Detection of Distributed Denial of Service Attacks in SDN using Machine learning techniques

K.Muthamil Sudar
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar, India
k.muthamilsudar@klu.ac.in

M. Beulah
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar, India
beulahlakshmi@gmail.com

P.Deepalakshmi
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar, India
deepa.kumar@klu.ac.in

P.Nagaraj
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Virudhunagar, India
nagaraj.p@klu.ac.in

P. Chinnasamy
Department of Information Technology
Sri Shakthi Institute of Engineering and Technology
chinnasamyponnusamy@gmail.com

2021 International Conference on Computer Communication and Informatics (ICCCI) | 978-1-7281-5875-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCCI50826.2021.9402517

Abstract- **Software-defined network (SDN) is a network architecture that used to build, design the hardware components virtually. We can dynamically change the settings of network connections. In the traditional network, it's not possible to change dynamically, because it's a fixed connection. SDN is a good approach but still is vulnerable to DDoS attacks. The DDoS attack is menacing to the internet. To prevent the DDoS attack, the machine learning algorithm can be used. The DDoS attack is the multiple collaborated systems that are used to target the particular server at the same time. In SDN control layer is in the center that link with the application and infrastructure layer, where the devices in the infrastructure layer controlled by the software. In this paper, we propose a machine learning technique namely Decision Tree and Support Vector Machine (SVM) to detect malicious traffic. Our test outcome shows that the Decision Tree and Support Vector Machine (SVM) algorithm provides better accuracy and detection rate.**

Keywords- Security, Distributed Denial of Service (DDoS), Machine Learning, Software-defined network (SDN), Support Vector Machine (SVM), Decision Tree.

I. INTRODUCTION

Software Defined Networking is an emerging paradigm which overcomes the limitations of conventional network architecture by separating the control from data plane devices. SDN consists of three planes such as data plane, control plane and application plane. Data plane carries the network traffic based on the decision made by controller. Control plane decides the flow of traffic by computing the routing tables. Application plane manages the other applications like load balancer, firewalls, Quality of Service (QoS) applications etc. SDN architecture improves the network performance by decoupling the network control and forward function. The control programs running in a logically centralized controller will control multiple routers across the network.

The SDN provides the sole ability to the applications to get to know the entire network information. During high traffic, the integration of different applications helps for load balancing and intrusion detection. If an anomaly is detected, the controller is instructed by the application to reprogram the data plane to alleviate it. Both control and data plane runs on routers that are distributed across the network, where the devices have open interfaces that can be controlled by the software.

In SDN architecture, it is possible to reconfigure the multiple devices at the same time. The application layer is used to configure network devices. The control layer (control plane) which consists of the same controller it is the brain of the SDN architecture. These two layers are communicated through API. The infrastructure layer (data

978-1-7281-5875-4/21/$31.00 ©2021 IEEE

plane) that communicates between the controller and the network devices use a central protocol. Figure 1 explains about the SDN architecture.

Since huge amount of traffic is passing through the controller, proper security mechanism is essential to analyze and identify suspicious traffic. We propose machine learning-based mechanism to identify the malicious activities in the SDN by investigating the traffic features.

