features extracted are shown in Table I.

C. Preprocessing Stage

An essential step in preparing the input data for the model's training in order to create an accurate detection system is data preprocessing. To prevent the overfitting issue, the created dataset does not include socket information such as source IP, destination IP, flow ID, etc.[2], where such data can be changed from network to network. In addition, since we aimed to record the features related to the behavior of attacker flow, we excluded source and destination ports because they are often constant values representing network identifier attributes, which may lead to overfitting. As the dataset does not contain a large number of samples, we deployed ML mechanisms instead of DL, since the latter proved to be effective with large datasets [18] [19]. As a first step in preprocessing, as depicted in Figure 4, the dataset is checked to see if it contains empty or null values to prevent any significant effect on the model efficiency. Moreover, duplicated samples were removed from the dataset. Normalization, which involves transforming data to a common scale, is an essential preprocessing step in machine learning. We used Min-Max scaling to make values between "0" and "1" to ensure that all features have the same scale, making it easier to compare their importance and contributions to the model. We divided the dataset, where 80% is used for training and the other 20% is kept aside for the model test to see how well we can predict the data.

We performed cross-validation with k = 5 to make sure there is no bias in performance estimation to prevent an imbalanced occurrence where the minority class is significantly underrepresented compared to the majority class. The fundamental principle of cross-validation is to divide the dataset into several folds, or subsets, and train and test the model repeatedly on various combinations of these folds.

D. Practical Implementation

The proposed work is implemented using the Python programming language. Mininet version 2.3.1, beta 4, a network simulator designated for SDN education and research, was utilized to implement and evaluate the suggested technique. Mininet was installed on a virtual machine, Ubuntu 20.04, with 12 GB of RAM and 4 CPU cores. The Ryu controller is used as the SDN controller in the testbed for managing compliant switches. It operated with Layer 4 (L4) learning capabilities to forward flow traffic based on the matching of the MAC, IP, protocol, and service port. Scikit-learn, an open-source library used in our work to implement machine learning. On the host side, to deceive the attacker, we used Honeyed as a low-interaction honeypot. During the experiment, we generated probe traffic using the NMAP tool, and for normal traffic generation, the Iperf and D-ITG tools were used. Moreover, we performed some real tasks such as browsing the internet, watching YouTube, using emails, downloading and uploading files, and SSHing different mininet hosts to generate realistic traffic.

4. Findings and Discussion

The performance of the suggested mechanism is evaluated using the evaluation metrics covered in depth in this section, which is then followed by a detailed presentation of the findings and outcomes. For evaluation, seven popular ML models were adopted to determine the best model, including DT, RF, NB, AdaBoost, LightGBM, and KNN classifiers. By employing this diverse set of models, we intended to select the one that best fits the characteristics of our data and the requirements of the mechanism. The equations in Table II. show the metrics used to evaluate the proposed mechanism in terms of accuracy, recall, precision, and F-measure, respectively.

Where TP, FN, TN, and FP stand for true positives, false negatives, and false positives, respectively. As depicted in [4], the majority of scholar articles used other metrics such as F1-score, precision, and recall because the average model's performance tends to favor the majority classes' performance, it is impossible to fairly assess classification performance under the imbalanced dataset using only the accuracy rate. A confusion matrix was utilized for each classifier in order to precisely assess the suggested mechanism and compute these performance indicators; the outcomes are displayed in Table III. Among the ML models used, their results were very good, except for NB, since they are commonly used for text classification tasks. The XGBoost model achieves the best result in terms of accuracy 94.73% as well as precision, recall, and f1-score.

