Detecting Backdoor Using Stepping Stone Detection Approach

Khalid Alminshid
Department of Information Technology
School of Computing
Universiti Utara Malaysia
Kedah, Malaysia
khalid_uum@yahoo.com

Mohd Nizam Omar
InterNetWorks Research Laboratory
School of Computing
Universiti Utara Malaysia
Kedah, Malaysia
niezam@uum.edu.my

***Abstract*—Several techniques are used by intruders to hide the track of intrusion in the network. One of these techniques executes by using series of hosts in network (stepping stones chain), which can be detected by using an approach, called Stepping Stone Detection (SSD). However, during all previous years, SSD was only confined to detect this type of intrusion. This paper discusses the using of SSD approach and potential applications in other emerging field by introduce the using of SSD concepts in backdoor attack detection field. This research shows that by using SSD to detect backdoor attack can be gained very low false negative and false positive rates and reduces the scan process time detection.**

Keywords—Stepping stone; stepping stone detection; backdoor; RTT

I. INTRODUCTION

The rapid growth in technology has changed the landscape of networks, in the same way, networks have changed the shape of business, working of organizations and offered a new high range of possibilities and flexibility, it is clear that, we cannot dispense with the using of these networks in our daily life, but on the other hand that appeared a new face of security threats.

Security threats are coming from various sources, like natural factors or users do not have authorization to access to the network and using the software. These threats are result to the lack of authentication, authorization and auditing control, cipher algorithms weaknesses, weak keys, the risk partnership, untrustworthy data center and disaster recovering failure. According to the last Microsoft's Security Intelligence Report [1], in over the last decade, the spread of malware became the crime story online. Nowadays estimates of the number of well known threats, such including viruses, worms, trojans, backdoors, exploits, password stealers and spyware in the millions, and a backdoor has a high rate of intrusions that happens to global networks in the world.

The backdoor attack is one of the threats, which caused serious concern. It was a major threat in recent years, as the outbound generated by backdoor attacks have several types of packages and gets dangerous control a range of host [2]. Therefore, detecting backdoors become a basic need nowadays.

There are many techniques suggested to detect backdoors attacks, in general these techniques involved in two mainly techniques: Signature based technique and Behavior based technique [3]. Most of antivirus systems and Intrusion Detection Systems (IDS) are signature based [4]. In all this type detectors, a sequence of features unique to a backdoor are used in order to detect backdoor. Behavior-based method focuses on analyzing the malicious behavior. Such behaviors include addresses of the backdoor destination and source [5]. Signature-based technique has less scanning time and few false positives. However, unknown backdoors can easily evade the detection. In addition, this technique has not ability to deal with obfuscation [6]. In the same way, behavior-based technique has best results in detecting of polymorphic backdoors, but cannot detect a lot of polymorphic backdoors present environment [7]. Moreover, it has two limitations. Firstly it has a high false alarm rate. Secondly, the complexity involved in the training phase to determine what features should be learned [8]. Given these facts, several researchers began to look at a new technique to detect backdoor attack. One of the successful techniques in recent years, which have given great outcome in network security field, is Stepping Stone Detection Techniques (SSD).

SSD approach is quite responsible for detecting the interactive connections [9]. Addition to that, the recent explorations indicate that SSD approach is quite flexible. It is possible to extend the theory of SSD to other emerging fields such spam, backdoor and proxy detection [10]. Consequently, this paper considers the issue of detecting backdoors attacks by using a simpler solution based on the concepts taken from SSD based-research.

ISBN: 978-1-4673-5256-7/13/$31.00 ©2013 IEEE

II. TERMINOLOGY

Firstly, we have to present some definitions that will be used in this proposed research. In this section there are two parts, backdoor terms and SSD terms.

A. Backdoor Terminology

Firstly, we present definitions that will be related to backdoor attack. Malware: Malware word comes from two words, malicious and software. It is a program that is designed to be harmful. Backdoor: A hidden technique is used for getting remote access to a machine or other system that without authentication Intrusion: An illegal act of entering to a computer or network or any system.

