Malware Capturing and Analysis using Dionaea Honeypot

Vasu Sethia
UG Student, Dept. of CSE, Faculty of Engg. & Tech, SRM Institute of Science and Technology
Chennai, Tamil Nadu, India
vasusethia_manoj@srmuniv.edu.in

Jeyasekar A
Associate Prof. , Dept. of CSE, Faculty of Engg. & Tech, SRM Institute of Science and Technology
Chennai, Tamil Nadu, India
ajeyasekar@yahoo.com

***Abstract*— The growth of internet and users has increased exponentially and drastically in this decade. It provides services inheriting various benefits to the users such as online banking, marketing, buying/selling and various facility management services etc. It attracts some people to develop programs that perform various malicious activities intentionally or unintentionally such as stealing sensitive information from computer, displaying advertisement, causing harmful, unwanted activities. The malicious software is referred as malware, viruses, worms, spyware, trojans, adware and botnets are also malwares. Therefore, in this paper, we run a honeypot and capture various zero-day attacks and malwares. The harmful and malicious activities of these malwares are studied by analyzing the network logs of the honeypot. Based on their properties and activities, the malwares are classified. These classifications of malware and their behavior help researchers to develop a security mechanism to prevent an organization from these kinds of harmful and malicious activities triggered by the malwares**.

Keywords*— *Malware, Honeypot, Dionaea, Malware capturing system, Malware analysis

I. Introduction

Internet plays an important role in our daily life starting from buying groceries to paying our bills. Just as technology makes life easier for everyone , it leaves top companies and their customers more vulnerable to cyber attacks. The global cost of cybercrime is expected to reach $2 trillion by 2019, triple the amount from 2015, and studies show most of the small businesses are at risk . A successful cyber-attack can totally sabotage the organization .Therefore it is necessary for all the organization to keep up their security mechanism and prevent them from the various cyber-attacks.

Malware is a broad term, but a very basic definition of it would be anything that performs any malicious activity .Viruses, worms, ransomware, Trojans etc comes under malware. The aim of most of the malwares is to steal sensitive information or encrypt the data and demand for ransoms [1]. It is also used to perform attacks on organizations by black hat hackers who have an extensive knowledge to break into secure networks and destroy, modify or steal sensitive data in the network and sometimes making the network unusable for authorized users. Hence each malware performs some kind of malicious activities. Therefore one could define an Intrusion Detection System (IDS) to block the malware in a network by creating a signature of a malware sample [5].In order to capture the malware entering into the network a honeypot is set up at the entry point of the network. It means that the malware analysis plays a vital role for an IDS system.

Hence in this paper, we present an analysis on malwares and propose a malware capturing system for an organization. Rest of the paper is organized as follows. In section 2, the overview and architecture of honeypot mechanism is presented and in section 3, the classifications of honeypot is described and in section 4, different techniques used for malware analysis is described. In Section 5, Dionaea based honeypot system and the testbed used to capture the malware are described. Section 6 discusses on experimental setup and result analysis and finally we conclude the paper in section 7.

II. Overview of Honeypot Mechanism

A honeypot is a fake system used at the entry point of network for trapping hackers, often deceiving the malware from attacking the real severs and to identify the malicious activities performed over the Internet. Honeypot serves as a basic network security mechanism for an organization [6]. It is deployed in order to distract the attacker from attacking the main server. Fig 1 shows the deployment of honeypot in an organization’s network. The firewall generally uses various policies to allow the packet to or from a local area network. Therefore the incoming packets are passed through the honeypot and firewall. Honeypot emulates various services depending upon its interaction level.

