DDoSNET: AN EFFICIENT DDoS ATTACK DETECTION AND MITIGATION MODEL IN THE SOFTWARE DEFINED NETWORKING

[1]Chandini C S, [1]Sumi Naushad, [1]Vaidehi S Jayan, [1]Vishnupriya C, [2]Shamin S
[1]UG Scholar, Department of Computer Science and Engineering,
[2]Asst. Prof, Department of Computer Science and Engineering,
UKF College Of Engineering And Technology, Parippally, Kollam, Kerala.

Abstract: ***Distributed Denial of Service(DDoS) attack is considered as one of the foremost complicated vulnerability threats in the Cyber Physical System. Detection and mitigation of such attacks in an efficient manner with better accuracy is the challenging need of the hour. In this paper, we proposed a model to detect the real time DDoS attack in the network and to block such assaults from entering the Software Defined Networking. The model displays the details of the system from which malicious packets have been sent and hence we can block that IP address from sending further packets to the destination host. So that the normal traffic and other network devices are not get affected and continues their service. And also it includes a provision to detect various categories of attacks in a log file. On passing a log file to the model it detects the types of attack as well as the count of the various detected attacks. The evaluation is done by using KDD 99 dataset, which is widely used for the detection of DDoS attack, using Naïve Bayes and Decision tree classifier. The performance of the model is found to be 98% accurate with less space and time complexity. This model is regarded as the par excellence of computer and network security fields.***

Keywords: Distributed Denial of Service, Software Defined Networking, Cyber Physical System.

I. INTRODUCTION

Distributed Denial of service(DDoS) over the Software Defined Network has become the most conspicuous attack in the computing world over the last decade. Detection and the prevention of DDoS have turned to be demanding and more challenging. Various successful attempts have been done to halt such attacks to protect the Cyber Physical System. Shang Gao et al. has proposed a model to detect and mitigate DoS attacks in SDN by using a protocol independent defense mechanism called FloodDefender, but the vulnerable structure caused severe network security problems [1]. The idea of Sumit Badotra et al. was an early DDoS detection by a SNORT based mechanism. It makes use of Opendaylight and open networking operating system techniques in software defined networking [2]. That suggested a further improvement in the DDoS framework. Jesus Arturo Perez Diaz et al. introduced a modular architecture for the identification and prevention of LR-DDoS attacks in SDN [3]. It found difficult to figure out analyze the characteristics of LRDDOS attack. Kshira Sagar Sahoo et al. used the idea of Support Vector Machine model for DDoS attack detection in SDN [4]. The request processing undergoes a delay due to the low memory space of the controller and hence it requires more training time too. Zakaria Abou El Houda et al. brought intelligence to SDN by mitigating the DDoS attacks by using a machine learning based DDoS detection module, flow statistics collection scheme and entropy calculation scheme [5]. This faced a problem of network flooding with illegitimate requests. By utilizing machine learning KNN approach Nguyen Ngoc Tuan et al. proposed a technique which was a novel approach in DDoS attack mitigation. They concentrate in SDN-based Internet Service Provider (ISP) networks basically for two types of attacks, TCP-SYN and ICMP flood attacks. When the attack intrudes the cloud its mitigation is found to be more complex. Collaborative detection and mitigation of distributed denial-of-service attacks on SoftwareDefined Network by Omer Elsier Tayfour et al. tried to reduce the controller overhead by making use of collaborative information sharing mechanism [6]. They recommended an improvement in the machine learning integration techniques.

N Satheesh, M V Rathnamma, G Rajeshkumar, P Vidyasagar et al. put forward a flow based anomaly intrusion detection using machine learning with SDN for open flow network [7]. Here the anomaly intruder is detected using the machine learning decision tree classifier by continuously monitoring the behavior of the network. It is a lightweight traffic anomaly intruder detection approach and is successful in finding the anomaly induced by high-efficiency network assaults. But they trained the model with very less traffic info and hence the speed and accuracy of the model has not met the mark. The feature selection was also found to be much more time consuming as well as intricate. Ramin Fadaei Fouladi et al come up with a time series analysis for the DDoS attack detection and Defense scheme for SDN [8] . The anomalies caused by the DDoS attacks are pointed out by monitoring each Open Flow individually but the model was found out to be more time complex. Oussama Hannache et al. implemented a Neural Network-Based approach for detection and mitigation of DDoS attacks in SDN environments by using artificial neural network and back propagation algorithm [9] . The samples of normal traffic and DDoS attack generated malicious traffic are collected and trained the dataset using a Traffic Flow Classifier based Neural Network and presented a mitigation process by using TFC- NN live classification. They suggested to deploy it in a Network Function Virtualization as the further enhancement. Most of the works found difficulties in the computational complexities of the model including the time and space complexity. Performance of the model using various classification algorithms and evaluation using different labelled datasets were challenging for most of them. It was hard while testing the model with more variants of parameter types and values. Some of the works failed in the real time implementation of the model and some referred it as

