Chapter 1

INTRODUCTION

1.1 PROBLEM STATEMENT

Network systems contain valuable data and resources that must be protected from attackers. Security experts often use honeypots and honeynets to protect network systems.
Honeypot is an outstanding technology that security experts use to tap new hacking techniques from attackers and intruders.
The purpose of Honeypot systems is to log every possible malicious activity of an attacker depending on the type of Honeypot system implemented within infrastructure.
Honeypot systems can be used to identify different types of malicious activities such as web applications attacks, known vulnerability exploitation, exploitation of outdated software/system and automated attacks by malicious bots.

1.2 Scope of the Project:

• The main aim is to analysis and detect the malicious actors and his attack pattern.

• Prediction of the pattern so that the know attack can be mitigated.

1.3 Objective

• To use free and open-source technologies and methods to reduce the amount of manual intervention needed to add to or modify high-interaction honeypots.

• Identify malicious IP addresses and vulnerabilities being exploited, so that

they can be patched as quickly as possible.

• Discover users that exhibit risky behaviour way before they execute the intended harm to the network.

• Deploy a number of decoys to lure attackers that get past other defenses.

• To detect attack patterns using machine learning and come out with solution to mitigate the attacks.

Chapter 2

REVIEW OF LITERATURE

Liberios Vokorokos, et.al, 2013 [1], proposed the urbane hybrid honeypot system. These systems Propagates and maintains the interaction with attackers and record all activities and perform data analysis, thus allowing improving security of computer systems. Furthermore in order to induce security authors also did managed to amalgamate passive fingerprinting technique. It also promotes the implementation of multiple Decoys (two Decoy Servers) in order to reduce the probability of missing the malicious activity on the server by changing the level of interaction.

Albert Sagala, 2015 [2], emerged with an idea of collaborative honeypot and intrusion Detection System where in the logs file from Honeypot server is passed on to the snort in order to generate the rules for Snort that acts similar to the firewall. The rules for the SNORT will be inevitably generated by the IDS using the logs provided by the honeypot tracked by the system. The rules generated are in the form of alerts for illegal activity.

Rahul koul [2017] Modern Attack Detection Using Intelligent Honeypot A Honeypot is a network system to determine the unauthorized use of information system by analyzing the behaviour of attacker in an isolated and monitored environment. But, there are tons of honeypot implementation implemented till date, however one thing missing from each of the honeypot implementation is continuous learning of trending attack scenarios and no human decision-making capabilities. In this paper, we proposed a solution for detecting modern attacks by introducing a semi automatic approach of attack detection via honeypot coupled with human decisionmaking capabilities.

Similarly, Jaiganesh, Sumathi, and Vinitha (2013) compared two data mining classification algorithms for intrusion detection system, Iterative Dichotomiser2 (ID3) algorithm and C4.5 algorithm. The results presented by Jaiganesh et al. (2013) shows C4.5 algorithm was best suited for intrusion detection, because it uses numeric and nominal data

Chaitanya D Patil, Thyagarajamurthy A "Integration of Honeypots and Machine Learning in Network Security" Conventionally, corporate information security focuses on deploying antivirus and antimalware tools that provided defensive security control of the business information systems. In this security approach, the information security officer relies on the signature or heuristic detection strategies adopted by the security solution to identify and mitigate threats facing the company. Although this framework offers robust security management for the known threats, it fails to effectively analyze the nature of the attack, review the risk, and understand the intent of the attacker. Such weaknesses limit the ability of this security management framework to effectively contain new threats, which do not have their signatures or heuristics extracted yet. It is due to these security management flaws that honeypots are developed and deployed at the corporate network. A honeypot is a security resource designed and implemented to lure attackers, enabling them to probe, attack, and compromise the system [4].

