Research and Implementation of a Network Secure System Based on Honeypots

Bingqing Luo[a], Zhixin Sun[b]
College of Internet of Things
Nanjing University of Posts and Telecommunications line
Nanjing, China
[a]ice820@126.com, [b]sunzx@njupt.edu.cn

***Abstract*—Aiming at the shortcomings of the traditional honeypot on massive global scanning attact, this paper constructs a novel network secure system based on honeypots (NSSH). Afterwards, this paper discusses the classic architecture of NSSH, which consists of four parts, firewall/router, analysis network, honeypot and monitor network. The paper also gives out analysis and implementation in detail of the key techniques of NSSH, including the honeypot model and audit analysis based on statistics. Our simulation experiments indicate that comparing to the traditional secure system based on honeypot, NSSH can make network's security performance to a high level and NSSH has important impact on forecast and monitor attack activities.**

Keywords-Honeypot; Trap system; Network security; Massive scanning; Audit analysis.

I. INTRODUCTION

A Potential network safety problems become outstanding, virus and hacker puzzle the Internet users almost everyday. Yet, despite decades of research and experience, we are still unable to make secure computer systems or even measure their security[1].

As a result, exploitation of newly discovered vulnerabilities often catches us by surprise. Exploit automation and massive global scanning for vulnerabilities enable adversaries to compromise computer systems shortly after vulnerabilities become known. One way to get early warnings of new vulnerabilities is to install and monitor computer systems on a network that we expect to be broken into. Every attempt to contact these systems via the network is suspect. We call such a system a honeypot [2]. A virtual Honeypot Framework [3] describes the design and implementation of Honeyd, a framework for virtual honeypots that simulates computer systems at the network level. Literature [4] propose that a honeynet can be used to assist the system administrator in identifying malicious traffic on the enterprise network. Literature [5] applies the honeypot concept to the email system in order to proactively detect bad providers where the honey consists of unique email address. Paper [6] presents a system that helps in the defence in depth of a network from distributed denial-of-services(DDOS) attacks, and he propose a honeypot for such attacks. However, the honeypot does not deal with complex setup and does not discuss data capture and data control techniques which our system has made clearly.

This paper constructs a novel Network secure system based on honeypots(NSSH), and presents the architecture and the key technology of NSSH. The paper gives out the analysis and implementation of NSSH in detail. Our simulation experiments indicate that NSSH can make network's security performance to a high level and NSSH has important impact on forecast and monitor attack activities.

II. A NETWORK SECURE SYSTEM BASED ON HONEYPOTS (NSSH)

We propose that using a honeypot provide an additional layer of network security. The honeypot can serve as a compliment to the use of the firewall and Intrusion Detect System (IDS) and help to overcome some of the shortcomings that are inherent to these systems. We have constructed a trap network successfully in NSSH, which can do better in protecting our real network. The trap system that we have constructed is shown in Fig.1. It includes four parts, i.e., firewall/router, honeypot , monitor network and remote analysis network.

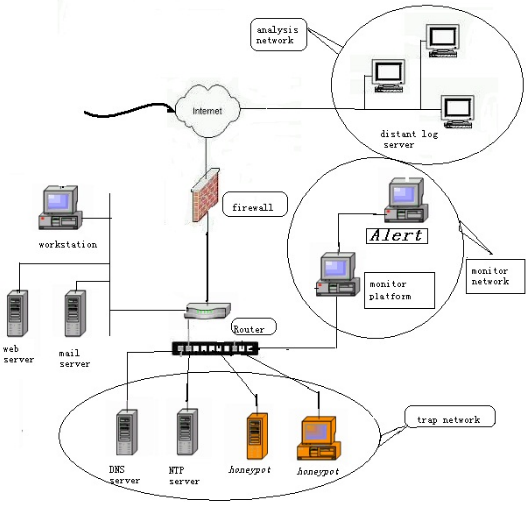


Fig.1. Example of a ONE-COLUMN figure caption.

Firewall controls the network traffic and restricts the connection to the outside. However it should offer hackers definite freedoms, such as obtaining tools, testing connection etc, except that, the router also controls the traffic, and the existence of the router makes NSSH much more real. In addition, the router is a backup for data control, and all traffic must pass through it. If the firewall fails to detect, the router will still execute the main functions of the firewall. It is convenient to network manager to manage network better through accessing control list (ALC).

