STEPPING-STONE DETECTION TECHNIQUE FOR RECOGNIZING LEGITIMATE AND ATTACK CONNECTIONS

Ali Yusny Daud[1], Osman Ghazali[2], and Mohd Nizam Omar[3]
Universiti Utara Malaysia, Malaysia, [1]aliyusny|[2]osman|[3]niezam@uum.edu.my

ABSTRACT. A stepping-stone connection has always been assumed as an intrusion since the first research on stepping-stone connections twenty years ago. However, not all stepping-stone connections are malicious. This paper proposes an enhanced stepping-stone detection (SSD) technique which is capable to identify legitimate connections from stepping-stone connections. Stepping-stone connections are identified from raw network traffics using timing-based SSD approach. Then, they go through an anomaly detection technique to differentiate between legitimate and attack connections. This technique has a promising solution to accurately detecting intrusions from stepping-stone connections. It will prevent incorrect responses that punish legitimate users.

Keywords: stepping-stone connection, stepping-stone detection, intrusion, legitimate, anomaly detection

INTRODUCTION

A stepping stone is an intermediate host or node that is used as a hop point usually in targeting a victim's host in network. Stepping-stone detection (SSD) is a process of detecting the route of intermediate points to the origin of the starting attack point. This process will help to detect the starting point but not the adjacent host where the connection comes.

Previous research in SSD mostly focused on detecting stepping-stone connections and stepping stones hosts. In SSD, every detected intrusion will be responded. However, not all connections which are characterized as stepping-stone connections are dangerous (Gilbert, Robinson, Butts, & Lacey, 2012; Gilbert, 2012). The main issue here is to correctly identify stepping-stone connections that are meant for intrusions. This has become vital because any false detection will affect the system because blocking the innocent channels may be considered as denial of service (DoS) attacks.

A few researches focused on identifying intrusions in stepping-stone connections. Study by Ding and Huang (2011) generalized that legitimate users do not make long connections or stepping-stone connections to remote host. They assumed only connections that use four or more hosts were considered as intrusions. Research by Zhang and Paxson (2000) also assumed only connection chains longer than three hosts are highly suspicious. However, this is contrary to the study conducted by Xinyuan Wang, Chen, and Jajodia in 2005 which indicated legitimate users also make long connections. Another research by Huang and Kuo (2011) believed that if stepping-stone connections are detected to have chaff, then those connections are considered as intrusions. However such explanation is doubtful because not all stepping

stone intrusions (connections) are being chaffed or the connections may have been perturbed by other pervasion techniques.

This paper proposes a preliminary technique to identify legitimate and attack connections from stepping-stone connections. This paper is organized as follows. In the first section, SSD, terminologies and variant SSD approaches are explained. The next section discusses legitimate stepping-stone connection and anomaly detection techniques. Different types of responses are also stated and enhanced SSD technique is presented at the end.

STEPPING-STONE DETECTION (SSD)

The stepping-stone connection is the chaining path from initiator to the victim through stepping stone machines. A stepping stone intrusion or attack occurs when attack commands or programs are transmitted through the connection chain from the attacker via intermediate hosts to the victim. When the victim performs an IP traceback, it will only recognize the adjacent host of the attacker but the initiator is undetected. This kind of attack will keep the attacker anonymous because the attacker is not directly connected to the victim's computer/host.

Terminologies

Figure 1 shows a stepping-stone connection scenario of five hosts. Host A is the initial attacker and Host E is the target. Host B, C and D are the intermediate hosts between Host A and Host E.

