Stepping Stone Detection: Measuring the SSD Capability

Ali Yusny Daud, Osman Ghazali, Mohd Nizam Omar

Abstract: ***The performance of Stepping Stone Detection (SSD) is measured by the accuracy to detect attacks that were initiated using stepping-stone hosts. The pattern of the attacks needs to be recognized to implement the detection. To evaluate the SSD, a variation of metrics have been used by many researchers but a benchmark should be introduced in calculating the measures. In this paper, we review the approaches used in evaluating the SSD and proposed the beneficial insights metrics in evaluating the effectiveness of SSD.***

Keywords: Stepping Stones, Intrusion, False Negative Rates, False Positive Rates, Percentage of Success.

I. INTRODUCTION

Intrusion can be put into practice by setting a connection path to the victim. Their intentions are to deploy an attack or steal data from the under attack host. Recent developments in computer attacks have heightened the need for network security. It will ensure the confidentiality, integrity and availability of data in network information system are safe [1].

Attackers in a stepping-stone attack hide their identity through a chain of compromised intermediate nodes while throwing attacks on the victims [2], [3]. Figure 1 shows the stepping stone attack from the attacker to the victim. An outside intruder might compromise a host in the administered network by taking advantages of some vulnerabilities and then use the host as the launch pad to achieve beneficial insight about the network and every host lying in it. The initially significant study of SSD was published in 1995 by Staniford-Chen and Heberlein[4].Ever since then, a lot of research has been conducted to detect stepping-stones attack which raises the relevancy of SSD till now.

Attackers will always try to hide their tracks to maintain the anonymity of the intrusion. Stepping stone is one of the most well-known techniques used by intruders to remain undetected by using the intermediate host as the attack path to the victim. Security detection software such as intrusion detection system (IDS) can only identify the adjacent host as the attacker as the real initiator remains undetected. Stepping stone detection (SSD) is the approach that can be used to track the real attacker by detecting the stepping-stone connection.

Revised Manuscript Received on March 08, 2019.

Ali Yusny Daud, School of Computing, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia,

Osman Ghazali, School of Computing, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia,

Mohd Nizam Omar, School of Computing, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia,

Measuring or evaluating the SSD is crucial to determine how 'good' is the SSD in detecting stepping stone attacks. It is essential for the administration to compare the capability of different SSD or to adjust the configuration of an SSD. In this paper, we look at the stepping stone attack and how it is initiated. SSD identify the stepping stone attack by examining the pattern of the traffic and identify the connection chains. The accuracy of detection of the SSD is measured by using a false negative rate (FNR), false positive rate and percentage of success.

The organization of this paper is outlined as follows. First, stepping stone attack and the SSD are explained in section 2. Then, the SSD evaluation design is discussed in section 3.Then, we look at the metrics used to evaluate the SSD in section4 and finally, we conclude the paper in the last section.

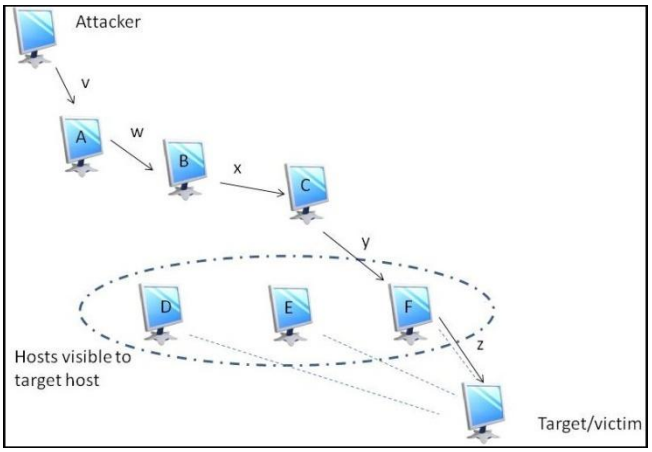

Fig. 1 Stepping stone attack

II. STEPPING STONE DETECTION (SSD)

Attack can be defined as an attempt to get access from the unauthorized system or manipulating information or using exploiting vulnerabilities in the system [5]. Stepping stone attack uses a sequence of hosts (stepping stones) to attack the victim. This kind of attack will keep the attacker remain anonymous because the attack is not directly connected to the victim's computer/host. It is very crucial to find the one responsible for attacks in stepping-stones connection. It is also important to detect the attack correctly for the SSD to response appropriately [6], [7].

Referring to Figure 2, host 1 to host 5represented by (H_1 to H_5).H_1 is the attacker while H_5 is the victim and the rest of the hosts are the intermediate hosts. C_1 to C_4 are the connections between hosts.