The honeypot is the core of NSSH, which is a kind of network to lure hackers to attack it. In the trap network, all operating systems are real, so there must be vulnerabilities that are easy to trap hacker to scan and attack. DNS is the necessary function, through which hackers often analyze or download tools. In order to endue honeypot with completed domain analysis performance, we add DNS server, which makes domain analysis from the honeypot to restrict to certain special host computer. NTP（network time protocol）server will guarantee that all time in trap network is synchronous. It is very useful to data analysis, which can guarantee the data captured from each system to be at the same time.

The monitor network monitors and controls all data in and out honeypot, however it can't communicate with honeypot (we adopt the method of not assigning IP address to it). Therefore it can guarantee that hackers will not destroy the obtained data. Where, Intrusion Detect System (IDS) captures data packages, and analyzes abnormal behaviors, then sends out alarms.

The analysis network carries out backup on system log and firewall log. The system log is important to NSSH and it is easy to be destroyed by hackers, so we deploy a remote server to save log files. Meanwhile, the remote log server has a concealment performance; it is also a complex honeypot. When hackers attack into a system that the safe grade is lower in honeypot, they may be aware that the system log has been sent to remote log server. In this case, hackers will try to attack this remote server and delete its log. If hackers focus their objects on the remote log server, we will learn more attack methods. Even though the remote log server has been broken into and log files have been deleted, NSSH will not lose any content. It is because the monitor network used as capturing all message will capture all log files that have been sent to the remote system log server. In the process of log analysis, NSSH adopts the statistics method to analyze and judge unknown attack actions, which is an important supplementary to IDS.

III. The Construction of the Honeypots

The realized honeypot model in NSSH is shown in Fig.2. Data control is the control of data passing in and out of the Honeypot. It is important that what kind of data can reach and what kind of destination is decided and controlled by network manager. The key of data control is visiting control equipment, such as firewall, which can separate Honeypot from the other parts in Internet[7]. All data passing in and out on the Honeypot must pass through firewall firstly. We employ transparent firewall over here, so hackers will not be aware that they are passing through a firewall system. For visiting control equipment, there must possess following rules [8]: (1) Any people can launch the connection from Internet to the Honeypot. In this way, it can allow hackers to scan, detect and attack the system in Honeypot finally. (2) The visiting control equipment controls the honeypot and can launch the way connected with Internet. For it can prevent hackers from employing the Honeypot to attack or destroy other systems in Internet. Once certain honeypot is broken into, it must accept activities of hackers. The mentioned activities indicate external connections allowed from the Honeypot. If we forbid all connections sending to Internet from the Honeypot, NSSH can decrease a majority of risks, however, this method is impractical. Once hackers break into a honeypot , and can't launch connections to the Internet, they will suspect on that quickly. They will leave this system even destroy the whole system before leaving, so there is no any value for the honeypot. Of course, visiting control equipment also can't allow too many external connections; otherwise, the attacked system is likely to be employed detecting or attacking other system in Internet.

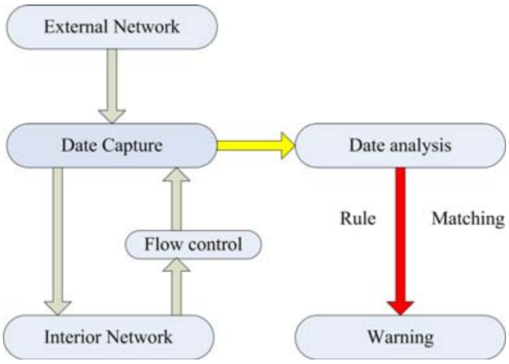


Fig.2. Honeypot model in NSSH

The captured data records all activities that hackers have done in Honeypot, which is just the value of Honeypot, for only the captured data can be analyzed and researched. Therefore, the right and correct capture is the key to the success of the whole trap system. We must consider selecting believable data capture sites and at the same time they situate in different place in NSSH. The captured data also need to be saved on believable network or hosts. For example, if activities of hackers are saved in inner of the honeypot, then hackers may likely detect this part of data and modify or destroy them. The backup of log files is very important in the honeypot. These files can be saved on a remote protected server in order to analyze conveniently. For ensuring the integrality of data, we should adopt encrypt and authentication techniques in log data transmission process.

IV. Abnormal analysis based on statistic in NSSH

Log analysis is always a heavy task to network manager [9]. However, it is an important reference to analyze a system abnormity and hackers' behavior. In log analysis, many certain phenomena obey calculation rules as follows. According to a system's state or behavior at the time t0, we can decide the state of the system or behavior at the time t>t0, and do not need depending on the process and state before t0. For the process of log analysis analogy to the rule in statistic system, the Markov process can be extended to log analysis.