a future addition. It is also found that models going wrong in classifying various kinds of attacks. The necessity of large scale labelled data to train the classifier was also not easy to get. In some cases finding the optimal values of hyper parameters related to ML techniques requires further elaborated study. The application of appropriate and most accuracy providing algorithm also found to be more complex.

We worked on reducing the overall complexities of the model by choosing one of the best machine learning algorithms which causes comparatively very less delay in producing the result. We dealt with producing an appreciable evaluation metrics and decided upon to implement the model in the real time environment. The aim of our model is to:

- Develop an accurate method to detect and block the compromised (attacker) hosts.
- Implement an efficient model in terms of time and space complexity with limited hardware requirements and well organized software requirements.
- Generate a model which could not raise any burden on the network devices as well as the legitimate users.
- Demonstrate the model in real time and prevent SDN from attack vulnerabilities.

The rest of this paper is structured as follows: section 2 detailing the materials and methods, section 3 describing the methodology of the proposed model, followed by the results and discussion in section 4 and ending with the conclusion and future work in section 5.

II. MATERIALS AND METHODS

II.I KDD CUP DATASET

KDD cup dataset has been the foremost widely used dataset for anomaly detection method, training dataset accommodates about four lakh single connection vectors within which contain 41 features which is labelled as normal or traffic. Although KDD CUP dataset has quite 15 years old, it is still being widely employed in many academic research like Machine learning Research(MLR) and Intrusion Detection Systems(IDS). The KDD data set could also be a well known technique within the world of Intrusion Detection techniques. Many labor goes on for the event of intrusion detection strategies while the research on the data used for training and testing the detection model is that the prime concern, because better data quality could improve offline intrusion detection. The KDD Cup99 data set was prepared on processing the TCPdump segment of 1998 supported DARPA Intrusion Detection System evaluation dataset by Stolfo et al. Of Lincoln Labs, U.S.A. DARPA98 is about 4 GB of compressed raw TCP-dump data of seven weeks of network traffic. It might be processed into about 5 million connection records, each with about 100 bytes. KDD cup99 is the dataset which was used for Third International Knowledge Discovery and processing Tools contest, held in concurrence with KDD Cup99, yet as within the Fifth International Conference on Knowledge Discovery and processing. Since 1999, KDD Cup99 has been the foremost wildly used data set for the evaluation of anomaly detection methods. The knowledge set consist an entire of 24 attack types(connections) that comprise one in every of the 4 major categories: Denial of service(DOS), Probe/scan, User to Root(U2R) and Remote to User(R2L). Most of the detection system is evaluated using the KDD cup data set for both normalized and un-normalized data.

II.II DDOS ATTACK DETECTION USING NAÏVE BAYES

Intrusion detection is that the strategy of monitoring and analysing the events occurring within the system, Dr. Saurabh Mukherjee[10] et.al has introduced in his paper about intrusion detection using naïve bayes. The data mining algorithm, naive bayes classifier are evaluated on the NSL KDD dataset to detect the attacks mainly on four attack categories which incorporates Probe (information gathering). Feature selection and reduction are the effective and are the vital steps. The foremost important includes feature reduction, and is applied using three standard feature selection methods which are Correlation based Feature Selection (CFS), Information Gain (IG), Gain Ratio (GR). The naive bayes classifier's results are computed, and is used for the comparison of feature reduction methods to point that the proposed model is best and more effective for network intrusion detection. In FVBRM, one input feature is deleted from the dataset at a time (feature reduction) and thus the resultant dataset is employed for the training and testing of the classifier, this process continues until it performs better than the initial dataset in terms of relevant performance criteria, stated as Feature Vitality Based Reduction Method(FVBRM). The output of the metrics are visiting be that the FVBRM performs far better than other feature reduction methods like CFS, IG and GR. The feature reduction is performed supported 41 features to induce 10 using CFS, 14 using GR, 20 using IG and 24 using FVBRM on NSLKDD dataset. By using well known performance metrics the empirical results are compared for different feature reduction methods. Here [10] have explained the algorithm for FVBRM. In this, initially apply naïve bayes classifier on dataset with 41 features. Then compute its performance output like classifier's accuracy, RMSE, average TPR value and set F is input to the algorithm.