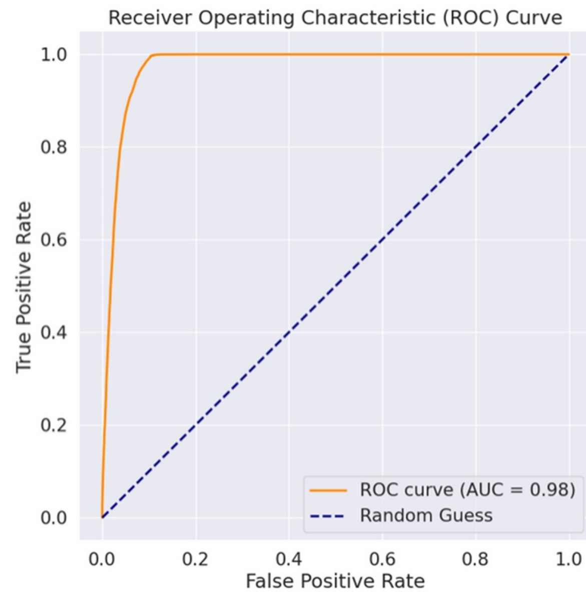

Figure 5. XGBoost AUC ROC Curve.

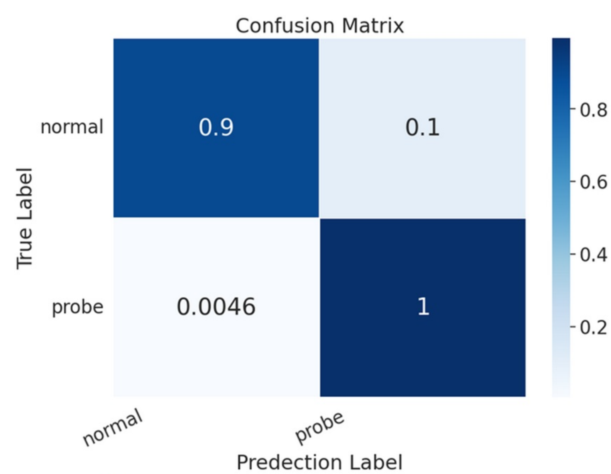

Figure 6. XGBoost Confusion Matrix.

Moreover, the author in [15] emphasizes the significance of the Area Under Curve (AUC) metric for the categorization of unbalanced datasets. The AUC of the XGBoost classifier obtained good results of about 0.98%, as seen in Figure 5. The confusion matrix of the XGBoost classifier in

TABLE I. Features Used in Proposed Mechanism.

Feature	Description
flow_duration (nsec)	The duration of flow remaining in the switch flow table in nanoseconds
ip_proto	The flow protocol
srcport	Source port
dstport	Destination port
byte_count	Total flow bytes
packet_count	Total flow packets
byte/s	Number of flow bytes per second
pkt/s	Number of flow packets per second

TABLE II. Evaluation Metrics Description

Accuracy	Precision	Recall	F1 Score
Percentage of true classified flow in relation to all classified flows (TP+TN)/(TP+TN+FP+FN)	Percentage of malicious flows identified correctly in relation to all flows classified as attacks TP/(TP+FP)	Percentage of correctly classified malicious flows versus all malicious flows presented in the dataset TP/(TP+FN)	The harmonic mean between precision and recall (2*Precison*Recall)/ (Precision+Recall)

TABLE III. Precision, Recall and F1 Score of Models Evaluation

Classifier	Accuracy	Precision	Recall	F1-score
DT	0.9390	0.9449	0.9378	0.9397
RF	0.9400	0.9457	0.9392	0.9410
NB	0.8449	0.8830	0.8356	0.8374
AdaBoost	0.9196	0.9260	0.9203	0.9218
XGBoost	0.9473	0.9544	0.9458	0.9480
LightGBM	0.9459	0.9532	0.9445	0.9467
KNN	0.9128	0.9170	0.9126	0.9139

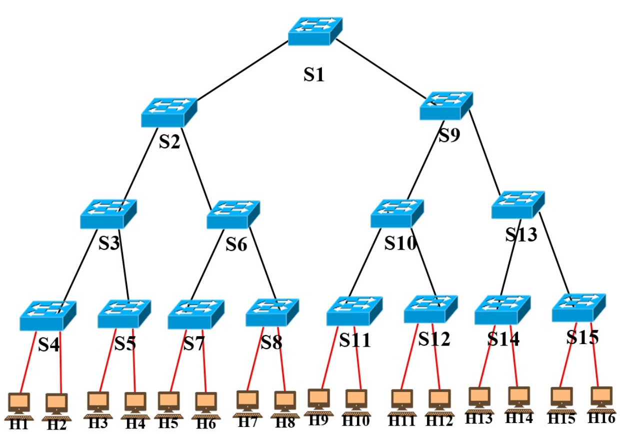

Figure 7. Tree Topology (Depth 4, Fanout 2).

