Honeypot : A Tool to Research Cyber Attack Techniques

Ramkumar Gupta[1] Kaiz Jessani[2] Manasiya Safan[3] Shweta Sharma[4]
[1,2,3]Students [4]Assistant Professor
[1,2,3,4]Department of Computer Engineering
[1,2,3,4]Atharva College Of Engineering,Mumbai - 400095, Maharashtra, India,University of Mumbai.

Abstract— The internet can be a dark and dangerous place; featuring viruses and cyber attacks. This paper aims to uncover some of these threats and reveal just how vulnerable the internet can be. The use of internet increases in daily life rapidly. Due to which many system run online such as payment system for online transactions, e - commerce system etc. Therefore there are some threat to these system . These are increasing in number very rapidly . The top five vulnerabilities are :

1) SSL/TLS Poodle Vulnerability
2) Cross-Site Scripting
3) SSL v2 support detected
4) SSL Weak Cipher Suites Supported
5) Invalid SSL certificate chain

There are various other vulnerabilities such as IP spoofing , SQL injection , DOS attack , DDOS attack. In the field of computer security, honeypots are systems aimed at deceiving malicious users who launch attacks against the servers and network infrastructure of various organizations. They can be deployed as protection mechanisms to an organization's real systems, or as research units to study and analyze the methods employed by individual hackers In this paper we will later present the results of a research honeypot's operation, which undertook the role of a web trap for attackers who target the SSH service in order to gain illegal server access Lastly we will present a visualization tool aimed at helping security researchers during the analysis and conclusions drawing phases, for use with the same SSH honeypot implementation software as outlined in this work

Key words: Cyber Attack, DOS attack, DDOS attack

I. INTRODUCTION

The massive use of computer systems and internet technologies has made information security one of the top concerns for today's reality. Malicious users and software launch attacks on a regular basis against specific or broad targets. The motivation behind these attacks varies and can be the simple act of a cyber vandal, corporate espionage for financial gain or state-sponsored attacks with geopolitical impact.

Attackers are searching the internet for servers that can be used for their malicious activities. One of the most prominent targets is servers on which the administrator has set up a remote access service, for example Secure Shell (SSH). Many times this canbe exploited by malicious users if a weak password is in a place as an authentication mechanism. Whenever an attacker finds such a server that runs the particular service, he will try to connect to it using various combinations of authentication credentials. If a successful login attempt is made, the attacker gains remote access to the server and then uses it for malicious activities such as malware installation or pivoting; using the compromised server to launch attacks against other systems.

Identifying and classifying these attacks along with their actors is a crucial step in developing security policies and strategies to effectively defend against them. To stay one step ahead of malicious users and get early warnings of new vulnerabilities and exploits, one can use devices known as honeypots. In their simplest form these devices act as decoy-based intrusion detection systems with full logging capabilities.

Many definitions for honeypots exist, but one of the most accurate belongs to Lance Spitzner who defines a honeypot as an information system resource whose value lies in unauthorized or illicit use of that resource [1]. A honeypot is a system with no production value. Its operation is based on the following concept [2]: there is no reason for a legitimate user to use it directly or interact with it, therefore any communication attempt with the system is automatically considered malicious and it is usually classified as one of the following: detection, network scanning or attack. Respectively, a honeypot that tries to connect to an external network resource has been probably compromised by an attacker [3].

Honeypots are both deceit tools and traps. They can lure malicious users by acting as systems that contain valuable data or interesting services. They allow to be exploited and then offer a simulated or real environment to the attacker to interact with, while logging all of his actions and activities. In such a way they help security professionals and researchers in the process of learning the techniques and methods used by attackers to compromise computer systems. Honeypots cannot prevent cyber attacks against the network by themselves, but they can help in identifying and detecting them when used alongside with other defense-oriented tools such as firewalls.

Honeypots often generate a small amount of data of high value [4] which is considered an advantage, but depending on the circumstances the analysis of this dataset can be a challenge for information security professionals. As the number of

attacks grows very large over time, it becomes impossible to manually analyze or compare each and every captured session. This is where data visualization solutions and visual analytics [5] can play a vital role in helping the defenders get a quick and detailed overview of a honeypot's operation.