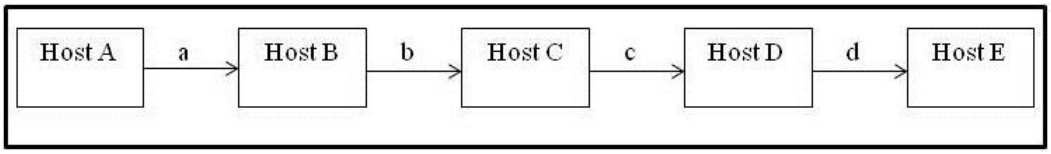

Figure 1. Stepping-stone connection

By referring to Host C; *b* is the incoming packet to Host C whereas *c* is the outgoing packet. The connection occurs when a host logs from one host to another host. Connection *a* connects Host A and B; *b* connects Host B and C; *c* connects Host C and D; and *d* connects Host D and E. So the sequence becomes a connection chain C = <a, b, c, d>. The downstream is the direction (arrows' direction) of a host's login and the upstream is the opposite direction.

SSD Approaches

Diversity of approaches or methods reflects the importance of conducting researches on SSD. Obviously most researchers agreed the first detection was content-based approach by Staniford-Chen and Heberlein (1995), unfortunately this approach seems inappropriate nowadays since most of today's data are encrypted (Almulhem & Traore, 2007; Mohd Nizam Omar, 2011). Content-based approach only works on unencrypted or unmodified data. Furthermore, it may compromise the confidentiality of data which are exposed in network traffics (Crescenzo, Ghosh, Kampasi, Talpade, & Zhang, 2011).

Similarly, deviation-based approach also cannot withstand modification at the stepping stone (Almulhem & Traore, 2007). Moreover, it is also costly because it has to measure all packets in the networks (Kuo, 2011). Watermark-based alike deviation-based is undeniably costly. The implementation of 'watermark' demands extra hardware such as watermarker and watermark detector which are required in performing detection (Xiaogang Wang, Luo, & Yang, 2012).

Recently timing-based and RTT-based approaches are mostly used in SSD. However, RTT-based was criticized for its nature of not accurately compute the packet timing (Kuo, 2011). Moreover, RTT-based consumes twice the time of timing-based approach (Shullich, Chu, Ji, & Chen, 2011) because it has to consider timing for 'send' and 'echo' packets. While the accuracy of time in RTT-based is questioned, the timing characteristics in timing-based approach is unique enough to be used in the detection of stepping stones. Timing-based approach was acknowledged as the current and future approach in SSD (Kuo, Huang, Ding, Kern, & Yang, 2010; Mohd Nizam Omar & Rahmat Budiarto, 2009).

LEGITIMATE STEPPING STONES

In the past years, SSD has attracted much interest and increased awareness from researchers. However, there was little attention has been paid to legitimate stepping stones. Most of the researchers concluded that the stepping-stone connection is considered as one of the ways to employ intrusion through network. However, some network traffic activities may seem as a stepping stones but are not harmful.

Wang et al. (2005), demonstrated that Voice over IP (VoIP) can be trace as stepping-stone connections. VoIP enables users to make telephone call through Internet rather than traditional Public Switch Telephone Network (PSTN). It becomes attractive to people because it offers more cost savings and more interesting features. The research proved that authorized traffic such as VoIP can be detected and interpreted as stepping stones.

In paper by Gilbert et al. (2012), it showed that administrators sometimes permitted communication access through certain gateways. In this case, the gateways that were connected to the computers were listed as stepping stones. Once again the connection was not meant for any malicious intention. So, it is very important to filter stepping-stone connections because legitimate traffic connections may look like malicious traffic.

ANOMALY-BASED DETECTION

Anomaly-based detection or behavior-based detection technique assumes normal behavior is different from unacceptable deviation of expected behavior. The model for the anomaly is pre-defined by rules or activities. Since its introduction (by Denning in 1987), many research has been done to enhance it. It considers anything that is not fit to the list of normal behaviour is considered as threat or attack. It is the best way to detect new attacks or also known as zero-day attacks (Mitchell & Chen, 2014). The major advantage of this technique is the ability to detect novel attacks which cannot be detected by existing signature (Casas, Mazel, & Owezarski, 2011).