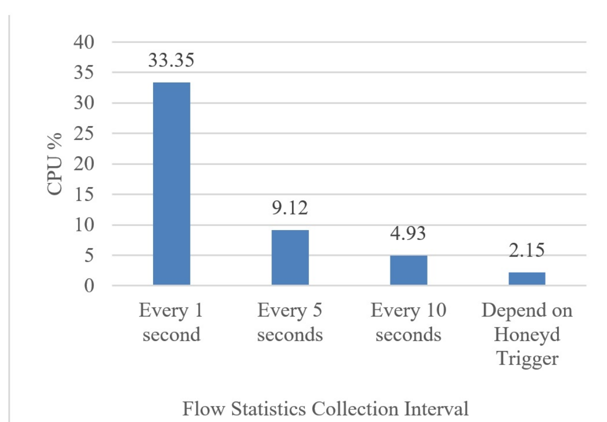

Figure 8. Ryu Controller CPU Utilization.

Figure 6 shows that our detection mechanism successfully recognizes probe attacks, but the ratio of normal traffic recognized as probe attacks (false positive) needs more consideration.

We conducted another experiment with a larger topology of tree with depth 4 and fanout 2, as shown in Figure 7, to support the hypothesis we previously highlighted in this article, which states that continuous traffic checking through statistics requests and reply messages by detection applications will lead to an increase in controller workload. This topology aims to increase the number of switches and hosts in the network, which will increase the number of flow rules in each switch. In this stage, normal traffic originated from each host to another, creating huge traffic.

Four separate scenarios were employed in order to assess the controller CPU load, as seen in Figure 8. Several traffic checking intervals, such as 1, 5, and 10 seconds, were chosen for the first three scenarios of statistics request message sending. In the final scenario, honeypot triggering solutions are employed in place of periodic checking when necessary. Only normal traffic was transmitted for all scenarios, and when controllers checked the traffic for each second, controller overhead was quite high 33%. Controller overhead significantly decreased when the interval was extended to every 10 seconds, but attackers in this case will have plenty of time to compromise the network. However, since the detecting mechanism is in sleep mode and would activate anytime there is an alarm, our idea of applying a triggering mechanism successfully lowered CPU demand to 2.1%.

5. Conclusions and Future Work

Due to the growing threat of cyberattacks in SDN environments and the lack of built-in security mechanisms, this model finds it difficult to replace traditional network architecture. One important attack that is neglected by the research community is the probe attack. Probe attacks are considered the first phase of other serious attacks such as DoS, botnets, MitM, brute force, etc. to exploit vulnerabilities in the network. The issue with this attack is that it is hard to detect since it works passively without being noticed. Machine learning and deep learning are widely used in current works for detection techniques because of their ability to detect novel attacks, contrary to classical methods of using independent agents in networks, which depend on fixed thresholds or attack signatures. In this work, we propose a novel online lightweight detection mechanism empowered by machine learning models. Our approach leverages the dynamic nature of SDN to embed honeypots within the network infrastructure. These honeypots on the network host serve a dual purpose: they lure potential attackers by presenting fake services, and they act as triggers for our detection mechanism in the SDN controller when any traffic contacts them. Machine learning models need a good dataset to provide accurate detection in real-time. Therefore, we collected a new dataset specific to probe attacks in SDN environments. Contrary to other datasets, which are either outdated, incompatible with SDN, or whose features cannot be easily obtained, our dataset features were collected through OpenFlow statistics messages that can be easily collected in an SDN environment. Through extensive experimentation, we have demonstrated the efficacy of our proposed system. Notably, our detection mechanism achieves an exceptional level of accuracy, while imposing minimal stress on the SDN controller's CPU. These results underscore the practical viability and effectiveness of our approach on strengthening SDN security against evolving cyber threats.