B. SSD Terminology

Also we have to present definitions that will be related to SSD approach. Assume attacker logs in from the first Host (1), and ultimately connects to the destination Host n, through Host i-1 … , and the Host *i*+1. Fig. 1, shows the stepping stone chain.

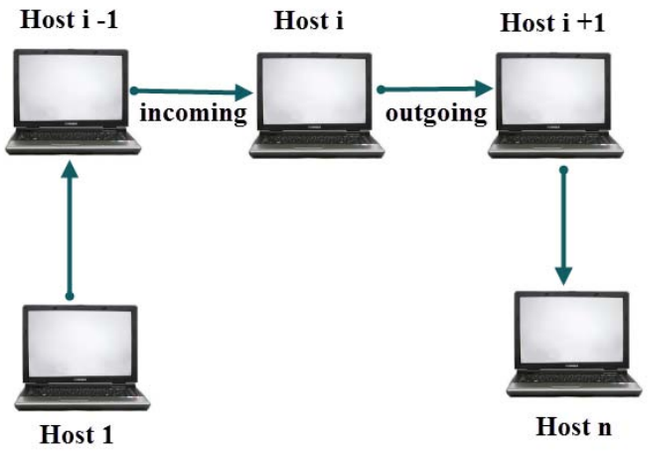

Fig. 1. Stepping Stone Chain.

Connection: The connection of user logging from host to next host is called connection session between the two hosts. Chain: Given n hosts H_1 to H_n, a series (chain) of connections is defined as a chain $C = <C_1, C_2, .., C_{n-1}>$ terms C_i is a chain of connection between Host H_i and Host H_{i+1} for a series i = 1 to n-1. Stepping-Stone: The intermediary host of a connection chain which is invaded by the attacker. Downstream and upstream: If a direction is along a user's login direction from attacker to victim (as shown in the arrows in Fig. 1), it is called downstream. Otherwise, it is called upstream. False positive: False positive means when the detector detects a backdoor in a file does not have backdoor. False negative: False negative means when the detector does not detect the backdoor in a file has infected by backdoor.

III. STEPPING STONE

There are two main concepts are involving in the name of stepping stone, the first is stepping stones attack (chain) and the second is Stepping Stone Detection approach (SSD). In fact nowadays, most of the attackers are using a chain of intermediate points to get attack in final destination, computer or any host. These hosts or points from attacker to his final destination are called a stepping stones (chain). The attacker is using this technique to avoid the detection by the intrusion detection system. While SSD approach is a system to analyze the traffic of the connection and identify which connections are stepping stone connections or identify which connection pair are correlated connections. Correlated connections are a pair of connections which are in the same way of connection chain.

SSD-based research was mostly limited to only field of detection stepping stones and all of the researchers focus to the main usage of SSD to detect stepping stone either in host or network-based environment [10]. No such a research that realized the other usage of the SSD in other fields of research without looking to the full capabilities of stepping stone detection in other fields of research. The study [10] suggested using the potential applications of SSD in other fields such as Spam Detection, Proxy Server Detection and Backdoor Detection. This study is emerging to this suggestion, to use the SSD concept in the detection backdoor problem.

A. SSD Models

Based on the location where the analysis takes place, SSD are divided into two types, which are: Network based SSD, (NSSD) and Host based SSD, (HSSD). Each type has a special technique for securing information and monitoring, and each type involved special pros and cons. HSSD is called on SSD when the management unit on each host. HSSD monitor traffic on its workstation by utilizing the host resources to find any intrusion attack. In (NSSD) model the management unit as stand-alone devices on all components of network. NSSD checks traffic on the network to detect intrusion attacks. Sometime, combine both NSSD and HSSD to identify the attacks. In this type, both kinds of SSD can be used simultaneously. Hybrid SSD based is called to this type.

According to [3] and [11], For network-based, it helps only to detect the external intrusions furthermore, it is not easy to detect intrusion from encrypted traffic in addition, it is not easy to detect network intrusion in a virtual network. Therefore, network-based is not suitable for detecting backdoors. On other hand Hybrid SSD based is difficult to understand, in other words, it is complex. While for HSSD, it is able to detect intrusion by monitoring the file system of the, network events or the system calls. Furthermore, it does not need to extra hardware. Therefore, the HSSD is the best architecture to detect backdoor attacks.