II.III DDOS ATTACK DETECTION USING DECISION TREE

In [11][12][13] DDOS attack is detected using decision tree algorithm. There have been several reason behind this, a number of them includes: Decision tree algorithm is extremely effective in data and text mining, information extraction. This idea is incredibly important because it enables modelling and extraction of information from set of information available. Decision trees are the foremost important techniques in information analysis, that adopts the knowledge through examining complex similarly as high mountain information designs. In [11] first of all the information is collected from KDDCup'99 intrusion detection and evaluation dataset. Then Select small portion of coaching and testing data from KDDCup'99 Dataset for the experiment. Then data pre-processing is frenzied, J48 classifier is employed for detection and categorization of data and by using J48 training algorithm for data training. After preparation and testing, the connections are classified into 15

classes. The test and train option is employed where the test option is tested for both before and after the attribute selection and a weighted average is calculated. From these analyses J48 algorithm provides most accurate and valuable solution compared with random forest decision tree algorithm.

In[12] DDoS attack detection and attack path construction are handled within the protection agent which is that the center for the whole system. A protection agent consists of mainly four components such are: - A packet aggregator (to aggregate the traffic signatures), a message manager (to construct the SSH tunnel and handle communication between the protection agent and sentinels), DDoS attack detection module and a trace back module.

In[13] proposes a Flow based mechanism to obtain information from the network, and uses this information to classify the information theory, then changes flow rules to block malicious flows. There were Several factors that affect accuracy in the detection of SDN based intrusion identification methods, which includes applicable data set for the volume of data in the dataset, feature selection, with proper ML model superior, learning rate, cross-validation, training time, and the list goes on. Many kinds of literature lack in the identification of all attack groups within SDN. However, they used both the techniques like ML and DL approaches for building the network for intrusion identification methods for all kinds of attack categories and calculate the performance by supporting through the dataset for measuring the comparative weaknesses and strengths in both the ways. From the experimental setup decision tree has the highest detection rate of 95.16 %, and the false alarm rate is 2.49 %, but it also has the lowest test time of 0.181ms. The decision tree is the right one to identify suspicious network traffic flow. From the above analysis we are clear that naïve bayes and decision tree classifiers are the best algorithms in detecting and mitigating DDoS attacks.

II.IV MOTIVATION OF NEW MODEL

Our aim is to develop a new model which could be able to detect the DDoS attacks on the network and mitigate it with less time and space complexity, and without causing any burden on the other network devices as well as the legitimate users. To solve the difficulties faced by the models proposed earlier which is described in the introduction section of this paper is the main motivation behind this work. Since most of the previous papers suggested real time implementation of the model as the future enhancement, we worked on that too.

III. PROPOSED SYSTEM

III.I EXPERIMENTAL SETUP

In our work, we use Python 3.8.6, cross platform Kivy, Visual Studio Code and Python Django framework to perform the experiments as it provides a better option to carry out the works in an easy and strategic way. In order to develop the model Python programming language with various support libraries such as sklearn, numpy, pandas have been used. The experiments are implemented in an environment of Intel i3 2.2 GHz, 6 GB RAM, 64 bit Windows 10 operating system.

III.II WORKFLOW

At the first stage we designed a user interface by using Python language, specifically Kivy, since the language has simple syntax and supports cross platform and is easy to use. We had imported various gadgets for the material designing, screen management, graphics texture, float layout etc. The user interface includes various modules as follows:

- In order to allow the user to type the username and password to login we go for login module.
- A register module to store the outputs of the commands and included XAMPP server to store the registration details.
- A homepage module consists of a file chooser to choose a particular file which should be a .gz file that indicates who all have already accessed the website.
- An attack detection module to monitor the traffic. It analyses the traffic and matches it with the library of known attacks.

Here we go for two kinds of attack detection methodologies.

III.II.I BY USING LOG FILE

In the home page module, we provided a button called choose file. On clicking that we could choose the downloaded log file. We have trained the KDD cup dataset by using commands. It's a train.py file. We run the dataset then it gets converted to model.py file. Whenever we have chosen the log file within the home page module this model.py file get compares with the log file and detects the type of attack already occurred in that file and the number of times the attack already happened.