The working of honeypot depends upon the level of interaction it does .Low interaction honeypots like Dionaea, cowrie etc are easier to maintain and deploy. They emulate various system services and make the system look vulnerable. High interaction honeypot runs and doesn’t emulate all the services which are there present in the production server. But these systems are more resource-intensive and are difficult to set up and maintain. They are often used to study attack vectors and mindset of an attacker. The system in which the attacker breaks in should be monitored properly .All the logs, TCP and UDP connections should be monitored to study the attack .The system running the honeypot must contain files which make it seems like a legitimate server. As our dependence on internet increases, a lot of people are carrying out attacks to steal information, ask for ransom to decrypt the data and other malicious activities .The motivation behind capturing the malwares was to study the mechanism of a malware and how it interacts with the operating system .What kind of activities it performs over network. Later we can easily develop and deploy a security mechanism to protect people and various organizations from these kind of malware attacks. It helps us in assessing the severity of a Cyber-attack and what all methodologies are being used by an attacker to bypass all the security mechanism[9][5] we have till now.

978-1-7281-1576-4/19/$31.00 ©2019 IEEE

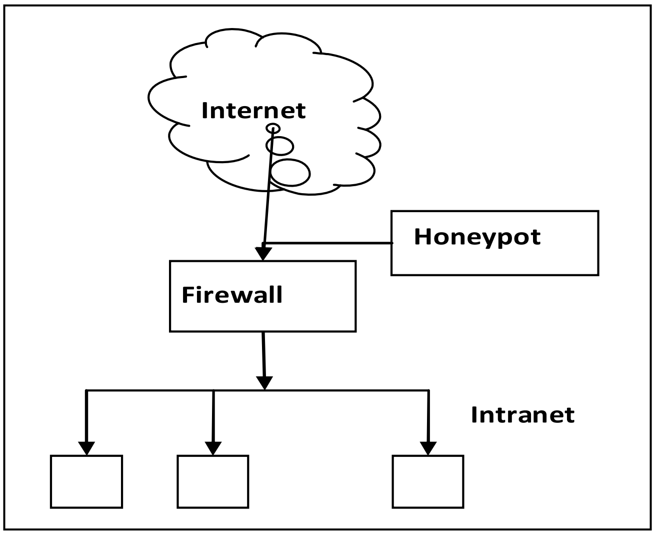

Fig. 1 Architecture of a network with honeypot and firewall

III. Classification of Honeypot

The growth of cybercrimes and cyber attacks has increased dramatically in the recent days which opens the need of high security mechanisms. Honeypot is one such mechanism which is used to trap the attackers, diverts the attention of the attackers from the real server and capture the new viruses , worms for future investigation. Honeypots are classified based on the level of interaction, type of application and physical deployment.

- Level of interaction: The level of interaction means range of attack possibilities that a honeypot allows an attacker to have with real system [3].
- Type of Application: The honeypot is deployed in a network for either research purposes or production purposes.
- Physical deployment: The honeypot is physically deployed either inside the network or outside the network

A. Level of Interaction

There are three types of honeypot under the category of level interaction: 1) High interaction honeypot, 2) Low interaction honeypot, 3) Hybrid honeypot. High interaction honeypot allows the attackers interacting with real system and does not assume anything about the possible behavior of the attacker. It helps the administrator to investigate the complete activities of the attackers. It utilizes the operating systems (eg Linux) and applications (eg FTP) without using the emulation software (eg dionaea) to record the activities of attackers [3]. Low interaction honeypot captures the information about the predetermined activities of attackers and limits the activities of attackers by using emulation software. It exclusively utilizes the emulation software and its services. Therefore for example, emulated FTP service using port 21 supports login and other support FTP commands and it does not link the real FTP server because it contains sensitive information. Hence the quality of low interaction honeypot is decided by the quality of emulation software. Example for low interaction honeypots are honeyd, Nepenthes, Mwcollect, Dionaea, Amun. Hybrid honeypot integrates the high interaction honeypots and low interaction honeypots which helps the administrator with far more information about the activities of the attackers.

B. Type of Application

Production of honeypot : It is deployed by the organization to mitigate the threat and protect themselves from various attacks. It is less resource intensive and can be easily deployed. These are generally low interaction honeypot e.g. Dionaea which emulates various services to trick the attacker. Use of low-interaction honeypot includes port scan identification, generation of attack signatures, trend analysis and malware collection. On the other hand, this is also a disadvantage. It is not possible to watch how an attacker interacts with the operating system as all the services are emulated.