B. SSD Techniques

Many techniques have been proposed to solve stepping stones problem. Generally, there are four main techniques used widely to detect a stepping stone problem, there are:

Round Trip Time (RTT) based, Timing based, Deviation based and Packet number based. However for Packet Number based is not enough in resisting the intruders while, for Timing based, it is not difficult to manipulate it by intruders [12], "backdoor is an intruder". For Deviation-Based, it has all the problems of the Time-Based Approach [13]. The fourth used technique is RTT. This technique uses both Send and Echo packets for detecting stepping stones attack. Furthermore, it has the ability to filter chaff packets and unsymmetrical Internet packets and, it is more resistant to intruder evasion and network imperfections than any other type of SSD techniques [9].

The time needed for a packet to travel between the departure source and the final destination host and return again is called, Round Trip Time (RTT). In this concept, the (exporter) departure source is the host that sending the packet and the destination is the host or the system that receiving the packet and retransmits it again. In fact, RTT is one of several factors affecting latency in the network. The RTT can range from a few milliseconds of time (thousandths of a second) under perfect conditions between very closely spaced two nodes to several seconds under negative conditions between two nodes separated by a long distance.

[13] proposed a new algorithm to find RTT in the real time and named it, Estimation Based Algorithm (EBA).The study explained, that we can use this algorithm at all times by RTT based SSD approaches, such as "Step-Function" study [14]. The study showed that EBA can also be used sparingly to find the value of parameters by other non-RTT based SSD approaches. The experiment for this study justified that EBA is far more accurate than other algorithms which used for getting RTT. Furthermore the study justified that EBA has a high matching-rate similar to the SDBA approach [15]. Given these facts, in addition to sum up from above all, RTT is chosen as a suitable technique for this study and EBA as algorithm to achieve RTT in real-time, Therefore we can use it in detect backdoor by using SSD approach based on [10] suggestion.

IV. BACKDOOR & SSD

In 2000, a general algorithm was proposed by Zhang and Paxson [16], for detecting the interactive of the backdoor, this algorithm based on three kinds of characteristics: directionality, the time of interarrival packet and the size packet. The main idea was, firstly by looking for small packets (because most backdoors commands are short, the size packet they used "20 bytes" to define "small" packets) which have large interval time (fall between 10 *m*Sec and 2 Sec). Then by looking for frequent for small size packets. For directionality key, they only consider the traffic that sent by the starter of a connection. In other words, this algorithm reflects, the risk of attack is coming with any flows consists of less than 8 packets or less than 2 seconds where a flow has one direction. This tool of the detection algorithms was clear and convincing.

Consequently, SSD has untapped potential in several emerging research fields, namely in backdoor detection. The study [10] demonstrated the potential of SSD in addressing current issue in backdoor detection by a novel SSD. The study proposed a simpler solution to detect backdoor by using concepts from SSD based-research. The study explained that backdoor detection using the same concept of SSD. Therefore, to detect the origin of the backdoor and a host of backdoor can be detected in similar way to a SSD host, when the incoming and outgoing flow through the host is matched. In the backdoor situation, the detection gets when a connection occurs repeated for a specific port. It usually happens when the affected host suddenly forwards a data to the outside network using the same port number and at the same period of time when the backdoor affects a number of hosts (used as a stepping stones).

V. METHODOLOGY

The operational framework of this study is divided into three sections. The first section (preliminary framework) includes finding the suitable technique that can be used in SSD approach to detect backdoor attack problem. By analyzing the characteristics of each technique depending on literature review and related works, this question has been answered and the preliminary study, RTT is the suitable technique. The second section of framework includes how we use the technique which is taken from the first part in SSD based by analyzing related works that already used the same technique in SSD approach. After that we have to use the new approach (SSD & RTT technique) in real environment. The real test will be conducted after that to produce the required results and analysis. Finally in the third section (Evaluation) the result and analysis will be evaluated before the research report is produced. Centrally, the result and analysis will be included the performance measurement for the developed approach. This methodology is quite similar to some of related works methodologies such as [17] and [18]. Most of this research based on experiment. So as well as the related work such as [19], for gathering the data, this study used the following steps:

1) A real backdoor used, manually entered a predetermined sequence of input on clean machines with ten different types of backdoors installed.
2) The proposed approach that used SSD approach (RTT) used to detect all backdoors behavior one by one.
3) Time function is used from the first time the proposed approach is run and stops at detection the backdoor by the proposed approach.
4) TPR and FPR also been captured.
   Then (1) and (2) are repeated by using Antivirus and IDS

For data analyses as well as the related work such as [20], TRP, FPR and Scan Process Time are the main data analysis approaches, TPR is detection rates and FPR is miss-detection rate. The high TPR and the low Scan Process Time (speeding) are proved successfully of the proposed approach. The three variables above are comparing with the same variables for antivirus and IDS to prove successfully of the proposed approach.

For evaluation, all results of this proposed research are digit numbers, so the evaluation, will compare the results of this research with the results of other intrusion detection systems and antivirus by including the three variables (TRP, FPR and Scan Process Time). That means the comparison will include the speed and the accuracy.

VI. EXPERIMENT

As describe in the Section V (Methodology), the experiment is run with four steps in (LAN) as a controlled environment. Different types of backdoor files are gathered from the Internet source such as [21] to use as samples. The experiment has done in two architectures design (Fig. 3, shows both two designs). The first by using LAN without concoction to the Internet by setup the backdoor (attacker) in the host B (192.168.5.45) and the backdoor server in host A (192.168.5.46) and the Source Port: (1177), Destination Port: (52361). Fig. 2, shows the backdoor sample that used in this architecture.

Fig. 2. Backdoor sample that used between two hosts.

While for the second design, we only need to setup the server of backdoor because, the backdoor attacker is already exists on the Internet network. For example, during the experiment the attacker (host) of the backdoor sample (Backdoor.Win32.Hupigon.sbo) is detected in destination host with IP (41.101.68.10) and ISP: (Algerie Telecom – FAWRI). And the Source Port: (62042), Destination Port: (751).

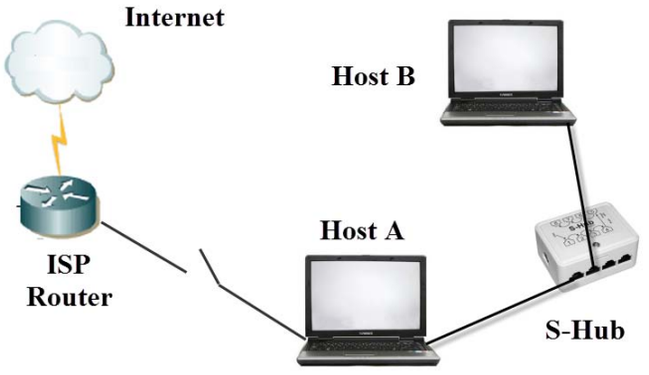

Fig. 3. Network Topology used.

After injection the backdoor to the host, the experiment starts with using of Wireshark V. 1.10.0 [22] to capture the incoming and outgoing network packets on the monitored hosts and to find the RTT for TCP sessions as show in Fig. 4. This operation is needed to less than one minute to export a file from Wireshark Tool that contains all the information that will be used as to provide next processes.