However, the effectiveness of the mechanism is contingent on attackers engaging with honeypots, a strategy that may not always be reliable against more sophisticated threats. Additionally, the attacker might be able to identify the honeypot and avoid it. Periodic traffic monitoring may be able to resolve this problem, but as this article has previously shown, it also increases controller CPU demand and generates network overhead. The future work of this research is to develop a method to prevent honeypot fingerprinting and continuously examine traffic without adding controller overhead.

References

[1] A. Alshamrani, “Reconnaissance attack in sdn based environments,” in *2020 27th International Conference on Telecommunications (ICT)*. IEEE, 2020, pp. 1–5.

[2] M. S. El Sayed, N.-A. Le-Khac, M. A. Azer, and A. D. Jurcut, “A flow-based anomaly detection approach with feature selection method against ddos attacks in sdns,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 8, no. 4, pp. 1862–1880, 2022.

[3] A. S. Alshra'a, A. Farhat, and J. Seitz, “Deep learning algorithms for detecting denial of service attacks in software-defined networks,” *Procedia Computer Science*, vol. 191, pp. 254–263, 2021.

[4] H. Y. I. Khalid and N. B. I. Aldabagh, “A survey on the latest intrusion detection datasets for software defined networking environments,” *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, pp. 13 190–13 200, 2024.

[5] M. S. Towhid and N. Shahriar, “Early detection of intrusion in sdn,” in *NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium*. IEEE, 2023, pp. 1–6.

[6] A. Almazyad, L. Halman, and A. Alsaeed, “Probe attack detection using an improved intrusion detection system.” *Computers, Materials & Continua*, vol. 74, no. 3, 2023.

[7] A. Althobiti, R. Almohayawi, and O. Bamsag, “Machine learning approach to secure software defined network: Machine learning and artificial intelligence,” in *Proceedings of the 4th International Conference on Future Networks and Distributed Systems*, 2020, pp. 1–8.

[8] A. Abubakar and B. Pranggono, “Machine learning based intrusion detection system for software defined networks,” in *2017 seventh international conference on emerging security technologies (EST)*. IEEE, 2017, pp. 138–143.

[9] B. Karan, D. Narayan, and P. Hiremath, “Detection of ddos attacks in software defined networks,” in *2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS)*. IEEE, 2018, pp. 265–270.

[10] M. W. Nadeem, H. G. Goh, V. Ponnusamy, and Y. Aun, “Ddos detection in sdn using machine learning techniques.” *Computers, Materials & Continua*, vol. 71, no. 1, 2022.

[11] J. Wang and L. Wang, “Sdn-defend: A lightweight online attack detection and mitigation system for ddos attacks in sdn,” *Sensors*, vol. 22, no. 21, p. 8287, 2022.

[12] V. Hnamte and J. Hussain, “An efficient ddos attack detection mechanism in sdn environment,” *International Journal of Information Technology*, vol. 15, no. 5, pp. 2623–2636, 2023.

[13] S. Wang, J. F. Balarezo, K. G. Chavez, A. Al-Hourani, S. Kandeepan, M. R. Asghar, and G. Russello, “Detecting flooding ddos attacks in software defined networks using supervised learning techniques,” *Engineering science and technology, an international journal*, vol. 35, p. 101176, 2022.

[14] H. Y. Ibrahim, P. M. Ismael, A. A. Albabawat, and A. B. Al-Khalil, “A secure mechanism to prevent arp spoofing and arp broadcasting in sdn,” in *2020 international conference on computer science and software engineering (CSASE)*. IEEE, 2020, pp. 13–19.