This is the example of direct connection is used for basic SSD environment.

Let say there are five hosts (H_1, H_2, H_3, $H_4$and H_5). I_n are inbound flows and O_n are the outbound flows for each *n*-host. Host H_1 begins to launch a stepping stone attack to Host

H_5 through Host H_2, H_3andH_4 by sending data C_1, C_2, C_3 andC_4. Therefore, O_{H1} = {C_1, C_2, C_3, C_4}, I_{H2} = {C_1, C_2, C_3, C_4}, O_{H2} = {C_1, C_2, C_3, C_4}, I_{H3} = {C_1, C_2, C_3, C_4}, O_{H3} = {C_1, C_2, C_3, C_4}, I_{H4} = {C_1, C_2, C_3, C_4}, O_{H4} = {C_1, C_2, C_3, C_4} and I_{H5} = {C_1, C_2, C_3, C_4}.

From here, O_{H1} = I_{H2}, O_{H2} = I_{H3}, O_{H3} = I_{H4}, O_{H4} = I_{H5} and these are identified as Flow $F_{H1,H5}$ from the host of origin, H_1, to host of destination, H_5. A stepping stone exists when $F_{H1,\ H2}$ = $F_{H2,H3}$ = $F_{H3,H4}$ = $F_{H4,H5}$ is identified as connection chain C which is a stepping stone connections.

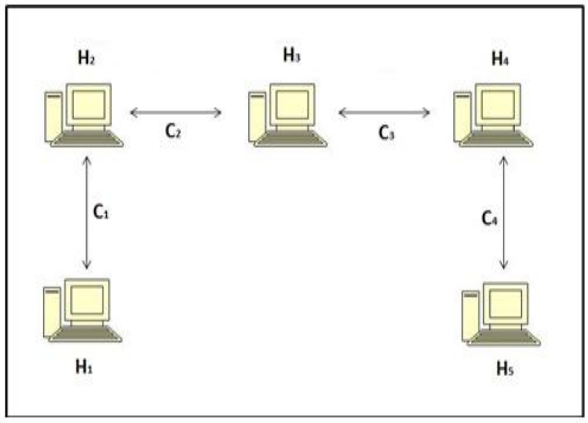

Fig. 2 Stepping stones connection

III. SSD EVALUATION DESIGN

The SSD evaluation design is illustrated in Figure 3. Table 1 explains the attributes in Figure 3. The effectiveness of the SSD depends on the value of evaluation calculated at the end of the process.

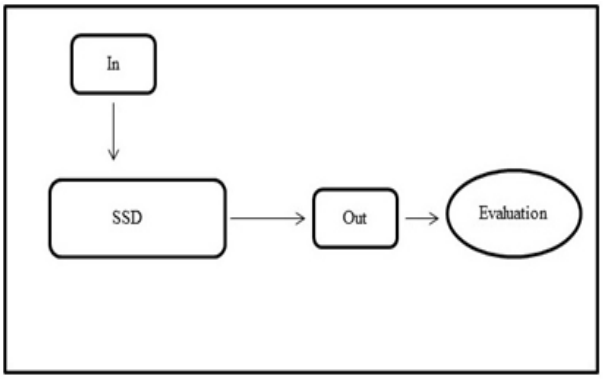

Fig. 3 SSD evaluation design

Table. 1 Attributes of SSD evaluation design

Attributes	Explanations
In	Input or raw traffic that going through the SSD
Out	The output or the result of detection from the SSD
Evaluation	Measuring the capability of the SSD

Raw traffics are the input in this process. It goes through the identifying process of SSD to determine if the traffic is stepping stone connection attack or not. 'Out' is the output or result of the identifying process. 'Evaluation' is the process of measuring how accurate (effective) the SSD in detecting the stepping stone attack.

IV. EVALUATION OF SSD

Evaluating the performance of SSD is fundamental in the field of detecting intrusion[8]. The performance of SSD generally can be evaluated in two standpoints [9]:

1. Effectiveness (also known as classification accuracy): Evaluate the capability of the SSD to differentiate between intrusion and non-intrusion actions.
2. Efficiency: Evaluate the allocation of the resources needed for the SSD including a main memory and CPU cycles.

Confusion Matrix

The confusion matrix is a matrix which represents the classification of the result. It represents the true and false of results classification. Majority of the researchers mostly focused on assessing the effectiveness and accuracy in terms of the rate of false alarm and the attacks percentage that is detected successfully. Researchers traditionally use True Positive (TP), True Negative (TP), False Positive (FP), and False Negative (FP) to measure and evaluates the performance of SSD [10].