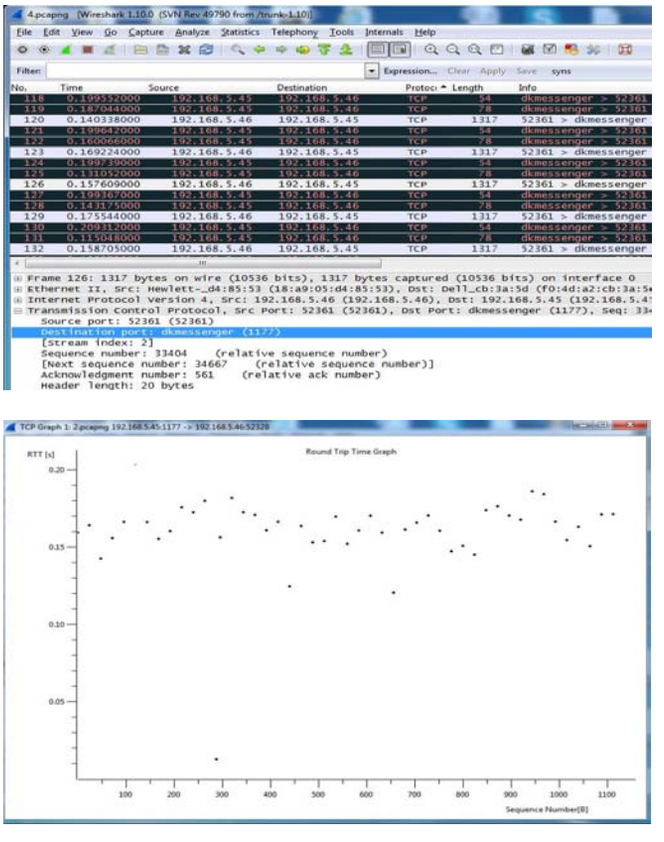

Fig. 4. Wireshark tool to get RTT and, to capture the needed network packets

This file will be transferred into own software that will be written by C++ programming language. This code analyzes the information that obtained from the first step to detect if there is an intrusion by the backdoor or not, based on what has already assumed in this research in Section IV. The detection backdoor attack algorithm in this research as follows:

- The risk of backdoor attack is coming with any flows consists of less than 8 small packets or less than 2 seconds where a flow has one direction.
- When the incoming and outgoing flow through the host is matched.
- When the affected host forwards a data to the outside network using the same port number.

The output of this software will be taken as the result on this study. The comparison is done between the result of this experiment and the results of more than 45 (IDS and antivirus) by using Virustotal website [23].

VII. RESULTS

Preliminary result for this research has ratio of 100% in the detection of the backdoor samples that used. Table 1 shows the results to five samples.

TABLE I. THE RESULTS TO FIVE SAMPLES

Filename (Backdoor)	File size (bytes)	Ratio/Antivirus & IDS Detection	Ratio/SSD Detection
Backdoor.Win32.Cmjspy.b	102709	3/46	1/1
Backdoor.Win32.CyberSpy.84	89286	4 / 46	1/1
Backdoor.Win32.Agent.iq	51498	5/46	1/1
Backdoor.Win32.Hupigon.sbo	299988	3 / 46	1/1
UUM-backdoor.exe	44544	1 / 46	1/1

This method is very fast compared to antivirus systems and IDS. Furthermore, most of these backdoors that detected by this technique are not detected by antivirus systems or IDS as show in Fig. 5. In general, the performance of this technique is quite well and high accuracy. But, the finding many legit backdoors is highly probable, if it has similar behavior and interactive to the real backdoor.

Fig. 5. Detecting backdoors by antivirus systems and IDS

VIII. CONCLUSION AND FUTURE WORK

This paper stated the complexity of IDS, that make researchers are looking for a simple approach to enhance the accuracy and speeding. One of the successful techniques in recent years, which have given great outcome in this field, it is Stepping Stone Detection (SSD). SSD can be used more than detecting stepping stone chain. The advantage from using SSD makes backdoor detection much move faster and this will reduce the time for detection of an intrusion. Also, this research study contributes to the body of knowledge in the domain of research by introducing the usage of SSD-based technique in the backdoor attack detection environment that usually used in SSD-based environment only.

On other hand, this study has been shown why choosing RTT technique as a suitable technique to detect backdoor attacks, and why choosing Host-based architecture. Furthermore, it justified why proposed a new approach, while there are another approaches to detect backdoor such as Signature-based detection approach and Anomaly based detection approach.

In future work, we will improve the performance of this technique by using more samples and by make sure that legit backdoors will not effect on the accuracy. In addition to the using of this technique for detecting other threats such as spam and proxy.

REFERENCES

[1] Microsoft, "Microsoft Security Intelligence Report," Technical Report 2012.

[2] B. Choi and K. Cho, "Detection of Insider Attacks to the Web Server," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA),* vol. 3, pp. 35-45, 2012.

