Anomaly Techniques in Stepping Stone Detection (SSD): A Review

Ali Yusny Daud, Osman Ghazali, Mohd Nizam Omar
School of Computing, College of Arts and Sciences, Universiti Utara Malaysia
aliyusny@uum.edu.my

***Abstract*—Stepping Stone Detection (SSD) can be used to trace back the real attacker in stepping-stone connection. Anomaly techniques are capable of identifying between normal and abnormal traffic. The collaboration of SSD and anomaly techniques enhanced the capability of detection of stepping-stone connection. Several SSD approaches and anomaly techniques have been proposed in the literature. In this paper, we review these approaches and techniques. Furthermore, we suggest a potential future of anomaly techniques in SSD.**

***Index Terms*—Anomaly; Attack; Stepping-Stone Detection; Trace Back.**

I. INTRODUCTION

Computer attacks can be easily made without being close to the victim by sending the malicious code through the network. The attackers can be everywhere all around the world. To make it even worst, the attackers can use an intermediate host to channel their attacks. This kind of attack will take over control of the intermediate host (or called stepping stones) and launch the attacks. It will keep the anonymity of the attackers because it seems that the attacker is not directly involved. Attack from an adjacent host can be easily exposed the attacker but to trace back the route of the initial attacker from a chain of attack is going to be even hard and trickier [1].

Figure 1 demonstrates how the initial attacker can get away from being detected. In this scenario, Host C is detected to be the attacker because the attack traffic from Host C is the visible traffic to the target host or the victim. Unfortunately, the real culprit has got away undetected. So, here is where SSD plays it role in detecting where the attack is really coming from.

In this paper, we survey stepping stone detection (SSD) approaches for detecting connection-chains and anomaly techniques used in SSD that have been discussed in the literature. In general, SSD approaches can be divided into content-based, timing-based, deviation-based, watermark-based and round-trip time-based (RTT).

The remaining of the paper is outlined as follows. First, SSD is explained in section II and the approaches in section III. Then, anomaly techniques are discussed in section IV. Then, we look at SSD that applied anomaly techniques in their detection in section V. Finally, we conclude the paper and present promising direction in the last section.

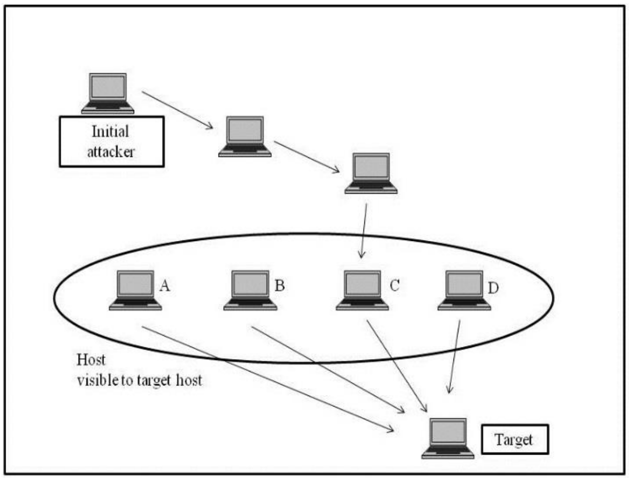

Figure 1: Stepping stone scenario

II. STEPPING STONE DETECTION (SSD)

Connection chains refer to a set of connections where someone logs into one host and then keep on recursively log into another host and so forth [2]. The intermediate hosts in between the first host and the target host are called stepping stones [3]. It has become a great way for attackers to hide their activities and remain anonymous to the victims. The victims can only identify attacks coming from the last host in the connection chain while the real attacker hides behind the stepping stones.

Figure 2 show Host A is the sender or the attacker, and Host E is the receiver or the victim. Host B, C and D are the intermediate nodes or stepping stones. Whereas connection a, b, c and d perform as the connection chain from Host A to E. Therefore, a stepping stone connection is detected when content in connection a = b = c = d. SSD is the process of tracing the initial attacker back to Host A.