[15] Q.-V. Dang, “Intrusion detection in software-defined networks,” in *Future Data and Security Engineering: 8th International Conference, FDSE 2021, Virtual Event, November 24–26, 2021, Proceedings 8*. Springer, 2021, pp. 356–371.

[16] A. Mzibri, R. Benaini, and M. B. Mamoun, “Case study on the performance of ml-based network intrusion detection systems in sdn,” in *International Conference on Networked Systems*. Springer, 2023, pp. 90–95.

[17] N. Mazhar, R. Salleh, M. A. Hossain, and M. Zeeshan, “Sdn based intrusion detection and prevention systems using manufacturer usage description: A survey,” *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 12, 2020.

[18] M. Said Elsayed, N.-A. Le-Khac, S. Dev, and A. D. Jurcut, “Network anomaly detection using lstm based autoencoder,” in *Proceedings of the 16th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, 2020, pp. 37–45.

[19] M. S. ElSayed, N.-A. Le-Khac, M. A. Albahar, and A. Jurcut, “A novel hybrid model for intrusion detection systems in sdns based on cnn and a new regularization technique,” *Journal of Network and Computer Applications*, vol. 191, p. 103160, 2021.

[20] H.-M. Chuang, F. Liu, and C.-H. Tsai, “Early detection of abnormal attacks in software-defined networking using machine learning approaches,” *Symmetry*, vol. 14, no. 6, p. 1178, 2022.

[21] E. M. Zeleke, H. M. Melaku, and F. G. Mengistu, “Efficient intrusion detection system for sdn orchestrated internet of things,” *Journal of Computer Networks and Communications*, vol. 2021, pp. 1–14, 2021.

[22] M. S. Elsayed, N.-A. Le-Khac, and A. D. Jurcut, “Insdn: A novel sdn intrusion dataset,” *IEEE access*, vol. 8, pp. 165 263–165 284, 2020.

[23] N. Abbas, Y. Nasser, M. Shehab, and S. Sharafeddine, “Attack-specific feature selection for anomaly detection in software-defined networks,” in *2021 3rd IEEE middle east and north africa communications conference (menacomm)*. IEEE, 2021, pp. 142–146.

[24] D. Stiawan, M. Y. B. Idris, A. M. Bamhdi, R. Budiarto *et al.*, “Cicids-2017 dataset feature analysis with information gain for anomaly detection,” *IEEE Access*, vol. 8, pp. 132 911–132 921, 2020.

[25] D. Kreutz, F. M. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig, “Software-defined networking: A comprehensive survey,” *Proceedings of the IEEE*, vol. 103, no. 1, pp. 14–76, 2014.

[26] H. Y. KHALID, P. M. ISMAEL, and A. B. AL-KHALIL, “Efficient mechanism for securing software defined network against arp spoofing attack,” *Journal of Duhok University*, vol. 22, no. 1, pp. 124–131, 2019.

[27] T. A. Tang, D. McLernon, L. Mhamdi, S. A. R. Zaidi, and M. Ghogho, “Intrusion detection in sdn-based networks: Deep recurrent neural network approach,” *Deep Learning Applications for Cyber Security*, pp. 175–195, 2019.

[28] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, “Deep learning approach for network intrusion detection in software defined networking,” in *2016 international conference on wireless networks and mobile communications (WINCOM)*. IEEE, 2016, pp. 258–263.

Harman Yousif Ibrahim Khalid is currently working as a Lecturer in University of Duhok, College of Science, Department of Computer Science, Duhok, IRAQ. he received the master’s degree in Computer Science Data in 2019 from Duhok University, Duhok, IRAQ and he is currently a Ph.D. student at Department of Computer Science, College of Computer Science and Mathematics, University of Mosul, Mosul, IRAQ. His research domains include Cyber Security and authored many publications in international journals and conferences.

Najla Badie Ibraheem Assistant prof. in Department of Computer Science, Computer Science and Mathematics College, University of Mosul, Mosul, Iraq.