[3] S. Sonawane, G. Prasad, and S. Pardeshi, "A survey on intrusion detection techniques," *World Journal of Science and Technology,* vol. 2, 2012.

[4] B. Kang, H. S. Kim, T. Kim, H. Kwon, and E. G. Im, "Fast malware family detection method using control flow graphs," in *Proceedings of*

the 2011 ACM Symposium on Research in Applied Computation, 2011, pp. 287-292.
[5] C. Modi, D. Patel, B. Borisaniya, H. Patel, A. Patel, and M. Rajarajan, "A survey of intrusion detection techniques in cloud," *Journal of Network and Computer Applications,* 2012.
[6] I. K. Lata, "Study and Analysis of Network based Intrusion Detection System," *International Journal of Advanced Research in Computer and Communication Engineering,* vol. 2, 2013.
[7] M. A. Maarof and A. H. Osman, "Malware Detection Based on Hybrid Signature Behaviour Application Programming Interface Call Graph," *American Journal of Applied Sciences,* vol. 9, 2012.
[8] N. Idika and A. P. Mathur, "A survey of malware detection techniques," *Purdue University,* p. 48, 2007.
[9] L. Ping, Z. Wanlei, and W. Yini, "Getting the Real-Time Precise Round-Trip Time for Stepping Stone Detection," in *Network and System Security (NSS), 2010 4th International Conference on*, 2010, pp. 377-382.
[10] M. N. Omar, A. Amphawan, and R. Din, "Evolution of Stepping Stone Detection and Emerging Applications," *11 WSEAS International Conference on Information Security and Privacy (ISP'12),* 2012.
[11] D. Mudzingwa and R. Agrawal, "A study of methodologies used in intrusion detection and prevention systems (IDPS)," in *Southeastcon, 2012 Proceedings of IEEE*, 2012, pp. 1-6.
[12] J. Yang and B. Lee, "Detecting Stepping-Stone Intrusion and Resisting Evasion through TCP/IP Packets Cross-Matching," in *Autonomic and Trusted Computing*, ed: Springer, 2008, pp. 2-12.
[13] P. Li, W. Zhou, and Y. Yu, "A quick-response real-time stepping stone detection scheme," in *High Performance Computing and Communications (HPCC), 2010 12th IEEE International Conference on*, 2010, pp. 677-682.
[14] Y.-W. Kuo and S. Huang, "Stepping-stone detection algorithm based on order preserving mapping," in *Parallel and Distributed Systems, 2007 International Conference on*, 2007, pp. 1-8.
[15] J. Yang and S.-H. S. Huang, "Matching TCP/IP packets to detect stepping-stone intrusion," *International Journal of Computer Science and Network Security,* vol. 6, pp. p269-276, 2006.
[16] V. Paxson and Y. Zhang, "Detecting backdoors," in *Proc. of 9th USENIX Security Symposium*, 2000, pp. 157-170.
[17] M. S. Prasad, A. V. Babu, and M. K. B. Rao, "An Intrusion Detection System Architecture Based on Neural Networks and Genetic Algorithms," *International Journal of Computer Science and Management Research,* vol. 2, 2013.
[18] M. N. Omar, "Approach for Solving Active Perturbation Attack problem in Stepping Stone Detection," Ph.D dissertation, School of Computer Science, Universiti Sains Malaysia, Malaysia, 2011.
[19] K. Borders, X. Zhao, and A. Prakash, "Siren: Catching evasive malware," in *Security and Privacy, 2006 IEEE Symposium on*, 2006, pp. 6 pp.-85.
[20] G. Tahan, L. Rokach, and Y. Shahar, "Mal-ID: Automatic Malware Detection Using Common Segment Analysis and Meta-Features," *The Journal of Machine Learning Research,* vol. 98888, pp. 949-979, 2012.
[21] G. T. I. S, Center. (2013, July 13). *Open Malware* [Online]. Available: http://oc.gtisc.gatech.edu:8080
[22] W. Foundation. (2013, July 13). *Wireshark* [Online]. Available: http://www.wireshark.org
[23] Virustotal, (2013, July 13) *VirusTotal* [Online]. Available: http://www.virustotal.com