Thanawat Tejapijaya

Weekly report Week 11 (31/7/2023 - 4/7/2023)

For my progress report this week, I have finished writing the paper as shown below,

Botnet Detection by Integrating Multiple Machine Learning Models

Thanawat Tejapijaya[1], Prarinya Siritanawan[2,*], Karin Sumongkayothin[1,*], and Kotani Kazunori[2]
[1]Mahidol University, Nakhon Pathom, Thailand
[2]Japan Advanced Institute of Science and Technology, Nomi, Ishikawa, Japan
oujeete@hotmail.com, prarinya@jaist.ac.jp, karin.sum@mahidol.edu, ikko@jaist.ac.jp
*corresponding author

Abstract—Botnets are persistent and adaptable cybersecurity threats, displaying diverse behaviors orchestrated by various attacker groups. Their ability to operate stealthily or on a massive scale poses challenges to conventional security monitoring systems like Security Information Event Management (SIEM). In this study, we propose a integrating machine learning methodology to effectively identify different types of botnet activities. Our approach involves training individual models using random forest and distinct network traffic characteristics. To evaluate the effectiveness of our methodology, we compare various integrating strategies. The evaluation is conducted using unseen network traffic data, revealing occasional high false positive rates but achieving a remarkable reduction in false negatives. The results demonstrate the potential of our integrating methodology to detect different botnet behaviors, enhancing cybersecurity defense against this notorious threat.

Keywords–botnet detection; models integration; anomaly detection; machine learning

1. Introduction

Botnets refer to groups of computers that have been compromised by malware and are under the remote control of a single malicious entity as known as botmaster. The term "botnet" is derived from combining "robot" and "network" due to the computer's robot-like behavior, executing wicked actions like spamming and Distributed Denial of Service (DDoS) attacks over the internet, local access, and other channels. These malevolent networks represent one of the most aggressive cyber attack threats, characterized by their elusive nature and evolving behaviors. As botnet techniques, life cycles, and behaviors constantly change, detecting their presence becomes an immense challenge for defenders, particularly Security Operations Center (SOC) personnel. Furthermore, botnet detection often demands a considerable amount of time and resources, and the limited availability of experts in this defensive field exacerbates the problem. The severity of botnet attacks were demonstrated in the 2018 incident involving the learning emerges as a promising solution. The motivation for this research lies in utilizing machine learning algorithms to expedite and enhance the detection of botnet activities across networks. By leveraging machine learning, the detection process can become faster and more efficient, increasing the likelihood of identifying botnets before they inflict severe damage. SIEM outputs often contain significant noise, allowing botnets to employ stealthy techniques that bypass conventional detection methods. Additionally, machine learning models can operate continuously without rest, unlike human analysts, thus alleviating the workload on SOC teams. Therefore, the primary goal of this research is to develop a robust and efficient model for botnet detection in computer networks. This research paper comprises five sections: literature review, methodology, experimental results, conclusion, and future work.

2. Literature Review

As botnets have evolved over time, research on using various machine learning techniques for botnet detection has also seen significant growth. The earliest botnets such as Puppe, Game Manager, and Bartender were created in 1988 with non-malicious intent along with the creation of internet relay chat as known as IRC where the purpose of it is used for real-time communication on the internet, unfortunately, due to it being anonymous, and efficient nature botmasters used it to control the botnet, but the first malicious botnets, Sub7 and Pretty Park, emerged in 1999, marking the beginning of botnet evolution. Since then, there have been numerous studies exploring different techniques to understand and combat these malicious botnets.

In 2006, researchers delved into the multifaceted approach to understanding and analyzing the phenomenon of malicious botnets, providing insights into their structure, life cycles, taxonomy, and more [15]. The same year witnessed the publication of an algorithm for Anomaly Botnet-Based detection, which effectively detected botnets in IRC channels. Another significant contribution came in the form of BotHunter, a new method that had a substantial impact on botnet detection by aiding operational use and stimulating further research in

This is just a part of it for this week it will be the final week in JAIST, as I will check through my stuff and preparing to go back to Thailand