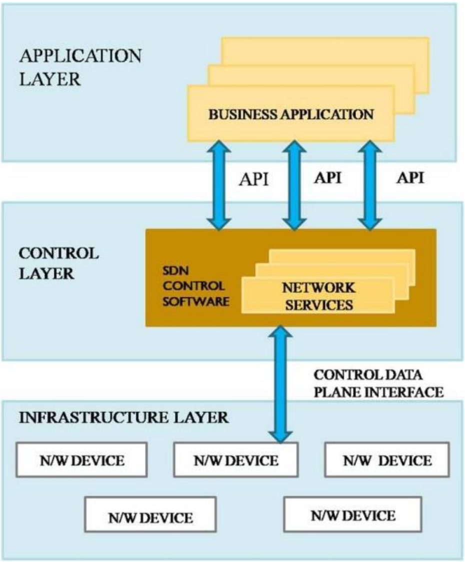

Fig1. SDN Architecture

II. RELATED WORK

In this section, we analyze the various research works done in the field of SDN to detect the DDoS attacks. In [1], authors constructed a network topology that establishes the DDoS attack in a host (including 1 server and 10 clients for network connection). They used an improved KNN algorithm to dig out the DDoS attack. These algorithms give out a good prediction rate of 0.912 respectively. In [2], the authors explained about the DDoS attack in cloud computing environments. They claimed that the datasets with optimal feature selection in the current DDoS attack will strengthen the accuracy of detection of DDoS attack in the cloud computing. In [3], the authors proposed semi-supervised k-means – hybrid feature selection (SKM-HFS) used to detect the malicious attack. They claimed that the proposed algorithm gives a good performance level of 80%. In [4], the authors proposed navie bayes, neural networks and SVM algorithm to identify the DDoS attack in the SDN. They stated that naïve bayes, SVM and neural network provide an accuracy of 70%, 80%, and 80% respectively. In [5], authors applied the SVM algorithm on the SDN environment to detect the DDoS attacks. The experiment result shows the accuracy of 0.998. In [6], authors developed an integrated frameworks to identify DDoS attacks such as analytical and machine-learning-based attacks. They also claimed that the methodology based on machine learning often aims to recognize zero-day assaults. In [7], the study presented hybrid machine learning approach, to recognize the DDoS assualts. They used SVM and Self-organized Map (SOM) to detect the malicious activities in the SDN environment. In [8], the authors proposed ensemble modeling model to identify the DDoS attacks using K-Nearest Neighbors (KNN), SVM, SOM, Naïve Bayes. The authors of [15], used techniques like SVM, MLP, Decision tree and Random forest in SDN to classify DDoS attack. They reported that perhaps the random forest algorithm provides good precision and the decision tree achieves the optimal time to be completed.

By analyzing the various research works, we have identified that there are various techniques to avert the DDoS attack i.e. Random forest, Naive bayes, KNN, Neural Network, SVM, SOM. In the proposed work, Support Vector Machine (SVM) and decision tree algorithms are used to detect the DDoS attacks by analyzing the essential features of traffic.

III. PROPOSED WORK

In this section, we discuss about our proposed work for identifying DDoS attacks using ML in SDN. We have used SVM and Decision tree algorithm to detect the attacks due to its accurate classification and less complexity.

The DDoS attack mainly categorized into three types such as (i) volume-based attack, which is mainly used to use drench the internet pipe of targeted server like UDP floods and ICMP floods (ii) protocol attacks like SYN flood, fragmented packet, ping of death, smurf DDoS which mainly focus on extracting the server resources (iii) application layer attacks include GET/POST floods which focus on web applications and its goal is to crash the web server.

A. MACHINE LEARNING

SVM is a machine learning technique supervised and used for the purpose of classification and regression. The aim is to mark the packets as malicious or natural in this article. To experience a DDoS attack, SVM and decision tree are used. Compared with other machine learning methods, SVM is more stable. The DDoS attack is launched towards multiple hosts and machine learning algorithm helps to identify the malicious attack in the SDN environment. Figure 2 explains about the system diagram.

The dataset is divided into training and testing data to train the model with both SVM and decision tree. The traffic flow data from the flow table entries are collected and analyzed using machine learning module to detect the malicious packet.

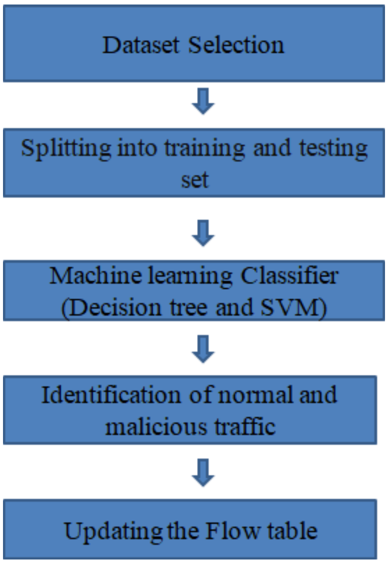
Dataset Selection
Splitting into training and testing set
Machine learning Classifier (Decision tree and SVM)
Identification of normal and malicious traffic
Updating the Flow table