Research Honeypot: It is deployed by various security researchers to assess the cyber threat , study the real motivation behind the attacks and capture various zero-day exploits. It helps in gathering information about the attack and what kind of hacking (attacking) techniques are being used by the attackers. They mainly consists of high interaction honeypots like the Honeynet Project is a volunteer project that runs honeypots to assess cyber threat in which attackers interacts with the real time operating systems and services. Various logging tools are used to log each and every activity happening on the honeypot which can really help researchers to study the various advance methodologies used by an attacker. The biggest disadvantage of a high interaction honeypot is the time and effort it takes to build the decoy system at the start.

C. Physical Deployment

Physical Deployment means that the malware capturing system or the honeypot is actually running on the physical system unlike running honeypot services virtually or on the virtual machines, which actually emulates the vulnerable services .Generally high interaction honeypot are physical deployed which are difficult to maintain and install.

TABLE I. Comparison of Honeypot based on level of interaction

Parameters	Low-Interaction Honeypot	High-Interaction Honeypot
Degree of Involvement	Low	High
Physical Deployment	No	Yes
Interacting with OS	No	Yes
Information Gathering	Yes	All connections
Cost of Maintenance	Low	High
Installation	Easy	Hard
Knowledge Required to Run	Low	High
Emulated Services	Yes	No

IV. Techniques Used for Malware Analysis

There are mainly two ways of analysing a malware [2] that are static analysis[7] and dynamic analysis [8] of the malware. Static Analysis of malware involves techniques which try to determine the nature of the malware without executing or running the malware .It can be done using various tools like IDA, gdb etc. One should have great understanding of windows Portable executable file format to perform static analysis whenever you are dealing with windows malware. It involves various techniques like checking the strings present in the binary, entropy of the binary, api imported etc. Virus total platform is used to analyse the malware which uses signature based detection techniques. Dynamic analysis of malware is the process in which malicious binary is actually made to execute, and then we try to study the behavior of the malware and learn more about its functionality. This can be done using various tools like windows sysinternals, Wireshark, debugger etc. Dynamic analysis of a malware exposes the IP it is trying to connect, domain name it is trying to resolve, what all process are spawned by the process and various other behavior related properties. Cuckoo sandbox is automated dynamic platform which can be used to analyze the malware.

V. DIONAEA BASED MALWARE CAPTURING SYSTEM

It is a low interaction honeypot which was initially developed under The Honeynet Project's 2009 Google Summer of Code (GSoC). Dionaea emulates various vulnerable services. Once the malware attacks the vulnerable services we obtain the malware binary without actually affecting the base system. Dionaea runs in a restricted environment without administrative privileges. Protocols emulated by the Dionaea are SMB, HTTP, FTP, TFTP, MSSQL and VOIP. The common techniques used by the attacker are

- Shell Binding / Connect Back
- ExecURLDownToFile API
- Multi-Stage Payload

Dionaea acts upon these techniques used by the attackers to capture the malwares. Fig 2 shows the basic architecture of Dionaea honeypot. Dionaea offers logging system which logs all the activities in the clear text log files which can be used for analysis and attack prediction. These logs can be used for various debugging processes and to report the errors. Dionaea uses some internal communication system which is called incidents and this ihandler is much superior to the text based logging. Ihandler provides exact overview of the attack which is happening.

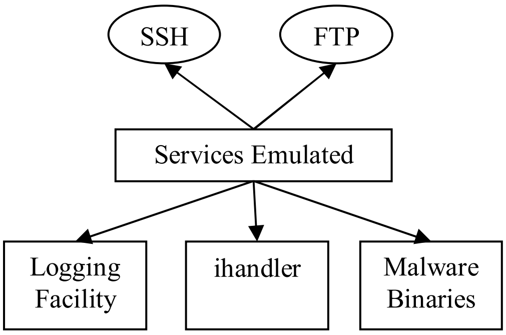

Fig.2 Architecture of Dionaea Honeypot

VI. EXPERIMENTAL TESTBED AND ANALYSIS

The experimental setup uses Amazon ec2 instance and Dionaea honeypot for capturing the malware. Dionaea honeypot runs on the amazon ec2 instance which is a virtual server in Amazon's Elastic Compute Cloud (ec2) for running applications on the Amazon Web Services (AWS) infrastructure.