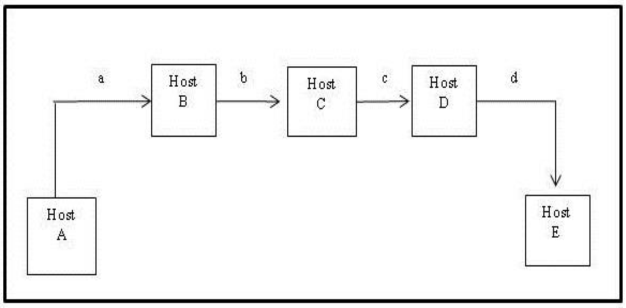

Figure 2: Stepping stone connection

III. SSD Approaches

During the past 25 years, an increasing amount of literature on SSD has been available. To date, various approaches have been developed and introduced to implement stepping-stone detections which are content-based, timing-based, deviation-based, watermark-based and round-trip time-based (RTT) approaches.

A. Content-based

SSD was first shown or demonstrated by Staniford-Chen and Heberlein in 1995 [2]. At that time, they introduced thumbprint, a content-based technique to trace connection chains. Thumbprint works by looking at the similarity of content in two connections. Unfortunately, thumbprint fails to detect stepping stones if the content is encrypted or modified during the transmission.

B. Timing-based

In 2000, Zhang and Paxson [3] demonstrated the replacement for content-based with a timing-based approach of detection. They ignored the contents and used other characteristics such as timing and packet size to detect stepping stones. Stepping-stone connection can be detected by introducing a pattern of the ON/OFF condition in the transmission. The critical problem with this research is that if the attacker injects evasion techniques such as time jitter or chaff, the detection rate of stepping stone will be decreased.

Initial ON/OFF algorithm [3] was improved by Kampasi et al. [4] by applying anomaly detection techniques. Three separate anomaly detection techniques were used to detect chaff and time jitter in stepping-stone connection. These three algorithms enhanced the existing timing-based approach. These algorithms worked in the presence of ‘magic numbers’ to conduct detection. However, authors failed to explain the computation nature and adjustment of that ‘magic numbers’. Controversy remains regarding how to create and use these numbers if we do not know how to tune them.

In 2002, Wang et al. [5] published a paper in which they described a timing-based approach that correlates connection based on inter-packet delay (IPD) timing characteristics. A correlation between two flows is detected by looking at the similarity of connection generated by the IPD.

The introduction of research using timing perturbation attack against timing-based correlation was first studied by Donoho et al. [6]. However, the study fails to show experimental results of the research. The study also fails to address other perturbation attacks such as chaff perturbation and the assessment of false positive and false negative rate of SSD.

Later study after that considers the effect of perturbation attacks on SSD. By introducing the method of Detect-Attack (DA), Blum et al. [7] investigated numbers of chaff packets to avoid detection. Their work found that there is ‘upper bound’ of packets for detecting stepping stone. In another study by Zhang et al. [8], they demonstrated SSD in the introduction of timing delay and chaff perturbation simultaneously. However, they only managed to detect the case of chaff perturbation.

He and Tong [9], [10] have developed a detection algorithm known as Detect-Match (DM). This algorithm takes account of detection without chaff packets. Perhaps the most contribution of this study is that DM reduced the complexity from exponential to linear of the result compared to DA [7].

The use of timing-based approaches by looking at the similarities in the incoming and outcoming stream is still considered the most capable and promising approaches till present time [1], [11], [12].

C. Deviation-based

Research by Yoda and Etoh [13], is where deviation method was introduced. Deviation assumed if the size of bytes transferred grows gradually at the same rate, then that connection is a stepping-stone connection. However, this method has drawbacks. It cannot be used on compressed data because it is dependable on the size of the data. It is also making the detection after the end of connection because the correlation metrics were defined after entire connections complete [14].

