Machine learning on network analysis

Susmita Rai - 908928

Contents

1 Introduction

With the rapid expansion of the Internet, it has become an essential part of our lives with over half the population connected [1]. However, this results in an increasingly complex and fragile network. Many systems are left vulnerable, waiting to be exploited. By 2021, it is predicted that cost of cyber-attacks will reach $6 trillion [2]. The importance of a good and secure security system is far too crucial.

"Offensive cyber capabilities are developing more rapidly than our ability to deal with hostile incidents" [3]. Attacks are becoming smarter, polymorphic viruses and obscured malwares are passing through current systems. Over a third of organizations believe that the threats they are facing cannot be blocked by their anti-virus [2]. Due to our rapid growth, we have left many openings for an attack, one of them is through the network. Our need for constantly being connected is causing a major gap in security. In 2017, 8 different network attacks dominated the market [4].

1. Browser attacks - malicious users target vulnerable websites to infect, infecting new genuine users.
2. Brute force attack - attempting to guess your way through to the system.
3. Denial of service (DoS) or Distributed Denial of Service (DDoS) – flooding a service by creating many requests in order to slow or crash the system.
4. Worm attacks – self propagating program that spreads through local system through exploitable vulnerabilities.
5. Malware attacks – programs that can take many forms, however their purpose is always malicious.
6. Web attacks – exploiting vulnerabilities found in the website such as SQL injection.
7. Scan attacks – indirect attack to gain knowledge of any vulnerabilities that exist such as an open port.

8. Other attacks – attacks that were out of scope, such as physically attempting to steal device.

Fortunately, methods such as Intrusion detection system (IDS) exist to deter most of these attacks. IDS constantly scan the network for any anomalous activity in the network. Some are even capable of stopping the attack completely rather than just alerting the user.

However, IDS face many issues such as explaining what an anomaly is in the first place. Robustness and accuracy also come into question. How often does an IDS system report false negatives or how many different types of malwares can they detect?

By using machine learning it is possible to overcome these problems. Its ability of learning patterns and understanding different classifications can assist IDS. Todo add more here about machine learning. and what i want to do

I aim to create a system that can detect malwares on a network, and also test its robustness and accuracy rates. todo think more about this

References

[1] Global internet usage in 2019. https://wearesocial.com/blog/2019/01/digital-2019-global-internet-use-accelerates. Accessed: 2019-10-18.

[2] Cybersecurity statistics of 2019. https://www.varonis.com/blog/cybersecurity-statistics/. Accessed: 2019-10-18.

[3] World Economic Forum. *The Global Risks Report 2018*, volume 13. 2018.

[4] Top 8 network attacks by type in 2017. https://www.calyptix.com/top-threats/top-8-network-attacks-type-2017/. Accessed: 2019-10-18.