Fig 2. System Diagram

Step 1:
For the development of a model, the linear kernel function $K(x,\ y)$ as well as the relevant parameter C are defined. For mapping input vector into high-dimensional feature space, the linear kernel function $K(x,\ y)\ =\ x^{T}y\ +\ c$ is used. C is the variable of regularisation, which has to be greater than unity. let the C is 1 $\alpha^{*}\ =\ (\alpha_{1}^{*},\alpha_{2}^{*},\ldots..,\alpha_{N}^{*})^{T}$ is the LaGrange multiplier vector.

$$\min\left(\frac{the\ 1}{2}\sum_{i=1}^{N}\alpha_{i}\,\alpha_{j}\,y_{i}\,y_{j}\,K(x_{i},x_{j})-\sum_{i=1}^{N}\alpha_{i}\right) \qquad (1)$$

$$s.t \quad \sum_{i=1}^{N} a_{i}\,y_{i}=0 \qquad (2)$$

$$0\leq\alpha_{i}\leq C,\qquad i=1,2,\ldots N$$

$$\alpha^{*}=(\alpha_{1}^{*},\alpha_{2}^{*},\ldots\ldots,\alpha_{N}^{*})^{T} \qquad (3)$$

Step 2: Set a positive element from $\alpha^{*}(0\leq\alpha_{j}^{*})$ to evaluate b^{*} which is the specification function of the object:

$$b^{*}=y_{j}-\sum_{i=1}^{N}\alpha_{i}^{*}\,y_{i}K\ (x_{i}.\,x_{j}) \qquad (4)$$

Step 3: construction of decision function

$$f(x)=sgin\left(\sum_{i=1}^{N}\alpha_{i}^{*}\,y_{i}K(x.\,x_{i})+b^{*}\right)(5)$$

The decision tree is used for the classification of different traffic and to differentiate between normal and malicious traffic.

B. DDoS ATTACK IDENTIFICATION

Initially the dataset is divided into training and test dataset. In the feature extraction phase, essential features are identified and selected for further detection process. In next phase, the dataset is passed through the SVM classifier and Decision tree module. Classifiers output the traffic dataset into two classes either attack or normal based on the flag value (0 or 1). In case of attack instance (flag=1), it alerts the controllers to drop the particular flow from the flow table. Otherwise, controller will formulate the routing path for the normal traffic packets.

The controller will send the forwarding table to process such payload whenever the DDoS issue is detected through using SVM classifier and decision tree. The SVM has high robust that it uses a kernel trick technique to gives out the possible outputs.

IV. EXPERIMENT SETUP

To simulate the SDN environment, we used the standard simulation tool called Mininet [9, 11-14]. It is used to establish the SDN network topology connection. By using mininet, we can add switches, hosts, and controller virtually and we can change or modify the network connection.

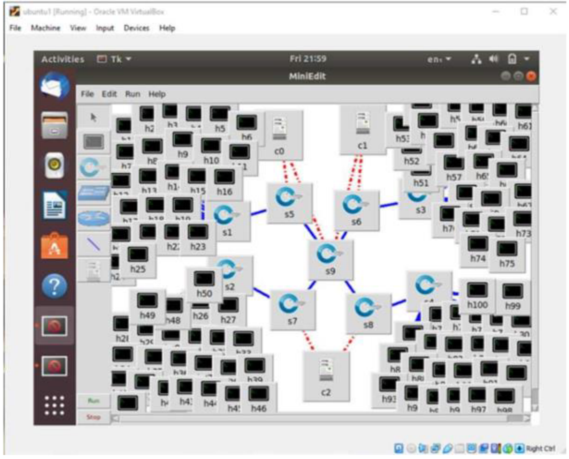

Fig. 4 Network topology for Proposed work