D. Watermark-based

The study of watermark-based approach [15]–[19] suggested embedding code or watermark into the network traffic flow. If this watermark re-appears again in another traffic flow, the two connections belong to the same stepping-stone connection. Watermark is actually created by modifying the inter-packet timing date in the traffic flow. The embedded watermark then can be recovered through detection process or can even be duplicated if it is improperly designed [20].

Peng et al. [21] studied the effects of chaff and limited perturbation using watermarking in the active timing-based algorithm. The algorithm injected the watermark into the upstream traffic and then detected the watermark in the downstream traffic. However, this method enables an attacker to counter back the detection because the detection process was exposed by the active intervention method.

Pyun et al. [22] demonstrated probabilistic watermarking without being precisely synchronised. This study is actually an extension of research in [18] using re-packetisation. The duration of traffic flow was chunk into fixed-length intervals. It also does not have to be precisely synchronised between encoder and decoder as a previous probabilistic watermarking method. One major drawback of this method is the high cost of implementation compared to other methods.

Wang and Reeves [23] introduced a novel watermarking method to detect stepping stones with the presence of timing perturbation. They managed to show theoretically that their method managed to detect almost 100 percent of True Positive Rate (TPR) and close to zero percent of False Positive Rate (FPR). However, this method needs a considerable amount of packet to do the detection. In 2012, Wang et al. [24] managed to demonstrate their Efficient Sequential Watermark Detection (ESWD) which experimentally reduced twenty-eight percent of packet needed for detection. Unfortunately, contradict from [23], their method was not robust against timing perturbation.

E. Round Trip Time (RTT)-based

The study on RTT was first carried out by Yung [25]. The idea of RTT is to find the length of the connection chain by comparing delayed acknowledgement gap and reply-echo gap. However, such explanation on calculating the length remains in doubt because the author simply multiplies the time by two that can be inaccurate.

Yang et al. [11], [26], [27] try to overcome the weakness of RTT in [25]. The authors computed the length by matching the 'send' and the corresponding 'echo' packet. They applied algorithms which determined the RTT-based on the changes of packet RTT. Yang and Huang [28] published research on a Clustering-Partitioning algorithm using TCP/IP packets RTT to compute the length of a connection chain. Their experiment managed to detect stepping-stone intrusion when they can find the length. The primary issue of this method is that it is not efficient. Linear increases of the dataset will increase the timing in quadratic. Sheng et al. [29] introduced Computing Dataset X (CDA) to be used in Clustering-Partitioning and able to lower the running time by decreasing the dataset.

Ding and Huang [30], show another way of using RTT in SSD which called upstream RTT or uRTT. However, uRTT was not a better measurement of RTT because of its dependency on the subjective ability of human keystrokes which created different timing. The timing measurement of RTT is also not accurate because of the occurrence of crossover. Crossover occurred when the 'echo' from the first message still did not reach the host, another message has been sent. This research also tolerates as high as 15 percent of FPR.

IV. Anomaly Detection Techniques

The anomaly-based technique is the process of analysing profile for normal traffic behaviour. The profile for 'normal' (legitimate) is first created to be the baseline. All new activity will be compared with the profile. Any deviation from the standard profile is called an anomaly. Most commonly used techniques for anomaly-based are statistical, machine learning, knowledge and data mining.

A. *Statistical-based*

Statistical-based techniques are based on capturing the activities of network traffic, and a profile of attack/normal behaviour is created. The profile is based on the number of packets, connection rates, transfer rates etc.

In the process of anomaly detection, two datasets of traffic are analysed. The first dataset corresponds to current profile over time while the second one is the trained statistical profile that previously created. An anomaly score will be approximated by comparing the two behaviours. The score points out abnormality of the specific event when the score exceeds a certain threshold.

The first statistical technique was using univariate model [31]. The parameters are modelled as independent Gaussian variables. It defined every variable the values of the acceptable range. Ye et al. in 2002 [32] proposed multivariate models which consider the correlation of two or more metrics or parameter. It had been shown that better level of discrimination could be produced by combining related measures compared to individually.