Sr No.	Year	Name of the Paper	Author	Content
1	2017	Modern Attack Detection Using Intelligent Honeypot	Rahul Koul, J. W. Bakal, Sahil Dhar	System Design
2	2016	Design and Implementation of a Real-Time Honeypot System for the Detection and Prevention of Systems Attacks	Farouk Samu, Dr. Amos O. Olagunju, Dr. Ezzat Kirmani, Dr. Jerry Wellik	System Set up
3	2019	DDos Attack Detection In Telecommunication Network using Machine learning	Dr.A.Pasumpon pandian, Dr.S. Smys	Machine Learning
4	2017	Detection of Severe SSH Attacks Using Honeypot Servers and Machine Learning Techniques	Gokul Kannan Sadasivam, Chittaranjan Hota, Bhojan Anand	SSH honeypot
5	2020	Integration of Honeypots and Machine Learning in Network Security	Chaitanya D Patil , Thyagarajamurthy A	Machine learning
6	2015	Threat Prediction Using Honeypot and Machine Learning	Vishal Mehta, Pushpendra Bahadur , Manik Kapoor , Dr. Preeti Singh and Dr. Subhadra Rajpoot	-
7	2010	Honeypots in Network Security	Deniz Akkaya	Honeypot Setup

Chapter 3

SYSTEM DESIGN AND ANALYSIS

3.1 BLOCK DIAGRAM

Block Diagram for System Set-up

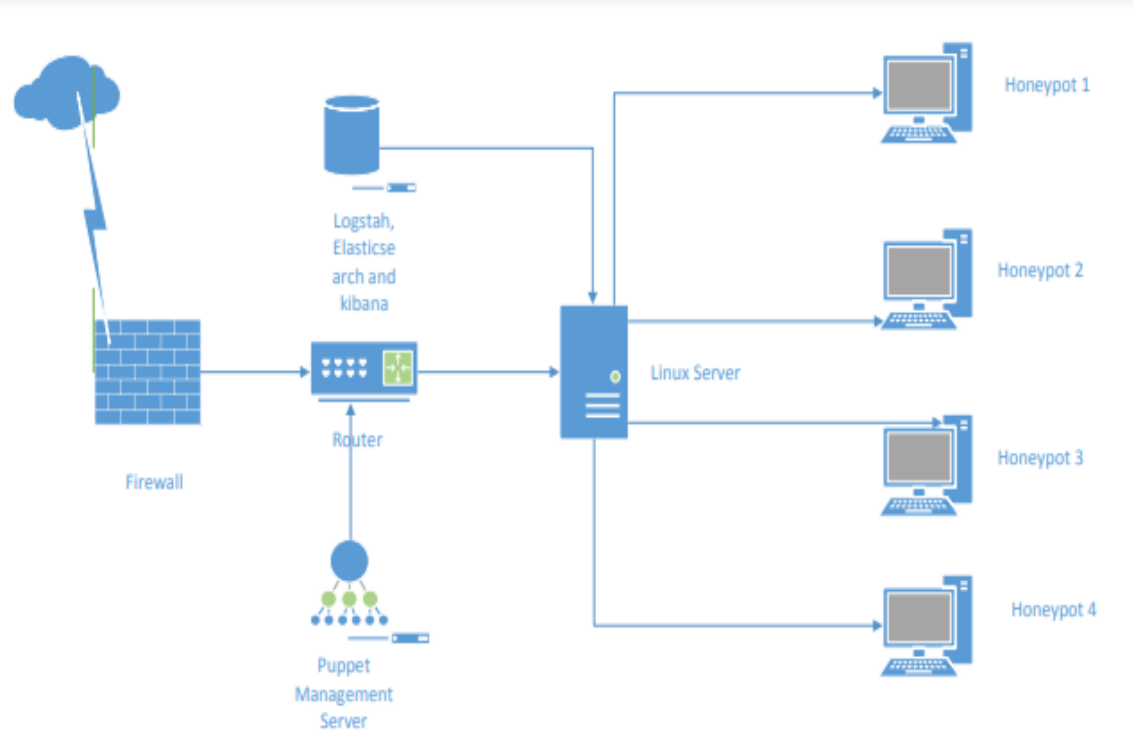

Figure 3.1: system set-up

Block Diagram for Machine Learning Model:

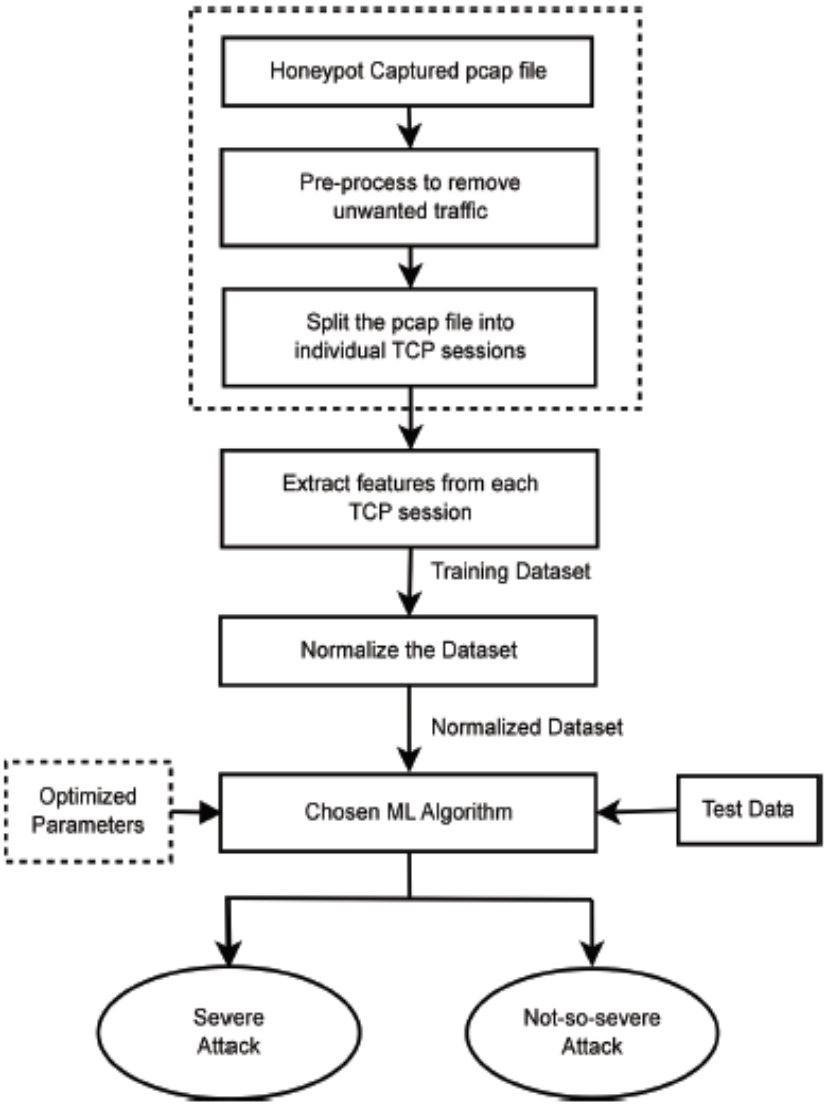

Figure 3.2: Machine Learning Model

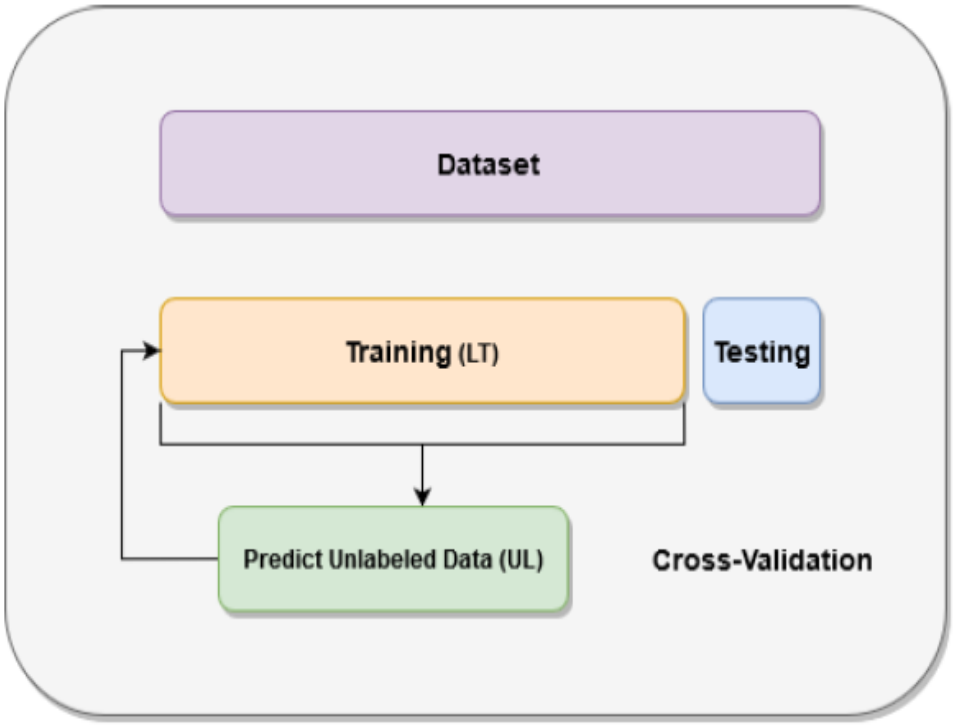

Figure 3.3: Splitting Data

Chapter 4

IMPLEMENTATION STRATEGY

4.1 METHODOLOGY

• The virtual machine based honeynet has of 4 honeypot systems, a centralized logging host and Puppet automation host. Each honeypot system is consisted with three individual hosts;a router, a firewall, and a Linux server host. The Linux server provided a Secure Shell (SSH) service to act as the target for attackers.

• Data will be collected using Logstash, an open source data collection engine with real-time pipelining capabilities. Logstash can dynamically unify data from disparate sources and normalize the data into destinations. It cleanses and democratizes data for diverse advanced downstream analytics and visualization use. This was chosen because of its centralized log management which was used to collect all logs from compromised honeypots.

• Analysis on source IP addresses, attack patterns, countries, username and passwords used to gain access, source ports, commands etc.

• Keylogging will be an important part of the analysis process thus we can also tell apart between a real human interaction or a bot trying to brute force its way through.

• Supervised machine learning algorithms will be used to classify SSH attacks based on attack nature Split into two data Severe and Not so severe attacks.

• Honeypotter will focus on detection, forensics and response including:

o Real-time customizable and detection from multiple sources, network behaviour, optional automation, full session assembly, protocol decoding and full content analysis.

4.2 Hardware and Software Requirement:

Hardware Requirement:

Environment	Version / Specifications
OS	Linux
CPU	I7 10th gen
RAM	12 GB
Storage	500 GB

Figure 4.1: Hardware Requirement

Software Requirement:

Puppet : It is an automation software for IT system administrators. It is used to automate repetitive tasks such as the installation of applications and services, patch management, and deployments.

Elasticsearch : It is used because it is able to achieve fast search responses instead of searching the text directly, it searches an index instead. Elasticsearch is designed to be scalable and distributed.

Kibana : Kibana is an open source (Apache Licensed) browser based analytics and search dashboard for Elasticsearch that visualized the data to provide a better interpretation. It was used to visualize captured logs from compromised honeypots.

HonSSH : A high-interaction HoneyPot solution designed to log all SSH communications between a client and server.