Markov process is a widely used random process, and it has been used in many fields [10]. We adopt Markov process model in log analysis, that is to say, every type log is defined as a corresponding system state, and we can express the change of state by state-transition matrix. If a log transfer probability is less than others in log state-transition matrix, then this log behavior may be abnormal behavior. In the process of log analysis, normal behaviors are in the majority, so transfer probability from normal behavior to normal behavior is in the majority, and transfer probability of abnormal behavior is less. Now we show an analysis on an intrusion trap example in NSSH.

In this example, we divide the log into abnormity and normal. Using 1 expresses normal, and 0 expresses abnormity. In a log file, we get a data sequence as follows:

1110010011111110011110111111001111111110001
1011011110110110101111011101111011111100110111
11100111

Suppose *Xn* is the *n*（n=1,2,…,97）behavior state, which can be supposed as a homogeneous Markov chain, state space I={0,1}. Instances of the 96 times state transfers are:

0→0,8 times; 0→1,18 times; 1→0,18 times; 1→1,52 times.

Where, the transfer probability can be shown as frequency approximately.

P00=P{Xn+1=0|Xn=0}≈8/(8+18) = 8/26,

P01=P{Xn+1=1|Xn=0}≈18/(8+18) = 18/26,

P10=P{Xn+1=0|Xn=1}≈18/(52+18) =1 8/70,

P11=P{Xn+1=1|Xn=1}≈52/(52+18) = 52/70.

According to Markov process model, from above results, it is easy for us to know that 0 is abnormity, which accords with the actual conditions. The example indicates that Markov process can enhance the efficiency of log analysis, and reduce the workload of network managers. The application of this model can find a system's abnormity behaviors quickly, and dispose them in time.

V. Application

In order to validate this model, we carry through a simulation experiment. We apply NSSH to a real network. We lay some honeypots in the trap network. In honeypot, we install Apache web server of Linux, and Solaris ftp server of Sun. We employ default install on these operating systems, so they have some ordinary system leaks. The trap network connects with our real network through a router. We install an IDS on the monitor platform. For the sake of safety, we do not allocate IP addresses to them. The monitor platform and warning hosts compose a local network, and they can communicate each other freely.

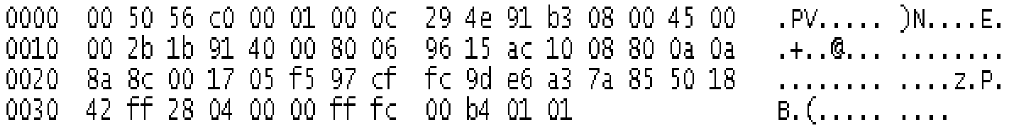
In order to restrict the traffic, we deploy a firewall. The TABLE 1 is configuration of the firewall. We restrict a honeypot to connect with outside network, and the restriction to other honeypots is analogous.

TABLE 1. CONFIGURATION TABLE OF THE FIREWALL

Source IP	Object IP	protocol	Connected number /time
Any	10.10.138.140	Any	absoluteness
10.10.138.140	Any	TCP	20/m
10.10.138.140	Any	UDP	15/m
10.10.138.140	Any	ICMP	50/m
10.10.138.140	Any	Other	20/m

Where, 10.10.138.140 is the IP address of honeypot; Any expresses any IP address or protocol; Other can be a ticket according to different network state.

In the monitor platform, we employ IDS to capture data packages. We have captured a package contents as show in Fig.3.



```
0000  00 50 56 c0 00 01 00 0c  29 4e 91 b3 08 00 45 00   .PV..... )N....E.
0010  00 2b 1b 91 40 00 80 06  96 15 ac 10 08 80 0a 0a   .+..@... ........
0020  8a 8c 00 17 05 f5 97 cf  fc 9d e6 a3 7a 85 50 18   ........ ....z.P.
0030  42 ff 28 04 00 00 ff fc  00 b4 01 01               B.(..... ....
```

Fig.3. The captured package contents

We make a simple analysis. 00 50 56 c0 00 01 expresses object MAC address is 00:50: 56 :c0: 00: 01. 00 0c 29 4e 91 b3 expresses source MAC address is 00: 0c: 29: 4e: 91: b3. 08 00 shows network layer protocol is IP protocol. 45 00 expresses IPV4 protocol, i.e. the IP version number. 00 2b shows the total length is 43 bytes, and 1b 91 is a sign. 80 expresses the value of TTL is 128. 06 shows transmission layer protocol is TCP, and 96 15 is check sum. ac 10 08 80 is source IP address (172.16.8.128), and 0a 0a 8a 8c is object IP address (10.10.138.140). 00 17 is the port number, which is 23, i.e. telnet protocol, and 05f5 is object port number. 97 cf fc 9d expresses the sequence number of this package is 2546990237. e6 a3 7a 85 shows the ACK number is 3869473413, and 50 expresses the head length is 20 bytes. 18 expresses having set the ACK and PUSH bit. 42 ff shows the dimension of the window is 17151, and 28 04 is the check sum. ff fc 00 is a control order of telnet, and it means not processing binary system transmission. b4 01 01 is the end symbol of the message. We can see from the above analysis, the host whose IP address is 172.16.8.128 tried to enter a honeypot in NSSH (IP address is 10.10.138.140) by telnet.