Later, time series model [33] using interval timer with event counter/ resource measure was introduced. It considers the inter-arrival times and order of the observation together with their values. The observed traffic will be treated as abnormal if, at the given time, the probability of occurrence is low.

B. *Machine Learning-based*

Machine learning-based techniques focused on constructing a model that adapt and improve performance based on it earlier results. It needs data with particular characteristics to train behavioural model. This procedure demands valuable resources. Several schemes of machine learning-based are discussed as follows.

C. *Bayesian Networks*

Bayesian networks code the probability relationship of variables of interest. Usually, it can be combined with a statistical scheme. It then will be capable of coding interdependencies between variables and predicting the event. However, the results from the Bayesian network are similar to threshold system, but with higher computational effort model [34].

D. *Markov Model*

There are two approaches for Markov models which are Markov Chain and Hidden Markov. Markov Chain consists of a set of states which are interrelated through transition probabilistic. It defines the capabilities and topology of the model. In the first training phase, the probabilities are estimated from the normal behaviour of the target system. Anomalies were detected by comparing the anomaly score from the observed sequence with the fixed threshold. Hidden Markov is where the target system is assumed a Markov process where states and transition are hidden. Only productions are available to be observed [35], [36].

E. *Neural Networks*

Neural networks simulate the human brain (neuron and synapses among them) in the anomaly detection. Neural networks offer adaptability and flexibility to environmental changes. It creates user profile [37], predicts next command from preceding activity [38], and identify attack from the traffic patterns [39]. However, neural networks do not provide an explanation why a particular detection decision has been taken [40].

F. *Fuzzy Logic*

Fuzzy logic is taken from the theory of fuzzy where reasoning is approximately deduced (rather than precise) from predicate logic. It is used in anomaly detection as the features to be accounted can be modelled as fuzzy variables [41]. The observation will be considered as normal when it lies within a given interval [42]. The main disadvantage of fuzzy logic is the use of high resource consumption. It is also rejected by most statisticians, and some engineers who stand for probability is the only description for uncertainty [40].

G. *Genetic Algorithm*

Genetic algorithms used other techniques of machine learning-based motivated by evolutionary biology. It capable of choosing parameters or features for detection [41] and derived classification rules [43]. The major setback of this technique is also the high consumption of resources.

H. *Clustering and Outlier Detection*

This technique creates clusters by grouping the observed data depending on distance or similarity measure. Then it will select a representative point in each cluster. Each data point belongs to a cluster depending on the proximity of the corresponding point [44]. Points that are not belonging to any clusters are called outliers or anomalies in the detection process. Clustering answered the question is the outlier an anomaly [45].

I. Knowledge-based

Knowledge-based or also known as expert system classifies audit data using three steps involving a set of rules [40]. Firstly, classes and attributes are recognised from the trained data. Secondly, the procedures or parameters or classification rules are deduced. Finally, the data are accordingly classified.

Specification-based methods are more restrictive. A human expert normally constructs the model. The experts determine the rules for legitimate behaviour to create a complete model. Furthermore, false positive will be reduced because it avoids the problems of legitimate activities being detected as intrusions. A formal tool Finite State Machine (FSM) can also be used to develop a specification for anomaly model. FSM can provide a sequence of states and transitions for modelling network protocols [46].

J. Data Mining

Data mining technique is used to decrease the complexity of dataset rather than work as a detection scheme. Two most common of data mining techniques are Principal Component Analysis (PCA) and association rule discovery.

A dataset becomes more complex and broad as different services or speed of the network propagates. PCA simplify the dataset. PCA makes a translation by ‘n’ correlated variables or ordered to reduce the number of variables ‘d’ so ‘d<n’. This will facilitate detection process [47].

Association Rule Discovery workings by obtaining a correlation between different features from the datasets. For example, find an internal relationship between data, in a specific connection. Some algorithms of association rules are given in [48]. The term ‘data mining’ has been commonly applied for IDS processing whereas data mining is actually used to correlate network traffic instances in the database. Almost every anomaly detection techniques can apply data mining in dealing with huge databases [40].