Anomaly Techniques in SSD

Most studies that apply anomaly techniques in SSD have only been conducted in small number of areas. Research by Yung in 2002 had been recognized as the leading research in this area. It detected stepping stones based on the difference of 'send' packet and 'echo' packet. Yang et al. (2004, 2005) proposed analyzing connections using anomaly in real-time. They managed to reveal the step-function like results that indicated a stepping stone for each 'jump'. A work by Kampasi et al. (2007) and Giovanni et al. (2011) developed three anomaly algorithms to detect the presence of chaff and jitter in enhancing timing-based approach for better detection. Huang and Kuo in 2011 demonstrated when a stepping-stone connection is being chaffed, then the connection is considered as part of the intrusion.

It is clear that anomaly technique is a powerful tool in SSD but so far it is not been applied for identifying legitimate connection. The anomaly technique is chosen for it capability in defining normal and abnormal (intrusion). Combining timing-based approach with anomaly

technique ascertain our intention in uncovering legitimate (normal) connection from stepping-stone connections.

RESPONSE TO DETECTION

The system can respond either passively or actively. A passive system notifies the administrator if any abnormality happens. It can also attempt to terminate the connection before any intrusion started such as terminating active TCP session. Active responses include modifying the attacked entity state or even reconfigure them to block several types of actions or direct the connections to somewhere else.

Inappropriate responses may arise if the detector gives a wrong indication of false positive, i.e. give a false attack alarm even though the attack never happen (Shameli-sendi, Ezzati-jivan, Jabbarifar, & Dagenais, 2012). The consequences of producing wrongful judgment of responses will i) decrease the network performance because if the network is disconnected, it will consume more energy to get back to its healthy stage; ii) mistakenly disconnect innocent hosts/users from the network; iii) increase cost in bringing back the services to the users in the network; and iv) deny services to the users of the network or perform DoS attack to the network.

ACCURATE SSD TECHNIQUE

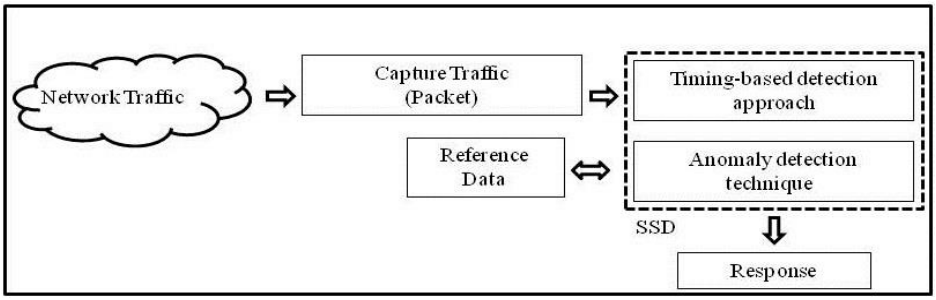

Figure 2. New SSD Technique

Figure 2 illustrates enhanced SSD is capable on identifying intrusion from stepping-stone connections. The key components of this technique are discussed below.

1) *Capture Traffic*. The network traffic (packet) is captured for analyze using common tool such as Wireshark or packet sniffer. The data (i.e. time) is used as input to timing-based approach.

2) *Timing-based detection approach*. Timing-based approach is used to classify raw data based on the necessary data from previous phase. This approach can recognize stepping-stone connections from raw network traffic.

3) *Anomaly detection technique*. An anomaly technique is responsible to determine which stepping stone-connection is legitimate or attack connection. Combining timing-based approach with anomaly technique ensures SSD is more accurate in identifying intrusion traffic.

4) *Reference data*. It stores information of normal behavior profiles. It provides anomaly detection with information for matching of legitimate or attack connection.

5) *Response*. Response acts based on the detection by SSD. An accurate detection ascertains responses are correctly executed.