III.II.II REAL TIME DDoS ATTACK DETECTION AND MITIGATION

First we have imported sklearn module in the code which is an environment including various well defined algorithm and function. We used naïve bayes algorithm which has high accuracy and speed on large datasets. The classifier trains the model on the given dataset during the learning phase and in the evaluation phase, it tests the classifier performance. We have trained the dataset by using commands. It's a train.py file. We run the dataset then it gets converted to model.py file. During testing this file get compared with the real time file and analyze the results. The inputs to the algorithm are the details of the log file including IP address, browser details, attack categorization etc. We have generated a real time DDoS attack by using High Orbit Ion Cannon. Selected count and IP address as the features to train the model. Here we provided 50 as the threshold value. If the number of requests generated is greater than 50 it is considered as a DDoS attack. The created data packets are send to the PythonAnywhere Django project hosting server. If more than 50 requests are trying to access the server at the same time

the model.py get compared with the file within the PythonAnywhere server and is detected as an attack and blocks the same. Such requests are send to the json file and get backlisted. Hence the DDoS attack got blocked from accessing the site. In the website it is indicated that service has denied and the attacker couldn't access the specific server. In the software it shows the IP address from which these much of requests came and the denial of the access. Hence the mitigation is done. On the same time the legitimate users can access the site without causing any delay.

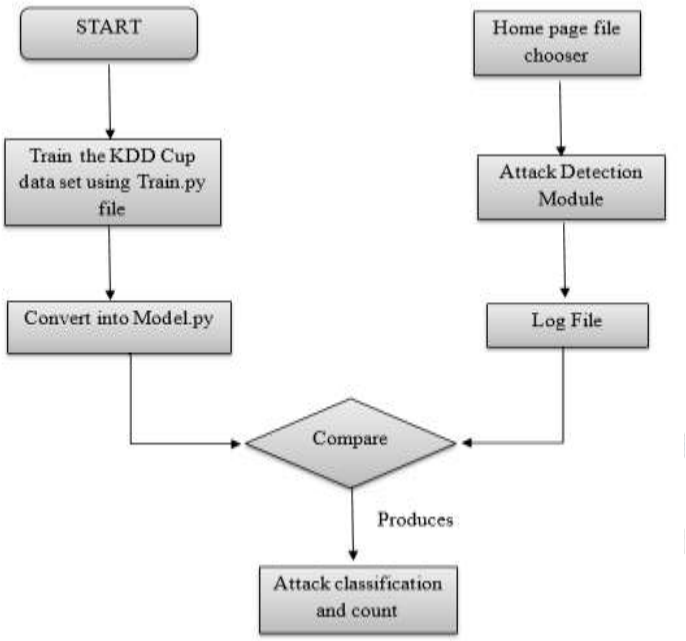

Fig 1. Workflow of Attack Detection in Log File

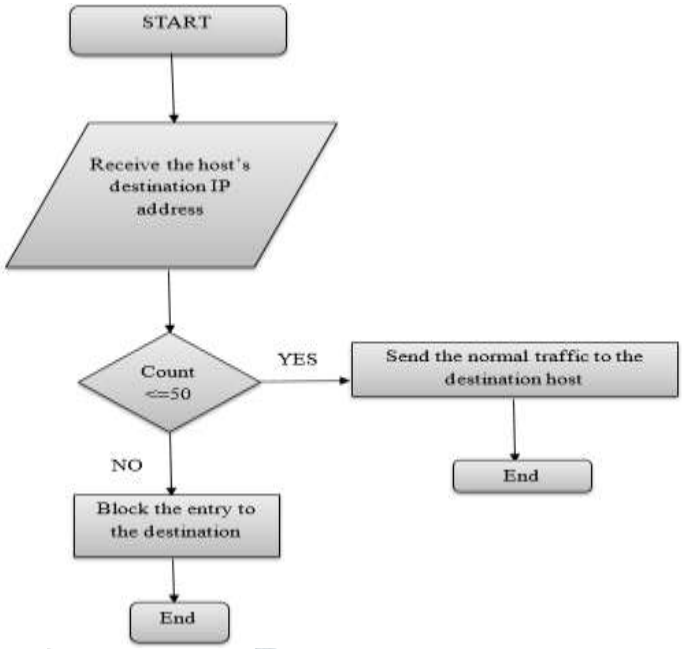

Fig 2. Workflow of Real Time DDoS Attack Detection and Mitigation

III.III EVALUATION METRICS

The validation scores are based on:

1. Accuracy

It is the fraction of the number of correct predictions among the total number of predictions.

Accuracy: (TN+TP)/(TN+TP+FP+FN) (1)

2. Precision

It tells us how many, out of all times that have been predicted to belong to elegance X, truly belonged to magnificence X.

Precision: (TP)/(TP+FP) (2)

3. Recall

It expresses how many times of class X have been anticipated correctly.