II. RELATED WORK

A substantial number of studies of SSH attacks have been carried out in recent years. Some of them have been the result of academic affiliated work [6]- [8], while others have stemmed from the effort of information security professionals or companies [9]- [11]. In some of these cases the study of SSH attacks has been a portion of a bigger study, which mostly included profiling attackers or following their activities after they gained illegal system access. In our research we have included post-compromise activities in our scope as well. This gives us a better understanding of the actors and motivations behind the attacks, while the captured SSH login traffic serves the goal of developing a deeper understanding of the techniques employed in SSH attacks.

Regarding the visualization of attacks on computer networks, a lot of research has been done, but it is mostly focused on visualizing NetFlow data coming from attacks logged by an Intrusion Detection System (IDS). For example, the tool NFlowVis [12] was created in 2008 and it can be used to visually analyze attacks in large-scale networks using NetFlow data from IDS attack logs. While SSH-related visualizations were presented alongside, only connection attempts were shown. In 2009 the tool VIAssist [13] was created, which can provide the details of specific network flows in need to be examined further. However, VIAssist has no valuable practical use for the analysis of SSH attacks since it only visualizes NetFlow data.

Focusing more on visualizing malicious activities using honeypots, a project that was started by security professional J. Blasco resulted in the creation of a visualization tool for the Nepenthes honeypot [14]. Nepenthes is a malware honeypot; a utility to assist malware researchers in the process of gathering and securely storing infectious binaries of malicious software. The aforementioned visualization tool [15] uses the AfterGlow and Graphviz software libraries in order to create several directed graphs. These depict the correlations between IP addresses, malware samples and geographical data.

Another honeypot-related visualization tool is called carniwwwhore [16] and was developed for the Dionaea malware honeypot [17]. Dionaea is considered to be the successor of Nepenthes honeypot and the aforementioned visualization tool has similar capabilities in comparison to our tool that was developed during the course of this work, although each employ different technologies.

III. MY IDEA

Looking at this different types of attack these are growing in number every day and different types of attack come into existent Every day about thousands of new attack are identified. Therefore we are going to implement a combination of signature based and host based Active intrusion detection system . We place signature based intrusion detection system at router and anomaly intrusion detection system at intrusion detection system manager . This is our new approach . This has some pros aand cons . The pros are This gives better security , Reliability as one gets failed other work, And cons are performance gets low , more databases are required to manage.

IV. DETAILS

There are two main aims for this project: the first is to build an SSH honeypot and the second is to research cyber-attack techniques. The first aim, building an SSH honeypot, will allow users interested in cyber-security (or information security) to easily and quickly deploy a medium-interaction SSH honeypot and be able to analyse the data produced from this to determine current threat levels. The specific objectives for building an SSH honeypot are:
Implement medium-interaction, research honeypot in programming language C

- Log all username and password attempts
- Allow user authentication and log all attempted commands
- Allow most common set of shell commands to emulate (e.g. ls, wget, w, ...)
- Allow attackers to upload files to honeypot using wget
- Emulate running of uploaded files from attackers

The second aim, research cyber-attack techniques, will analyse current threats (with data produced from the deployment of created honeypot) and help those interested in information security and those wanting to tighten existing SSH systems. Specific objectives for researching cyber-attack techniques are:

- Deploy honeypot to public server so anyone in the world can attack it
- Produce data from honeypot deployment
- Analyse most commonly used username and passwords
- Analyse how username and password lists are created (dictionary, other lists etc)
- Analyse most commonly attempted shell commands
- Analyse uploaded files from attackers

Analyse how host (honeypot) is used by attackers once compromised (such as building a medium- interaction honeypot and analyzing cyber-attack techniques) .And by using these cyber attack techniques and other record we build signature based Intrusion detection system.