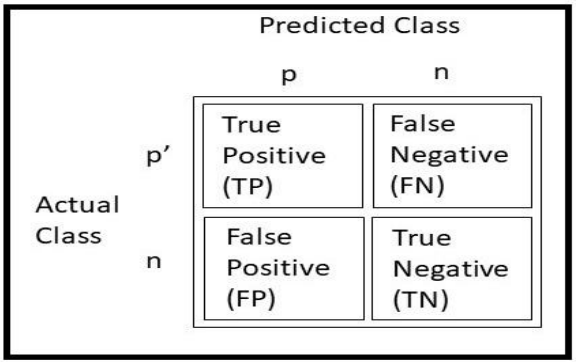

Fig. 4 Confusion matrix

Nowadays, computer systems are overloaded with vague information or false alarm [11]. Figure 4 shows the confusion matrix of the decision made by SSD either it is TP, TN, FP or FN. SSD may place its judgements whether the traffic is malicious or non-malicious. Table 2 shows the expression used in the context of TP, TN, FP, and FN in the successfulness of detection.

Table. 2 Truth table

Terminology	Actual Attack Situation	Alarm
TP	1	1
TN	0	0
FP	0	1
FN	1	0

TP and TN are the desired detection state for a security system. These four categories define the SSD efficiency in detecting stepping stone traffic.

The following explains the details of TP, TN, FP, and FN.

1. TP is defined as when the attack actually occurs; the security system effectively detects it. This is where the intrusions are detected successfully by the SSD.
2. TN is defined when the attack does not actually occur; the security system does not sound an alarm to it. This is where the non-intrusive/normal behaviour is successfully detected as non-intrusive/normal by the SSD.
3. FP is defined when the attack does not actually occur; the security system falsely sounds an alarm to it. This is where the non-intrusive/normal behaviour is classified wrongly as an intrusion by the SSD.
4. FN is defined when the attack actually occurs; the security system fails to detect it. This is where the intrusions are not detected by the SSD and classified as non-intrusive/normal.

Performance Metrics Derived from Confusion Matrix

Confusion matrix can represent a good classification but it is not enough to be a comparison tool for SSDs. Different performance metrics can be defined by looking at the variables of the confusion matrix. These metrics are easily comparable and are briefly explained as follows.

False Negative Rate (FNR)refers to the fraction of negative instances that are falsely detected by the algorithm versus all attack connections. In this case, the algorithm has incorrectly detected the attack connection as non-attack. The formula is shown in Equation 1.

$$\text{FNR} = \frac{number\ of\ attacks\ detected\ as\ normal}{total\ number\ of\ attack\ connections} \times 100$$

$$\text{FNR} = \frac{FN}{FN+TP} \quad (1)$$

False Positive Rate (FPR)refers to the fraction of positive instances that are falsely reported by the algorithm as being negative. In this case, the algorithm has falsely detected the normal connection as an attack. FPR is calculated based on the formula in Equation 2.

$$\text{FPR} = \frac{number\ of\ normal\ instances\ detected\ as\ attack}{total\ number\ of\ normal\ instances} \times 100$$

$$\text{FPR} = \frac{FP}{FP+TN} \quad (2)$$

Classification Rate (CR)is defined as the proportion of instances that correctly classified and the total number of instances. The formula is shown in Equation 3.

$$\text{CR} = \frac{correctly\ classified\ instances}{total\ no.\ of\ instances}$$

$$\text{CR} = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

Detection Rate (DR) is defined as the proportion between the number of attacks detected correctly and the total number of attacks. The formula is shown in Equation 4.

$$\text{DR} = \frac{correctly\ detected\ attacks}{total\ no.of\ attacks}$$

$$\text{DR} = \frac{TP}{TP+FN} \quad (4)$$

4.2 Percentage of Success

The quality and efficiency of SSD can be evaluated by using False Negative Rate (FNR) and False Positive Rate (FPR) [12]. In evaluating the identifying process, we will use a percentage of FNR and FPR. Most of the research in SSD and IDS apply FNR and FPR as their evaluation approach [10], [13], [14].

It is difficult sometimes to determine which SSD is better when using only FPR and FNR. If the first SSD can detect 10% more attacks, but the second SSD can produce 10% lower false alarms, so which one is better. By combining both uses of FPR and FNR, we define the percentage of success as shown in Equation 5.

Percentage of success =

$$\frac{total\ number\ of\ connections\ without\ FPR\ and\ FNR}{total\ number\ of\ connections} \quad (5)$$

An SSD is accurate when the values of FNR and FPR are low. It shows that the SSD making a correct decision in identifying attacks. The low value of false alarm indicates high accuracy. Furthermore, getting a high value in the percentage of success, it ensures the effectiveness of the SSD.