Repeated scans across NSSH for a specific port indicate that an infected machine may be looking for a vulnerability within our campus network. We analyze the data collected by the IDS session in the data capture mode to look at the time lapse between the various port scans that occur across NSSH.We speculate that scans that have occurred in under one second across numerous systems on the trap network are most likely automated worm type exploits. We can inform the network managers of these systems that are suspected of being exploited.

VI. Conclusions

NSSH is employed to obtain hacker group's wisdom, using tools, means and strategy, and their motives. NSSH hasn't simulated anything; however it is composed of some real operating systems and application programs. In data analysis, we adopt a new technique of audit data statistic, which can analyze data with higher efficiency. Once the system is attacked, the system can not only tell us the operation means of hackers, but also identify its existent risks and vulnerabilities.

Of course, NSSH is still in its infancy, we achieved first promising results with the presented initial setup. Further work will construct more intelligent honeynet to lure more brilliant hackers for our IDS to learn higher-level attack means in order to enhance our network system security. At the same time, we will research on more intelligent decoy route and gateway algorithm in honeynet, and make the trap network more intelligent.

VII. Acknowledgment

This work is supported by the National Natural Science Foundation of China (60973140, 61170276, 61373135), The research project of Jiangsu Province (BY2013011), The Jiangsu provincial science and technology enterprises innovation fund project (BC2013027), the High-level Personnel Project Funding of Jiangsu Province Six Talents Peak and Jiangsu Province Blue Engineering Project. The major project of Jiangsu Province University Natural Science (12KJA520003)

References

[1] Zeng, W. and M.-Y. Chow:*Optimal Tradeoff Between Performance and Security in Networked Control Systems Based on Coevolutionary Algorithms.* Ieee Transactions on Industrial Electronics,vol. **59**,(2012), p. 3016-3025.

[2] Zhang, F., et al.:*Honeypot: a supplemented active defense system for network security.*in: Parallel and Distributed Computing, Applications and Technologies, Pdcat'2003, edtied by P. Fan and H. Shen. (2003). 231-235.

[3] Zhan, Z., M. Xu, and S. Xu: *Characterizing Honeypot-Captured Cyber Attacks: Statistical Framework and Case Study.* Ieee Transactions on Information Forensics and Security, Vol. **8**(11), (2013), p. 1775-1789.

[4] Lin, F.Y.-S., Y.-S. Wang, and M.-Y. Huang: *Effective Proactive and Reactive Defense Strategies against Malicious Attacks in a Virtualized Honeynet.* Journal of Applied Mathematics, (2013).

[5] Sun, X., et al.: *Collecting Internet Malware Based on Client-side Honeypot*. in: Proceedings of the 9th International Conference for Young Computer Scientists, Vol. 1-5, edtied by G.J. Wang, et al.Hunan, China (2008). 1493-1498.

[6] Sardana, A., et al.:*Deciding optimal entropic thresholds to calibrate the detection mechanism for variable rate DDoS attacks in ISP domain: honeypot based approach.* Journal of Intelligent Manufacturing, Vol. **21**(5), (2010), p. 623-634.

[7] Li, L., H. Sun, and Z. Zhang: *The research and design of honeypot system applied in the LAN security*. in : *2011 IEEE 2nd International Conference on Software Engineering and Service Science, ICSESS 2011*. IEEE Publising, Beijing, China (2011).

[8] Mansoori, M., O. Zakaria, and A. Gani: *Improving Exposure of Intrusion Deception System through Implementation of Hybrid Honeypot.* International Arab Journal of Information Technology,Vol. **9**(5), (2012), p. 436-444.

[9] Yang, Y., H. Yang, and J. Mi: *Design of distributed honeypot system based on intrusion tracking*. in: *2011 IEEE 3rd International Conference on Communication Software and Networks, ICCSN 2011,*IEEE Publising, Xi'an, China (2011).

[10] Gao, B., et al.: *Page importance computation based on Markov processes.* Information Retrieval, Vol.**14**(5), (2011), p. 488-514.