A. Technical Details:

There are two phases-

1) Building the honeypot
2) Building the data gathering website

1) Building The Honeypot:

a) Basic Server :

The first objective for building the honeypot was to create a basic server written in C. This basic server will allow clients to connect to the server, send text to the server and then the server will echo the text back to the client.

One of the main advantages of building a basic server as a first step is that it introduces the programming language C. Since the echo server can also be created in an already familiar language such as Java, this makes aids the process of learning a new programming language.

b) Creating A Secure Connection:

The next step is to turn the standard connection into a secure one and to begin implementing the SSH protocol

c) Ssh Protocol Implementation:

The main SSH RFC protocol Introduction, explains that there are three major components to the SSH protocol:

- The Transport Layer Protocol [SSH-TRANS]
- TheUser Authentication Protocol [SSH- USERAUTH]
- The Connection Protocol [SSH-CONNECT]

Connection Setup describes how a new SSH connection is initially setup. The RFC reveals that an SSH connection is established over a standard TCP connection (and not a secure TLS connection).

d) Authorising Ssh Clients:

Once the SSH library had been implemented and a basic SSH connection had been established the next step was to implement user authentication by requiring clients to enter a username and password. The purpose of this authentication is to produce the appearance to an attacker that this honeypot requires a username and password to login to the system.

2) Data Gathering Server:

This section of the build stage involves creating the data gathering server which will act as a website that serves a number of purposes:

- Receive authentication and CLI data from honeypot
- Store received data into a database • Process data from database to produce statistics and graphs
- Provide a simple blogging platform to share project progress with wider cybersecurity community. database can be queried for the most frequent

a) Database Design:

- A MySQL database was used for this project to store honeypot data since it is a free and open source database and is easy to install on a Linux server running Apache web server.
- The database table ssh login attempts is used to store the data collected from the honeypot. It stores the timestamp the authentication attempt was made, the username and password used, the IP address of the client and the IP address of the honeypot (since 18 there will be multiple honeypots deployed). These values correspond to the table columns: date, username, password, client ip and honeypot ip.
- The database table ssh shell log stores the commands entered by authenticated attackers on the emulated CLI. This table records the command, IP address of the client, IP address of the honeypot and the timestamp that the command was entered into the honeypot

b) Data Collection:

The main PHP script is executed remotely by the honeypot which sends a POST HTTP request to the PHP script. The code shown below uses data sanitisation to ensure data being saved in the database cannot contain malicious code. This is achieve by using the method mysqli real escape string which is built into the mysql library. The IP addresses of the honeypots (in the below code) have been masked along with the database username and password. This will ensure that the project can continue to run in the future without compromise after this dissertation has been published.

c) Data Processing & Presentation:

One of the main purposes of the website is to display data collected from the honeypot in a clear and concise manner in order to to share results with the wider cyber-security community. This objective was achieved by querying the MySQL database for various statistics. One of the main honeypot statistics addresses what the most popular usernames and passwords are. The code below shows how the passwords of the day.

For Eg:

```
SELECT password, COUNT(password) AS
passCount
FROM ssh_login_attempts
WHERE DATE(date) = DATE(NOW())
GROUP BY password
ORDER BY COUNT(password) DESC
LIMIT 10
```

This query can also be adjusted to display the most frequent passwords of all time by simply removing the line WHERE DATE(date) = DATE(NOW()). This query would produce a set of results similar to table

id	password	passCount
0	123445	1045
1	changeme	754
2	password	367

V. FUTURE WORK

Our future work will consists of researching the cyber attack techniques which the attackers used and also to provide a result of the honeypot deployment which is a visualization tolls by charts and graphs showing the most of the common passwords and commands used. Also it will track the ip address of the attackers and show the number of times the attackers attacked with logged username and paswords

Furthermore, many such honeypot appliances can be spread across one large or many different networks, while a centralized database could collect all the activity from each one of them. The data logged from an operation of this sort, especially the source IP addresses responsible for the attacks, could be used in the creation of a blacklist database. This in turn could be used in conjunction with software implementing countermeasures, in order to block malicious connections before they are processed by the SSH service itself.

