Information Security Management (CSE3502)

Review-1

Topic: Machine Learning approach for DoS Attack Detection

Submitted By:

Gauransh Arora (18BIT0393)

Ayushi Gupta (18BIT0367)

Submitted To:

Prof. Sumaiya Thaseen

VIT®

Vellore Institute of Technology

(Deemed to be University under section 3 of UGC Act, 1956)

I. Idea:

The aim of the project is to develop a mechanism based on machine learning to detect DoS attacks, that may be performed using different protocols such as ICMP, TCP or UDP.

II. Scope:

A Denial of Service (DoS) attack is an attempt by an attacker to prevent legitimate users from accessing services. When this attack is carried out by multiple attackers concurrently, the attack is known as a Distributed Denial of Service (DDoS) attack.
For example, a DoS attack can be carried out by initiating an ICMP flood from the attacking system to the victim's IP address. This results in bandwidth consumption and the victim system is overloaded, leading to denial of service.

III. Novelty:

We will be using different python libraries and cloud platform for detecting Dos attack. Also, we will use predefined and trained machine learing model sets for detecting at a faster pace.

IV. Comparative Statement:

X. Yuan, C. Li and X. Li, "DeepDefense: Identifying DDoS Attack via Deep Learning," 2017 IEEE International Conference on Smart Computing (SMARTCOMP), Hong Kong, China, 2017, pp. 1-8, doi: 10.1109/SMARTCOMP.2017.7946998.

This paper proposes a deep learning-based DDoS attack detection approach (DeepDefense). Deep learning approach can naturally extract high-level features from low-level ones and gain powerful representation and inference. A recurrent deep neural network is designed in order to learn patterns from sequences of network traffic and trace network attack activities. The experimental results demonstrate a better performance of the model compared with conventional machine learning models. The error rate is reduced from 7.517% to 2.103% compared with conventional machine learning method in the larger data set.

Tuan, T.A., Long, H.V., Son, L.H. *et al.* Performance evaluation of Botnet DDoS attack detection using machine learning. *Evol. Intel.* 13, 283–294 (2020)

To over the problem of DDoS attack, various machine learning methods typically Support Vector Machine (SVM), Artificial Neural Network (ANN), Naïve Bayes (NB), Decision Tree (DT), and Unsupervised Learning (USML) (K-means, X-means etc.) were proposed. With the increasing popularity of Machine Learning in the field of Computer Security, it will be a remarkable accomplishment to carry out performance assessment of the machine learning methods given a common platform. This could assist developers in choosing a suitable method for their case studies and assist them in further research. This paper performed an experimental analysis of the machine learning methods for Botnet DDoS attack detection. The evaluation is done on the UNBS-NB 15 and KDD99 which are well-known publicity datasets for Botnet DDoS attack detection. Machine learning methods typically Support Vector Machine (SVM), Artificial Neural Network (ANN), Naïve Bayes (NB), Decision Tree (DT), and Unsupervised Learning (USML) are investigated for Accuracy, False Alarm Rate (FAR), Sensitivity, Specificity, False positive rate (FPR), AUC, and Matthews correlation coefficient (MCC) of datasets.

Zekri, M., El Kafhali, S., Aboutabit, N., & Saadi, Y. (2017, October). DDoS attack detection using machine learning techniques in cloud computing environments. In *2017 3rd International Conference of Cloud Computing Technologies and Applications (CloudTech)* (pp. 1-7). IEEE.

In a DDoS attack, the attacker usually uses innocent compromised computers (called zombies) by taking advantages of known or unknown bugs and vulnerabilities to send a large number of packets from these already-captured zombies to a server. This may occupy a major portion of network bandwidth of the victim cloud infrastructures or consume much of the server's time. Thus, in this work, a DDoS detection system based on the C.4.5 algorithm to mitigate the DDoS threat is designed. This algorithm, coupled with signature detection techniques, generates a decision tree to perform automatic, effective detection of signatures attacks for DDoS flooding attacks. This paper also selected other machine learning techniques and compared the obtained results in order to validate the system.

Aamir, M., Zaidi, S.M.A. DDoS attack detection with feature engineering and machine learning: the framework and performance evaluation. *Int. J. Inf. Secur.* 18, 761–785 (2019).

This paper applies an organized flow of feature engineering and machine learning to detect distributed denial-of-service (DDoS) attacks. Feature engineering has a focus to obtain the datasets of different dimensions with significant features, using feature selection methods of backward elimination, chi2, and information gain scores. Different supervised machine learning models are applied on the feature-engineered datasets to demonstrate the adaptability of datasets for machine learning under optimal tuning of parameters within given sets of values. The results show that substantial feature reduction is possible to make DDoS detection faster and optimized with minimal performance hit. The paper proposes a strategic-level framework which incorporates the necessary elements of feature engineering and machine learning with a defined flow of experimentation. The experiments show that approximately 68% reduction in the feature space is possible with an impact of only about 0.03% on accuracy.

Smys, S. (2019). DDOS attack detection in telecommunication network using machine learning. *Journal of Ubiquitous Computing and Communication Technologies (UCCT)*, *1*(01), 33-44.

The telecommunication network that is the assemblage of the terminal nodes enables the whole to be connected. The swift progress in the telecommunication networks and the information technology has enabled a seamless connection and the capacity to store and communicate vast scale of information in the form of text and voice that are sensitive. This makes the telecommunication networks prey to multiple cyber-threats of which the DDOS (distributed denial of service) are the more predominant type of the cyber-threat causing the denial of the services to the users. So, the paper utilizing the combination of the neural network and the support vector machine presents the detection and the classification method for the DDOS attacks in the telecommunication network. The performance evaluation using the network simulator-2 enables to have the enhanced detection accuracy for the proposed method.

V. Dataset/Platform:

We will be using the following tools and packages for this project:

a. Python 3.9: Code for the project will be written using Python programming language as it offers many packages to develop and implement machine learning algorithms.

b. Packages:
    i. Pandas: working with dataset required for model development.
    ii. NumPy: working with data multidimensional arrays and matrices.
    iii. Scikit-learn: implementation of various machine learning algorithms.
    iv. Matplotlib: to perform basic analysis before model development.
    v. Pickle: to save a trained ML model for later use.
    vi. tqdm: to show progress through progress bars.

c. Dataset: The dataset we will be using is the KDD Cup 1999 Data which includes a standard set of data about different network connections, simulated in a military network environment.

d. Jupyter Notebook: To present analysis results and model development process easily, we will use Jupyter Notebooks.