V. SSD Using Anomaly Techniques

Studies in applying anomaly techniques in SSD have only been done in a small number of research. Research by Yung in 2002 [25] had been renowned as the leading research in the extent of this area. The research [25] identified stepping stones based on the difference of ‘send’ packet with ‘echo’ packet. Unfortunately matching the packet of ‘send’ and ‘echo’ can be imprecise if the traffic is encoded [49]. Yang et al. [11], [26] recommended analysing the traffic connections using anomaly techniques in real-time. Their study managed to disclose the result of step-function like that specified a stepping stone for each indicator of ‘jump’. One main issue in this study is they were established on RTT-based approach. RTT does not have a capability to estimate time of ‘send’ and ‘echo’ accurately.

Research by Giovanni et al. [50] and Kampasi et al. [4] suggested anomaly used for detection of chaff and jitter in network connections rather than detecting stepping stones. They established three anomaly algorithms to sense the presence of chaff and jitter in timing-based approach for enhanced detection.

Huang and Kuo (2011) [51] established their anomaly for detection of chaff in stepping-stone detections. They have inferred the conclusion of identifying an attack in stepping-stone connection by only detecting the existence of chaff in the internet connections. They believed if chaff is detected in the stepping-stone connection, then the connection is being part of the attack. Though, such explanations overlook the fact that they are other evasion techniques rather than chaff such as jitter and dropped packet [52]–[55] to distract the detection of stepping-stone connection.

VI. Future Potential of Anomaly Techniques in SSD

Previously, a study on SSD using anomaly techniques had little attention or do not consider on identifying the attack and legitimate traffic for stepping-stone connection. So far, earlier research applied anomaly to detect stepping stones [11], [25], [26]; to detect chaff and jitter in connections chain [4], [50]; and to detect chaff to verify the connection is attack connection [51].

Next potential study can be focussing on identifying attack connections and legitimate connections. Preliminary research in SSD naturally assumed that all stepping-stone connections are attacks. Furthermore, every attack must be responded in an adequate way such as disconnecting the networks.

However, not all stepping-stone connections are attacks [56], [57] and we may misleadingly respond to the false alarm. Some network traffic activities may seem like stepping stones but are not harmful. One example is given in [56] was automated polling systems. Wang et al. [15], demonstrated that Voice over IP (VoIP) could be traced as stepping-stones connections. It is very crucial to identify which are legitimate and which one attacks connections in detecting stepping stone. This will prevent legitimate users to suffer from wrongful responses by the system.

In ensuring SSD can enhance identifying process between legitimate and attacks, anomaly detection techniques have to be integrated with the SSD. This is because anomaly technique is well known for their capability in categorising normal (legitimate) and abnormal (attacks). Hence, combining SSD with anomaly detection techniques will recognise the right connections to be responded rightfully.

VII. Conclusion

In this review, approaches in SSD have been classified in detecting stepping stone connections. Furthermore, we also reviewed techniques in anomaly which capable of identifying between normal or abnormal traffic. In parallel with the anomaly capacity to do classification, we conclude this paper by emphasising a possible directions SSD research. Looking at this review as an initial point, SSD can enhance their detection in identifying attack connection.

References

[1] Y. Kuo, “Algorithms to detect stepping-stone intrusions in the presence of evasion techniques,” (Doctoral dissertation), Available from ProQuest Dissertations and Theses database (UMI No. 3492359), 2011.

[2] S. Staniford-Chen and L. T. Heberlein, “Holding intruders accountable on the internet,” in *Security and Privacy, 1995*, 1995, pp. 39–49.

[3] Y. Zhang and V. Paxson, “Detecting stepping stones,” *9th USENIX Secur. Symp.*, vol. 171, pp. 1–11, 2000.

[4] 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.

[5] X. Wang, D. Reeves, and S. Wu, “Inter-packet delay based correlation for tracing encrypted connections through stepping stones,” *Comput. Secur. 2002*, pp. 1–20, 2002.