VI. CONCLUSION

Honeypots present a unique security concept and they are a powerful technology. In this paper we presented a practical implementation of a specialized honeypot for the Secure Shell (SSH) service, which is commonly targeted by attackers. Many Linux distributions install an SSH server by default, many times without proper security in place or additional protection mechanisms such as firewalls. Thus, otherwise fully patched and updated systems can be compromised by malicious users, dueto a carelessly chosen password.

In this paper we have run a relative lengthy experiment using the aforementioned SSH honeypot system, which yielded interesting results. It was shown that common credential dictionaries exist and they are exchanged between malicious users for repeated use in SSH attacks. Intruders seem to have in place specific tools they download and utilize immediately, aiding them mostly in conducting further attacks against other servers. Thus, the compromised systems are used as pivots; launch pads for more intrusion attempts. In such cases the affected organizations run the additional risk of law accusations for liability and damages, which is another important factor to consider when designing network and security architectures. System administrators should also use password-checking tools to ensure the security of their chosen credentials and implement password security policies for the users of their systems.

Lastly, a visualization tool was presented for use with the specific honeypot software fielded in our study. This type of utility aids in the graphical presentation of a honeypot's operation and can give a quick and detailed overview of the activity to information security professionals and researchers.

REFERENCES

[1] L. Spitzner, "Honeypots: Catching the Insider Threat," in Proceedings of the 19th Annual Computer Security Applications Conference, 2003.
[2] L. Spitzner, Honeypots: Tracking Hackers. Boston, MA: Addison Wesley, 2003.
[3] L. Spitzner, "Strategies and issues: Honeypots - sticking it to hackers," Network Magazine, 2003.
[4] A. Obied, "Honeypots and Spam." 2007.
[5] D. Keim, F. Mansmann, J. Schneidewind, J. Thomas, and H. Ziegler, "Visual analytics: Scope and challenges," Visual Data Mining, pp. 76- 90, 2008.
[6] E. Alata, V. Nicomette, M. Kaaniche, M. Dacier, and M. Herrb, "Lessons learned from the deployment of a high-interaction honeypot," in Proc. Dependable Computing Conference (EDCC06), 2006, pp. 39- 46.
[7] D. Ramsbrock, R. Berthier, and M. Cuckier, "Profiling Attacker Behavior Following SSH Compromises," in Proceedings of the 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, 2007, pp. 119-124.
[8] J. Owens and J. Matthews, "A Study of Passwords and Methods Used in Brute-Force SSH Attacks." 2008.
[9] "Observations of Login Activity in an SSH Honeypot," Cisco Security Intelligence Operations, 2009. [Online]. Available: https://www.cisco.com/web/about/security/intell igence/ssh-security.html.
[10] J. C. Klein Keane, "Using Kojoney Open Source Low Interaction Honeypot to Develop Defensive Strategies and Fingerprint Post Compromise Attacker Behavior," HITB Magazine, Volume 1, Issue 3, pp. 4-14, 2010.
[11] C. Seifert, "Analyzing Malicious SSH Login Attempts," Security Focus, Infocus 1876, 2006. [Online]. Available: http://www.securityfocus.com/infocus/1876.
[12] F. Fisher, F. Mansmann, D. A. Keim, S. Pietzko, and M. Waldvogel, "Large-scale network monitoring for visual analysis of attacks." 2008.

[13] J. R. Goodall and M. Sowul, "VIAssist: Visual analytics for cyber defense," in IEEE Conference on Technologies for Homeland Security, HST'09, 2009, pp. 143-150.
[14] P. Baecher, M. Koetter, T. Holz, M. Dornseif, and F. Freiling, "The Nepenthes Platform: An Efficient Approach to Collect Malware." 2006.
[15] J. Blasco, "An approach to malware collection log visualization." 2008.
[16] "carniwwwhore." [Online]. Available: http://carnivore.it/2010/11/27/carniwwwhore.
[17] "Dionaea honeypot." [Online]. Available: http://dionaea.carnivore.it/.