Amazon ec2 instance is created using the steps shown in Table 2. After creating the ec2 instance, Dionaea is setup by following the steps shown in Table 3. This amazon ec2 Dionaea honeypot is connected to the computer used for analyzing the malware and another two computers are connected with amazon ec2 Dionaea honeypot that generates the attacks on ec2 instance.

A. Steps for Creating an Amazon ec2 Instance

- Open amazon ec2 console at https://console.aws.amazon.com/ec2.
- Click launch instance and select the operating system you want dionaea to run on and select ubuntu 14.04 image from amazon machine image.
- Now configure the network details of the honeypot. Instance Details, select "Auto-assign Public IP" and set it to "Enable"
- For storage, just add the default and click Next.
- Ignore adding Tags and click next .Default security policy in AWS only allows SSH connections (port number 22) to the ec2 instance open. Change the default security to open all the ports on the server. This way we can capture different malwares.
- Then Launch

B. Steps for Implementing Dionaea in Ubuntu 14.04

Login in to amazon ec2 instance and give the following commands into the terminal for implementation of Dionaea in Ubuntu 14.04.

- To update the package cache

  ```
  sudo apt-get update
  sudo apt-get install software-properties-common
  ```

- Install PPA and update the package

  ```
  sudo add-apt-repository ppa:honeynet/nightly
  sudo apt-get update
  ```

- Installation of Dionaea

  ```
  sudo apt-get install dionaea
  ```

- Once the service has been installed we can start it by using following unix command

  ```
  sudo service dionaea start
  ```

As stated by the website the logs are stored under /var/log/dionaea/ and everything else is under /var/lib/dionaea/ directory.

C. Analysis

In this experiment, it is observed that SQL 2000XP, SMB and FTP protocols are mostly attacked. Further to it, there are some high confidence malwares like WannaCry ransomware, Slammer and GandCrab. These malwares have high level of severity on the system. WannaCry ransomware attacks the SMB protocol of windows operating system. Slammer attacks the SQL 2000 XP version of windows. GandCrab ransomware is an infamous family of cryptoviruses that was first introduced in early 2018. All malware samples are made available for the analysis at the address http://18.223.255.191/binaries/. From the samples, the malwares are categorized based on their behavior and protocols attacked by the malwares. They are Dropper, Rootkit, Ransomware, Trojan, Adware, Coin Miner, Botnet, RAT (In RAT, Windows meterpreter shell was found). Windows meterpreter shell is the payload delivered by the attacker after exploiting the system. It gives the full control of the system to attacker and attacker then can perform any malicious task that they want to perform. The windows meterpreter shell can be a reverse shell in which the target machine communicates back to the attacking machine or it can be a bind shell in which attacker connects to the target machine.

Fig 3 shows the various types of malwares observed by implementing the Amazon ec2 instance based Dionaea honeypot. The proposed honeypot detects 3 high profile ransomwares and various categories of malwares. Many of the malware samples are Droppers or downloaders followed by Trojans. These malwares attack mostly SQL server 2000 for XP, SMB and FTP protocols and the number of the attacks over them is shown in Fig 4. These attacks are possible because the administrator configure the FTP server poorly due to which anonymous user are allowed to login and put data in the server.