Another way to evaluate SSD is by using the Area under ROC curves (AUC). It is used to show the relationship between the false positive rate and the detection rate of a classifier. Moreover, it can compare the accuracy of several classifiers. Even though it is very effective, it has limitations.

Firstly, it is reliant on the ratio of attacks/intrusions to normal traffic. However, the comparison of the SSDs done on numerous data sets which are the ratio of attack/intrusion to normal instances are different. Secondly, AUC might be misleading and incomplete to understand the strengths and weaknesses of the system[9].

V. CONCLUSION

In this paper, the process of SSD has been explained by looking at the process of detection from raw traffic to the most trivial process; the evaluation of the SSD. The process includes the input, the output and the evaluation. The capability of how accurate an SSD in detecting stepping stone attacks depends on the performance metrics that have been presented. A high percentage of success, low FPR and FNR determined the high accuracy of SSD.

VI. ACKNOWLEDGMENT

This work was made possible by the Fundamental Research Grant Scheme (FRGS) (S/O 12896) by the Ministry of Education of Malaysia.

REFERENCES

1. Z. Lu and Y. Zhou, “The Evaluation Model for Network Security,” in 2014 Fourth International Conference on Communication Systems and Network Technologies, 2014, pp. 690–694.
2. S. H. S. Huang, H. Zhang, and M. Phay, “Detecting stepping-stone intruders by identifying crossover packets in SSH connections,” in Proceedings of the International Conference on Advanced Information Networking and Applications (AINA), 2016, pp. 1043–1050.
3. J. Yang, Y. Zhang, and G. Zhao, “Integrate stepping-stone intrusion detection technique into cybersecurity curriculum,” Proc. 31st IEEE Int. Conf. Adv. Inf. Netw. Appl. Work. WAINA 2017, pp. 1–6, 2017.
4. S. Staniford-Chen and L. T. Heberlein, “Holding intruders accountable on the internet,” in Security and Privacy, 1995, 1995, pp. 39–49.

5. O. Al-Jarrah and A. Arafat, “Network Intrusion Detection System using Attack Behavior Classification,” 2014 5th Int. Conf. Inf. Commun. Syst. ICICS 2014. IEEE Comput. Soc. https//doi.org/10.1109/IACS.2014.6841978, 2014.
6. S. Anwar, J. M. Zain, M. F. Zolkipli, Z. Inayat, A. N. Jabir, and J. B. Odili, “Response option for attacks detected by intrusion detection system,” in 2015 4th International Conference on Software Engineering and Computer Systems, ICSECS 2015: Virtuous Software Solutions for Big Data, 2015, pp. 195–200.
7. S. Anwar et al., “From intrusion detection to an intrusion response system: Fundamentals, requirements, and future directions,” Algorithms, vol. 10, no. 2, 2017.
8. G. Gu, P. Fogla, D. Dagon, W. Lee, and B. Skorić, “Measuring intrusion detection capability: An information-theoretic approach,” in Proceedings of the 2006 ACM Symposium on Information, computer and communications security, 2006, pp. 90–101.
9. G. Kumar, “Evaluation Metrics for Intrusion Detection Systems - A Study,” Int. J. Comput. Sci. Mob. Appl., vol. 2, no. 11, pp. 11–17, 2014.
10. R. Mitchell and I. Chen, “A Survey of Intrusion Detection Techniques for Cyber-Physical Systems,” ACM Comput. Surv., vol. 46, no. 4, p. 55:1-29, 2014.
11. R. Zuech, T. M. Khoshgoftaar, and R. Wald, “Intrusion detection and Big Heterogeneous Data: a Survey,” J. Big Data, vol. 2, no. 1, 2015.
12. R. Kumar and B. B. Gupta, “Stepping stone detection techniques: Classification and state-of-the-art,” in Proceedings of the international conference on recent cognizance in wireless communication & image processing, 2016, pp. 523–533.
13. A. Kampasi, Y. Zhang, G. Di Crescenzo, A. Ghosh, and R. Talpade, “Improving stepping stone detection algorithms using anomaly detection techniques,” Rep. TR-07-28 (regular report), no. The University of Texas at Austin, 2007.
14. G. Di Crescenzo, A. Ghosh, A. Kampasi, R. Talpade, and Y. Zhang, “Detecting anomalies in active insider stepping stone attacks,” J. Wirel. Mob. Networks, Ubiquitous Comput. Dependable Appl., vol. 2, no. 1, pp. 103–120, 2011.