Recall: (TP)/(TP+FN)

(3)

IV. RESULTS AND DISCUSSION

Classifier	DT	NB
Accuracy(%)	98	97.02
Sensitivity(%)	94.36	98.21
Specificity(%)	6.5	1.61
AUC(%)	98.89	91

Table 1. Performance Metrics

In our experiments, we relied upon two machine learning algorithms namely naïve bayes and decision tree classifier. In order to find out the accurate results, we go for confusion matrix which compares the performance of the used algorithms. This tables gives the possible outcomes for given classifications. In our results, '1' denotes that an attack is detected and '0' indicates legitimate users. Upon analyzing it we can identify that the combination of both the algorithms provides a better accuracy for the proposed system and hence reducing the complexity of the model.

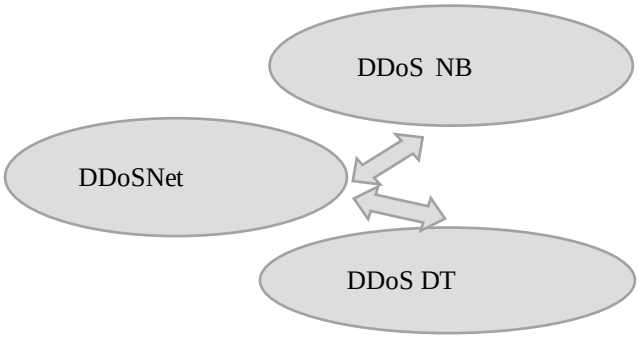

Fig 3 Combined Result

V. CONCLUSION

In this paper, we proposed a model to detect and mitigate real time DDoS attack in the network. Our aim is to develop an accurate model with less complexities and without causing any burden on the network devices. The evaluation is done by using KDD 99 dataset, which is widely used for the detection of DDoS attack. We implemented the model in a real time environment by generating the DDoS attack by using HOIC and blocking the generated requests from entering the target. In the targetor site it is notified to the attacker that access is denied and in the software we get the details of the IP address from which these much of attacks generated and blocks the same. And also proposed model includes a method to detect the attacks in a log file. It shows the various classification of attack as well as the number of times it occurred. It has been shown that the proposed model gives a very high accuracy of 98% in the detection of attacks and hence the model is highly significant in computer and network security fields. In the future, it can be worked on identifying the various classification of attacks in the real time environment.

REFERENCES

[1] "Detection and mitigation of DDoS attack in Software Defined Network" by Shang Gao, Zhe Peng, Bin Xiao, Senior Member, IEEE, Member, ACM, Aiqun Hu, Yubo Song, and Kui Ren, Fellow, IEEE, Member, ACM.

[2] "SNORT based early DDoS detection using opendaylight and opennetworking on SDN" by Sumit Badotra, Surya Narayan Panda.

[3] "A flexible SDN based architecture for identifying and mitigating low rate DDoS attack using ML" by Jesus Arturo Perez- Diaz, Ismael Amezcua Valdovinoz, Kim-Kwang Raymond Choo.

[4] "An evolutionary SVM model for DDoS detection in SDN" by Kshira Sagar Sahoo, Bata Krishna Tripathy, Kshirasagar Naik, Somula Ramasubbareddy.

[5] "Bringing intelligence to SDN: mitigating DDoS attacks" by Zakaria Abou El Houda, Lyes Khoukhi, and Abdelhakim Senhaji Hafid.

[6] "Collaborative detection and mitigation of DDoS attack on SDN" by Omer Elsier Tayfour1 & Muhammad Nadzir Marsono1.

[7] "Flow based anomaly intrusion detection using machine learning with SDN" by N Satheesh, M V Rathnamma, G Rajeshkumar, P Vidyasagar.

[8] "A DDoS attack detection and defense scheme using time series analysis of SDN" by Ramin Fadaei Fouladi.

[9] "Neural network based approach and detection of DDoS attacks" by Oussama Hannache.

[10] "Intrusion detection using Naïve Bayes with feature reduction" by Dr. Saurabh Mukherjeea, Neelam Sharma.

[11] "Detecting DDoS attack decision tree algorithm" by S. Lakshminarasimman, S. Ruswin.

[12] "DDoS detection and traceback with decision tree and grey relational analysis" by Yi-Chi Wu.

[13] "Journal Pre-proof Flow-based Anomaly Intrusion Detection using Machine Learning Model with Software Defined Networking" by N Satheesh, M V Rathnamma, G Rajeshkumar, P Vidyasagar.

(1)(2)(3) https://towardsdatascience.com/accuracyprecision-recall-or-f1-331fb37c5cb9