Grenze International Journal of Engineering and Technology, July Issue

Detection of Network Intrusions using Machine Learning

Ruthvik V[1], Keshav Mittal[2], Manasa U Hegde[3], K Suhas[4] and Annapurna D[5]
[1-5]PESIT Bangalore South Campus, Bangalore, India

***Abstract*—The aim of this project is to design and develop a software solution intended for enterprises to use to detect and classify intrusions and abnormalities in their network. If any such anomalies are detected, a designated administrator is notified to facilitate appropriate action. Machine learning is used to detect any such anomalies in the enterprise network. Our focus is on detecting and classifying DoS attacks.**

I. INTRODUCTION

An intrusion in a network is some illegal or unsanctioned activity on that network. Intrusions aim to steal network resources and/or usually put the security of networks and their data at risk. To detect network intrusions in a timely manner, organizations need to understand how network intrusions work and accordingly implement appropriate network intrusion, detection, and response systems.

An Intrusion Detection System, or IDS, monitors network traffic for abnormal activity and issues alerts if it discovers any. IDSs used in a network are installed at an optimal point in the network, which allows them to monitor traffic from all devices on it. This passing traffic is compared with known attack signatures, and if any matches are found the network's administrator is alerted.

This project intends to produce a software enterprise solution for the purpose of monitoring network traffic for malicious Denial of Service (DoS) attacks. DoS attacks aim to render machines and/or network resources unavailable for use. They do so by inundating the target resource with requests, thus overloading the system and disrupting its services. Using this application, enterprises will be able to ensure that their networks are secure and protected from attackers with malicious intent. A machine learning model is implemented as part of the application to identify and classify any network attacks and/or anomalies. The algorithm is executed over captured network traffic data and uses pre-selected features and parameters to identify and classify DoS attacks. Upon detection of an anomaly, a designated administrator is notified to facilitate appropriate action.

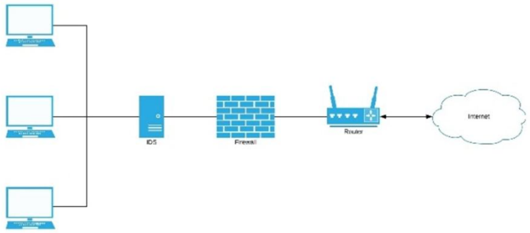

Figure 1. Network Intrusion Detection System

Grenze ID: 01.GIJET.6.2.13

II. Proposed System

The proposed system will monitor an enterprise's internal network for possible DoS attacks. A log will be generated at the end of the day which can be viewed by a designated administrator, summarizing the possible threats detected during the day. Attacks identified as extremely malicious will cause the system to immediately generate an alert.

III. Components

A. Network Tap

The network tap is used to monitor and capture the network traffic in the enterprise network. A tap has at least 3 ports: ports for points A and B in the network, and a monitor port.

B. CICFlowMeter

CICFlowMeter is a network traffic flow generator and analyser. It is used to analyse the enterprise network traffic and extract important statistical network features from the traffic data.

C. Flask

Flask is a web framework written in Python which provides functionality for building web applications, managing HTTP requests and rendering templates. It is used here to develop the front-end interface for the IDS.

D. Classification Model

The classification model used here is a Random Forests Classifier. It classifies the captured network traffic data into 6 categories: Benign, DoS Hulk, DoS Slowloris, DoS GoldenEye, DoS Slowhttptest and Heartbleed.

IV. Implementation

CICFlowmeter, the Flask service, and the network tap are deployed on a Raspberry Pi, which is connected to the enterprise network. From here, the network traffic is captured, and a CSV file is generated which contains the details of each distinct network flow in the network, along with important network statistical features. This CSV file is uploaded to the cloud and passed as input to the Random Forests Classifier, After the model classifies network flows as either benign or malicious, the results are sent to the Raspberry Pi, where they can be accessed using the Flask front end service.

The machine learning model is trained using the CSE-CIC-IDS2018[3] dataset, which was developed by the Communications Security Establishment (CSE) and the Canadian Institute for Cybersecurity (CIC) with the aim of constructing a comprehensive database of network attacks and intrusions. To improve accuracy, the feature selection technique known as bagging was used to eliminate unnecessary features from the CSV file, before it was passed to the machine learning model.

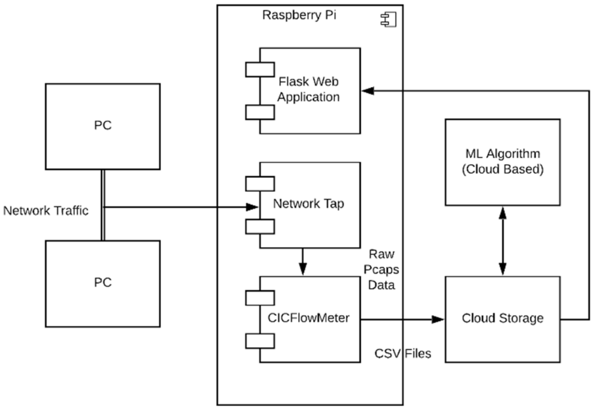

Figure 2. Implemented Design of Network Intrusion Detection System

V. RESULTS

The Random Forests Classifier was trained on the IDS2018 dataset using bagging and a train-test-split methodology. On this dataset, the model achieved an accuracy rate of 98%, indicating that it was able to correctly classify DoS attacks in the dataset with a high degree of reliability. To further cement the model's reliability, a DoS Slowloris attack was simulated in the test network for the duration of 1 hour, and the captured network traffic data was passed on to the model for classification. The captured network traffic data consisted of 3442 distinct network flows, with unique characteristics, but all were instances of a DoS Slowloris attack. The model achieved an accuracy rate of 93% over this simulated DoS attack data, correctly classifying most of the network flows as instances of DoS Slowloris attacks.

TABLE I. CLASSIFICATION RESULTS OF DEVELOPED INTRUSION DETECTION SYSTEM

Benign	**DoS Hulk**	**DoS GoldenEye**	**DoS Slowloris**	**DoS Slowhttptest**	**Heartbleed**
407	34	1	2994	5	0

VI. CONCLUSION

In this paper, we showcased an Intrusion Detection System designed to be used by enterprises to protect their networks from DoS attacks. The intrusion detection system utilizes a Random Forests Classifier to classify individual network flows into one of 6 different categories.

REFERENCES

[1] V. Kanimozhi and T. Prem Jacob, "Artificial Intelligence based Network Intrusion Detection with Hyper-Parameter Optimization Tuning on the Realistic Cyber Dataset CSE-CICIDS2018 using Cloud Computing", *2019 International Conference on Communication and Signal Processing (ICCSP)*, in press.

[2] Francisco Sales de Lima Filho, Frederico A. F. Silveira, Agostinho de Medeiros Brito Junior, Genoveva Vargas-Solar, and Luiz F. Silveira, "Smart Detection: An Online Approach for DoS/DDoS Attack Detection Using Machine Learning", Hindawi, Security and Communication Networks, Volume 2019, Article ID 1574749, in press.

[3] CSE-CIC-IDS2018, https://www.unb.ca/cic/datasets/ids-2018.html; https://registry.opendata.aws/cse-cic-ids2018/