T. C.
ANKARA UNIVERSITY
ENGINEERING FACULTY
COMPUTER ENGINEERING DEPARTMENT

AN INTRUSION DETECTION SYSTEM FOR NEW CYBER ATTACKS

GRADUATION PROJECT PROPOSAL

Ayça Nur VANLI

17290128

SUPERVISOR

Asts. Prof. Ömer Özgür TANRIÖVER

ANKARA – 2020

01/11/2020

1. INTRODUCTION

1.1.Problem Definition

Intrusion Detection System (IDS) is a commonly used device for cyber security solutions. The aim of this device is to monitor the network traffic and alert the administrator for any malicious actions in the traffic. We can classify three types of IDS based on their input data which are Network IDS that takes input from the network packets, Host IDS that takes input from the specific host's traffic, and Application IDS that takes input as high-risk applications traffic. [1]

With the soaring expansion of the Internet over the years, every network has been facing a tremendous amount of traffic with an exponential growth rate. For this reason, it has become inevitable to abandon the traditional network IDS. This traditional approach to detecting an intrusion is called signature-based IDS where the IDS's database tries to match the actions with the already detected attacks. Hence, the industry started to consider a more intellectual way to deal with their big data and they started to improve the efficiency of their IDS. When we are considering big data like network traffic, Machine Learning (ML) algorithms can be a remedy o handle them. Therefore, machine learning algorithms can be designed to detect the bad intended traffic in a network. So, the other approach to detect an intrusion is called the anomaly-based approach where the IDS is capable to determine the obscure attacks based on abnormal behaviors.

Nowadays, we can see the third approach which is the hybrid approach. The hybrid approach is a great solution where the IDS can detect both the unknown attacks and the known. We can see the application of such IDS in the internet of things, smart city, big data environment, fog computing, and mobile [2].

Some of the severe attacks that the IDS try to prevent are Denial of Service (DoS), Remote to Local (R2L), User to Root (U2R), and Probing. These attacks usually form the main attack categories for IDS. DoS attack aims to exhaust the resources such as memory, CPU, or bandwidth on a computer to prohibit legitimate users. R2L attack is an attempt to authenticate into the local services where the attacker has access to the network. U2R attack is the privilege escalation attack where the attacker authenticates the system but tries to authenticate to be the root/superuser. Probe attacks are malicious activities where the attacker aims to gather as much information as s/he can such as using the Nmap[1] tool. [3]

1.2.Aim of the Project

The aim of this project is to develop a new generation network IDS that is enhanced with machine learning algorithms. The IDS will use the anomaly-based approach where it can detect the malicious packets in a given network. The scope of the given network will be generally assumed to be an enterprise that uses TCP/IP [4].

[1] Nmap is an open-source network scanner that provides the information about hosts and services within the network.

Moreover, I will examine the possible way where the attackers can crack the IDS. I will study the fragile aspect of the IDS where the hackers can bypass the IDS. This aspect will mainly display the penetration test of the ML based security systems. The common approaches are about evading ML malware classifiers and using ML models to evade ML enhanced security devices. I will research the backdoors of ML.

2. METHOD

The ML algorithms will be used to implement an advanced IDS. Python and scikit-learn will be used for the most of the ML. The Keras will be used for deep learning. For the evaluation of the ML algorithms, the two metric systems will be referenced. These systems are accuracy and False Positive Rate (FPR).

For the classifier algorithms, I will consider using the Decision Table classifier because it can give the lowest value of FPR. Also, I will consider using the Random Forest classifier because it provides the highest average accuracy rate [5]. Moreover, I am planning to try the N-grams classifier.

Speaking of the data set of the project, it is a common approach for researchers to use KDDCup99[2] or NSL-KDD[3] for their projects. Yet, it is stated that this data set is already expired in the industry where it is more than twenty years old. With the light of this information, some researchers point this aspect to cause static growth in IDS [1]. After this clarification, I will consider using these data sets in my project but I will continue to search for further data sets that are more diverse and relatively new such as UNSW-NB-15[4] [6] or CICIDS2017[5] [7].

3. STUDY PLAN

The first step in my project will be implementing a machine learning model that can detect malware. I will train the classifier and handle the errors. Then, I will dive into the deep learning section. When I can detect malware, I will further my project with intrusion detection where I try data set and evaluate the success of my ML model. Then, I will try to expand the ratio of my security device by performing a side ML CAPTCHA[6] system which is a test to verify the user's humanity. I will foresee that implementing this ML will take a minimum of one semester. If I successfully complete my ML model development, I will proceed with the second phase of my project which is red teaming.

After the spring semester begins, I will start to pentest my own ML model and try to hack into the system. I will search for the nooks and crevices of the model and try to

[2] KDDCup99 is a data set that was formed based on a military network traffics in 1999. The last update was on October 28, 1999.
[3] NSL-KDD is a derivative of the KDDCup99 data set.
[4] UNSW_NB15 is a data set that formed by the University of South West in 2015. It has more attack types than the KDDCup99 does.
[5] CICIDS2017 is one of the newest datasets where it contains up-to-date attacks and scenarios. It is published in 2017.
[6] CAPTCHA stands for Completely Automated Public Turing test to tell Computer and Humans s Apart.

determine weak-points and take counter-measures. I will start with hacking into my CAPTCHA system using a deep neural network. After these tasks finish, I will use smart fuzzing to exploit the system. Then, I will try to evade the ML malware classifiers. When I complete that, I will do white-box and black-box attacks to my ML model. After these tasks, I will try to do ML poisoning and backdoor attacks on ML. Finally, I will examine the deepfake portion and possible attacks about it. I foresee that this pentesting and abusing my ML system will take at least one semester.

REFERENCES

[1] A. A. N. J. M. S. SH Kok, "A Review of Intrusion Detection System using Machine Learning Approach," *International Journal of Engineering Research and Technology,* vol. 12, no. 1, pp. 8-15, 2019.

[2] T. Saranya, "Performance Analysis of Machine Learning Algorithms in Intrusion Detection System: A Review," *Procedia Computer Science,* vol. 171, pp. 1251-1260, 2020.

[3] K. K. O ATA, "Network Intrusion Detection Using Machine Learning Technologies," *Aurum Journal of Engineering Systems and Architecture,* vol. 2, no. 1, pp. 115-123, 2018.

[4] E. B. CJ Ugochukwu, "An Intrusion Detection System Using Machine Learning Algorithm," *International Journal of Computer Science and Mathematical Theory,* vol. 4, no. 1, pp. 39-48, 2018.

[5] M. A. M. K. S. A. M. Almseidin, "Evaluation of Machine Learning Algorithms for Intrusion Detection System," *IEEE 15th International Symposium on Intelligent Systems and Informatics (SISY),* pp. 277-282, 2017.

[6] B. L. H Liu, "Machine Learning and Deep Learning Methods for Intrusion Detection System: A Survey," *Applied Sciences,* vol. 9, no. 4396, pp. 1-28, 2019.

[7] S. R. VR Varanasi, "Intrusion Detection using Machine Learning and," *International Journal of Recent Technology and Engineering (IJRTE),* vol. 8, no. 4, pp. 9704-9720, 2019.