At the end of the process, legitimate connections are identified from raw stepping-stone connections. Combination of timing-based approach with anomaly detection technique capable to identify between legitimate and attack connection. Prior SSD treats all stepping-stone connections as intrusions. It responses to such detection by actively disable or reduce the network performance. This kind of act costs a lot of money and energy to the user to rebuild the network to its peak. Our new SSD technique ensures that only attack connections being penalized. Legitimate connections remain operating as usual.

As a performance indicator in evaluating SSD, we use percentage of True Positive Rate (TPR) as in Eq. (1) and False Positive Rate (FPR) as in Eq. (2). Most of the research in SSD applied TPR and FPR as their evaluation approaches. Basically TPR will determine how efficient the process to detect normal and attack connections and FPR will calculate the misdetection of the process (García-Teodoro, Díaz-Verdejo, Maciá-Fernández, & Vázquez, 2009).

$$TPR = \frac{number\ of possible\ true\ instances}{number\ of\ true\ positive} \tag{6}$$

$$FPR = \frac{number\ of\ possible\ negative\ instances}{number\ of\ false\ positive} \tag{2}$$

True Positive Rate (TPR) refers to the fraction of true instances detected by the technique versus all possible true instances. In this case, the technique has correctly detected the normal attacks from the list of connection chains. FPR refers to the fraction of negative instances that are falsely reported by the technique as being positive. In this case, the technique has falsely detected the attack connections even though they are normal connections. Proof of concept for enhanced SSD (eSSD) is as in Eq. (3).

$$p \Leftrightarrow q \tag{3}$$

p: eSSD with anomaly technique; *q*: accuracy in defining legitimate user. The compound statement *p* if and only *q*, denoted as in Eq. (3) is equivalence or biconditional. The notation is true only when both *p* and *q* are true or both are false. The equivalence can also be stated as *p* is necessary and sufficient condition for *q*.

CONCLUSION

Previous research in SSD simply assumed all stepping-stone connections as intrusion connections. In this paper, we propose to combine timing-based approach with anomaly detection to identify legitimate and attack connections. Identifying legitimate connections from stepping-stone connections is crucial to avoid wrong punishments to the users in the network. Furthermore re-establishing the network after a wrong judgment affects users in the network and requires high cost and timely recovery. By identifying attack connection from stepping-stone connection, misjudgement in punishing legitimate users and negative actions on the network can be prevented.

REFERENCES

Almulhem, A., & Traore, I. (2007). A survey of connection-chains detection techniques. In *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, 219–222.

Casas, P., Mazel, J., & Owezarski, P. (2011). Steps towards autonomous network security: Unsupervised detection of network attacks. *2011 4th IFIP International Conference on New Technologies, Mobility and Security, NTMS 2011 - Proceedings.* doi:10.1109/NTMS.2011.5721067

Crescenzo, G. Di, Ghosh, A., Kampasi, A., Talpade, R., & Zhang, Y. (2011). Detecting anomalies in active insider stepping stone attacks. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *2*(1), 103–120.

Denning, D. E. (1987). An intrusion-detection model. *Software Engineering, IEEE Transactions on*, (2), 222–232.

Ding, W., & Huang, S. S. (2011). Detecting intruders using a long connection chain to connect to a host. *2011 IEEE International Conference on Advanced Information Networking and Applications*, 121–128. doi:10.1109/AINA.2011.109

García-Teodoro, P., Díaz-Verdejo, J., Maciá-Fernández, G., & Vázquez, E. (2009). Anomaly-based network intrusion detection: Techniques, systems and challenges. *Computers & Security*, *28*(1-2), 18–28. doi:10.1016/j.cose.2008.08.003

Gilbert, J. I. (2012). *Scalable wavelet-based active network stepping stone detection*. (Master's thesis, Air Force Institute of Technology, Air University, USA. Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a560009.pdf

Gilbert, J. I., Robinson, D. J., Butts, J. W., & Lacey, T. H. (2012). Scalable wavelet-based active network detection of stepping stones. In *SPIE Defense, Security, and Sensing* (p. 84080I–84080I).