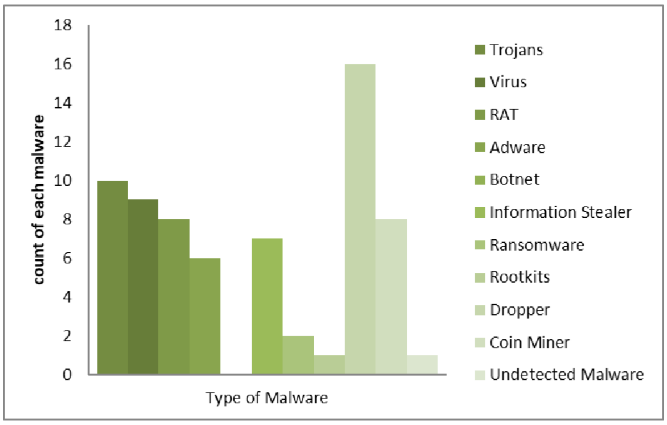

Fig. 3 Type of malwares observed

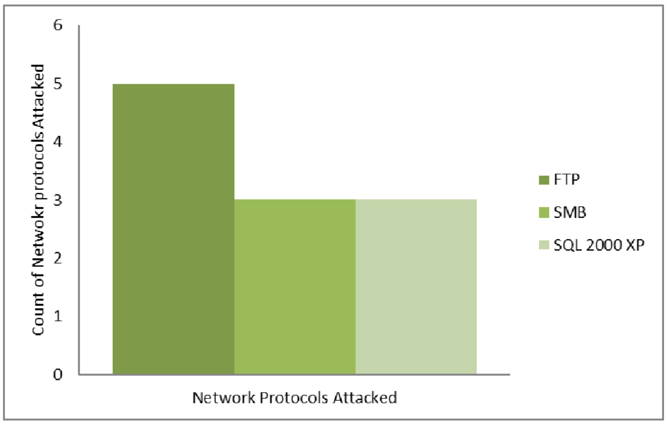

Fig 4. Count of each protocol attacked by the malware

VII. Conclusion

This paper introduces a mechanism to capture the malware present on the internet using the cloud. It presents how the ec2 instance, Dionaea are integrated together to capture the malware samples. The proposed mechanism finds various categories of malwares that target mostly the SQL server 2000 for XP, SMB and FTP protocol and one undetected malware. It is observed that it is very easy to exploit windows 7 or lower version which runs SMB version 1. Exploiting the SMB can be done by NSA and released by the shadow brokers. The proposed mechanism limits itself to low interaction honeypot and in future, it can be extended to high interaction honeypot. Further to it, the proposed mechanism can be used to detect zero-day malware and prevent them.

References

[1] Aru Okereke Eze, Chiaghana Chukwunonso E, “Malware Analysis and Mitigation in Information Preservation”, IOSR Journal of Computer Engineering, Vol. 20(4), pp. 53-62, Aug-2018.

[2] Belal Amro, “Malware Detection Techniques for Mobile Devices”, International Journal of Mobile Network Communications & Telematics, Vol. 7(4), pp. 1-10, Dec 2017.

[3] Dilsheer Ali P., Gireesh Kumar T., “Malware Capturing and Detection in Dionaea Honeypot”, International Conference on Innovation in Power and Advanced Computing Technologies, i-PACT2017, pp. 1-5.

[4] Gerard Wagener, Radu State, Alexandre Dulaunoy, “Malware Behaviour Analysis”, Journal of Computer Virology and Hacking Techniques, Vol. 4 (2008), pp. 279-287.

[5] Hassan Artail, Haidar Safa, Malek Sraj, Iyad Kuwatly, Zaid Al-Masri, “A Hybrid Honeypot Framework for Improving Intrusion Detection Systems in Protecting Organizational Networks”, Computers and Security, Vol 25 (2006), pp. 274-288.

[6] Kyi Lin Lin Kyaw, “ Hybrid Honeypot System for Network Security”, International Journal of Computer and Information Engineering, Vol. 2(12), pp. 4085-4089, 2008.

[7] Monirul Shanf, Vinod Yegneswaran, Hassen Saidiz, Philip Porras, Wenke Lee, “Eureka: A Framework for Enabling Static Malware Analysis”, Springer-Verlag LNCS 5283, pp. 481-500, 2008

[8] Waqas Aman, “A Framework for Analysis and Comparison of Dynamic Malware Analysis Tools”, International Journal of Network Security & Its Application, Vol. 6(5), pp. 63-74, 2014.

[9] Mr Mohit Tiwari, Raj Kumar, Akash Bharti, Jai Kishan “INTRUSION DETECTION SYSTEM”,International Journal of Technical Research and Applications e-ISSN: 2320-8163,www.ijtra.com, Vol 5(2) (March - April 2017), PP. 38-44