4.3 FUTURE PLANS

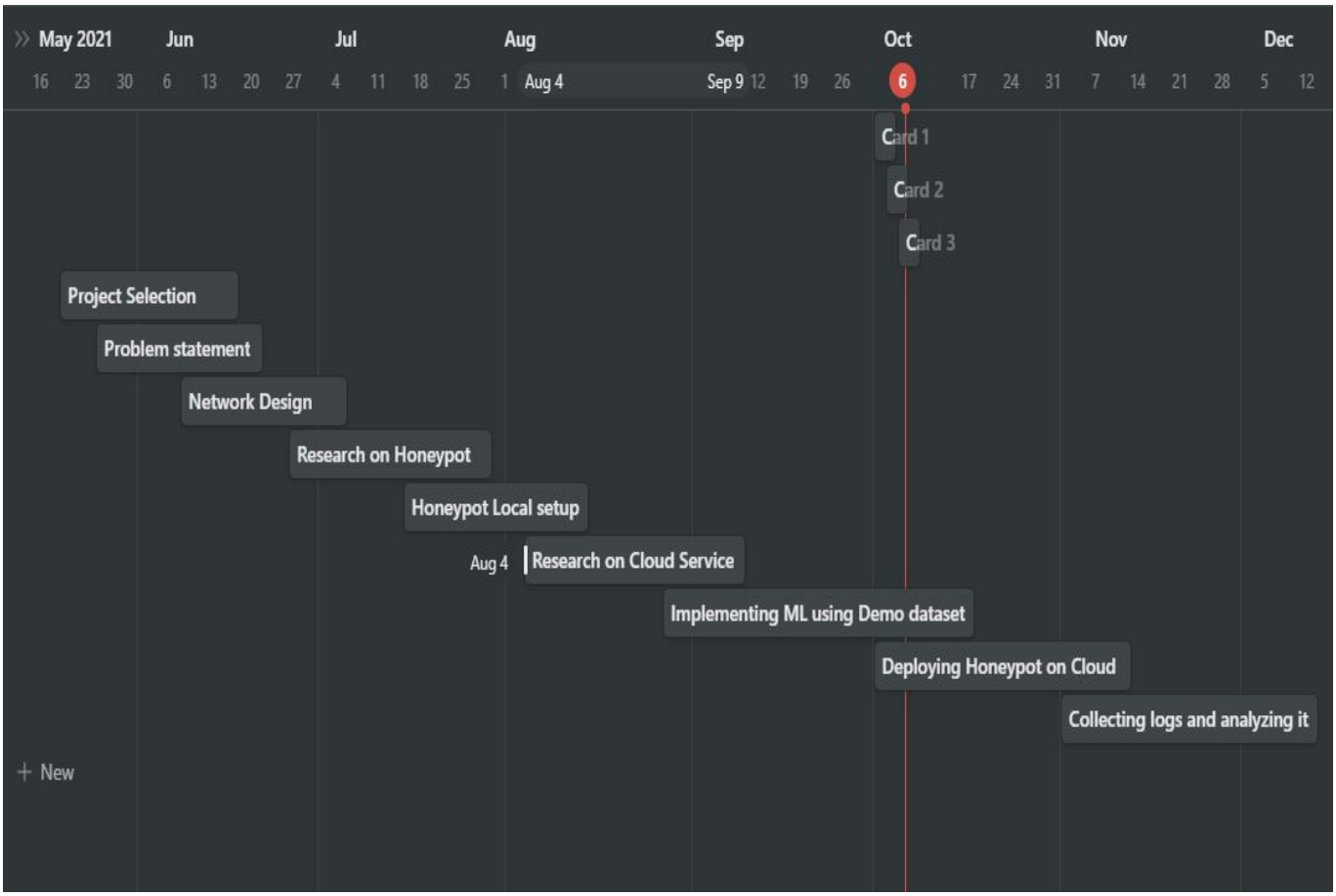

Figure 4.2: Time Plan

Chapter 5

CONCLUSION

The research designed and implemented a real-time Honeynet system using Machine learning for detecting and preventing system attacks. System services on Apache Webserver, MYSQL, FTP and SMTP were used to lure attackers.The problem nowadays is that a very good hacker will most likely be able to understand when he is attacking a honeypot. Low interaction honeypots will be able to identify mostly automated attack and will hardly be able to understand new hacker method. On the other hand, high interaction systems are here to entrap the hacker and make him give away his techniques and tools to the forensic team. The network administrator implementing this kind of honeypot should make sure that the system is completely isolated from the production network. This is the best defense if the hacker compromises the honeypot.We have proposed a machine learning model to classify SSH attacks based on the attack nature

In this semester we are done with literature survey and in next semester we will implement it.

REFERENCES

[1] Rahul koul, J. W. Bakal, Sahil Dhar "Modern Attack Detection Using Intelligent Honeypot "

[2] Farouk Samu , Dr. Amos O. Olagunju, Dr. Ezzat Kirmani , Dr. Jerry Wellik "Design and Implementation of a Real-Time Honeypot System for the Detection and Prevention of Systems Attacks"

[3] Dr. A. Pasumpon pandian ,Dr.S. Smys, "DDos Attack Detection In Telecommunication Network using Machine learning"

[4] Gokul Kannan Sadasivam, Chittaranjan Hota, Bhojan Anand "Detection of Severe SSH Attacks Using Honeypot Servers and Machine Learning Techniques"

[5] Chaitanya D Patil , Thyagarajamurthy A "Integration of Honeypots and Machine Learning in Network Security"

[6] Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011

[7] Newman, Sean. "Under the radar: the danger of stealthy DDoS attacks." Network Security 2019, no. 2 (2019): 18-19.

[8]Attack Scenario Prediction Methodology Fayyad, S. ; Meinel, C. Information Technology: New Generations (ITNG), 2013 Tenth International Conference