Huang, S. S., & Kuo, Y. (2011). Detecting chaff perturbation on stepping-stone connection. *2011 IEEE 17th International Conference on Parallel and Distributed Systems*, 660–667. doi:10.1109/ICPADS.2011.51

Kampasi, A., Zhang, Y., Di Crescenzo, G., Ghosh, A., & Talpade, R. (2007). Improving stepping stone detection algorithms using anomaly detection techniques. *Report TR-07-28 (regular Report)*. The University of Texas at Austin.

Kuo, Y. (2011). Algorithms to detect stepping-stone intrusions in the presence of evasion techniques. *Doctoral dissertation*. Availabe from ProQuest Dissertations and Theses database (UMI No. 3492359).

Kuo, Y., Huang, S. S., Ding, W., Kern, R., & Yang, J. (2010). Using dynamic programming techniques to detect multi-hop stepping-stone pairs in a connection chain. *2010 24th IEEE International Conference on Advanced Information Networking and Applications*, 198–205. doi:10.1109/AINA.2010.132

Mitchell, R., & Chen, I. (2014). A survey of intrusion detection techniques for cyber-physical systems. *ACM Computing Surveys (CSUR)*, *46*(4), 55.

Mohd Nizam Omar. (2011). Approach for solving active pertubation attacks problem in stepping stone detection. *Unpublished Doctoral thesis*. Universiti Sains Malaysia, Malaysia.

Mohd Nizam Omar, & Rahmat Budiarto. (2009). Stepping stone detection (SSD): Towards to provide future SSD-based research. *MASAUM Journal of Basic and Applied Sciences*, *1*(2).

Shameli-sendi, A., Ezzati-jivan, N., Jabbarifar, M., & Dagenais, M. (2012). Intrusion response systems : Survey and taxonomy. *IJCSNS International Journal of Computer Science and Network Security*, *12*(1), 1–14.

Shullich, R., Chu, J., Ji, P., & Chen, W. (2011). A survey of research in stepping-stone detection. *International Journal of Electronic Commerce Studies*, *2*(2), 103–126.

Staniford-Chen, S., & Heberlein, L. T. (1995). Holding intruders accountable on the internet. In *Security and Privacy*, 39–49.

Wang, X., Chen, S., & Jajodia, S. (2005). Tracking anonymous peer-to-peer VoIP calls on the internet. In *Proceedings of the 12th ACM conference on Computer and communications security*, 81–91.

Wang, X., Luo, J., & Yang, M. (2012). An efficient sequential watermark detection model for tracing network attack flows. In *Proceedings of the 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design*, 236 – 243.

Wang, X., & Reeves, D. S. (2003). Robust correlation of encrypted attack traffic through stepping stones by manipulation of interpacket delays. *Proceedings of the 10th ACM Conference on Computer and Communication Security - CCS '03*, 20. doi:10.1145/948112.948115

Yang, J., & Huang, S. S. (2004). A real-time algorithm to detect long connection chains of interactive terminal sessions. In *Proceedings of the 3rd international conference on Information security*, 198–203. doi:10.1145/1046290.1046331

Yang, J., & Huang, S. S. (2005). Matching TCP packets and its application to the detection of long connection chains on the internet. In *Proceedings of the 19th International Conference on Advanced Information Networking and Applications (AINA'05)*, 1, 1005–1010. IEEE. doi:10.1109/AINA.2005.240

Yung, K. H. (2002). Detecting long connection chains of interactive terminal sessions. In *Proceedings of the 5th International Conference on Recent Advances in Intrusion Detection*, 1–16. Springer-Verlag. Retrieved from http://portal.acm.org/citation.cfm?id=1754703

Zhang, Y., & Paxson, V. (2000). Detecting stepping stones. *9th USENIX Security Symposium*, *171*, 1–11. Retrieved from http://www.cs.jhu.edu/~fabian/courses/CS600.424/course_papers/Stepping-Stones.pdf