[6] D. L. Donoho, A. G. Flesia, U. Shankar, V. Paxson, J. Coit, and S.

Staniford, “Multiscale stepping-stone detection: detecting pairs of jittered interactive streams by exploiting maximum tolerable delay,” in *International Symposium on Recent Advances in Intrusion Detection*, 2002, vol. 2516, pp. 16–18.

[7] A. Blum, D. Song, and S. Venkataraman, “Detection of interactive stepping stones: Algorithms and confidence bounds,” *Recent Adv. Intrusion Detect. Springer Berlin Heidelb.*, pp. 258–277, 2004.

[8] L. Zhang, A. G. Persaud, A. Johnson, and Y. Guan, “Detection of stepping stone attack under delay and chaff perturbations,” in *Conference Proceedings of the IEEE International Performance, Computing, and Communications Conference*, 2006, vol. 2006, pp. 247–256.

[9] T. He and L. Tong, “A signal processing perspective to stepping-stone detection,” in *2006 IEEE Conference on Information Sciences and Systems, CISS 2006 - Proceedings*, 2007, pp. 687–692.

[10] T. He and L. Tong, “Detecting encrypted stepping-stone connections,” *IEEE Trans. Signal Process.*, vol. 55, no. 5 I, pp. 1612–1623, 2007.

[11] J. Yang and S. S. Huang, “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)*, 2005, vol. 1, pp. 1005–1010.

[12] Y. Kuo and S. S. Huang, “Detecting Stepping-Stone Connection Using Association Rule Mining,” *2009 Int. Conf. Availability, Reliab. Secur.*, pp. 90–97, 2009.

[13] K. Yoda and H. Etoh, “Finding a connection chain for tracing intruders,” in *Computer Security-ESORICS 2000*, Springer, 2000, pp. 191–205.

[14] J. Yang and E. Bosworth, “An efficient TCP/IP packet matching algorithm to detect stepping-stone intrusion,” *2009 Inf. Secur. Curric. Dev. Conf. - InfoSecCD ’09*, p. 1, 2009.

[15] X. Wang, S. Chen, and S. Jajodia, “Tracking anonymous peer-to-peer VoIP calls on the internet,” in *Proceedings of the 12th ACM conference on Computer and communications security*, 2005, pp. 81–91.

[16] Y. H. Park and D. S. Reeves, “Adaptive Watermarking against Deliberate Random Delay for Attack Attribution through Stepping Stones,” in *Proc. Of the Ninth International Conference on Information and Communications Security (ICICS 2007)*, 2007.

[17] X. Wang, S. Chen, and S. Jajodia, “Network flow watermarking attack on low-latency anonymous communication systems,” in *Proceedings - IEEE Symposium on Security and Privacy*, 2007, pp. 116–130.

[18] X. Wang and D. S. Reeves, “Robust correlation of encrypted attack traffic through stepping stones by manipulation of interpacket delays,” *Proc. 10th ACM Conf. Comput. Commun. Secur. - CCS ’03*, p. 20, 2003.

[19] A. Bates, B. Mood, J. Pletcher, H. Pruse, M. Valafar, and K. Butler, “Detecting Co-Residency with Active Traffic Analysis Techniques Categories and Subject Descriptors,” in *Proceedings of the 2012 ACM Workshop on Cloud computing security workshop - CCSW ’12*, 2012, pp. 1–12.

[20] P. Peng, P. Ning, and D. S. Reeves, “On the secrecy of timing-based active watermarking trace-back techniques,” in *Proceedings - IEEE Symposium on Security and Privacy*, 2006, vol. 2006, pp. 334–348.

[21] P. Peng, P. Ning, D. S. Reeves, and X. Wang, “Active timing-based correlation of perturbed traffic flows with chaff packets,” *25th IEEE Int. Conf. Distrib. Comput. Syst. Work.*, 2005.

[22] Y. J. Pyun, Y. H. Park, X. Wang, D. S. Reeves, and P. Ning, “Tracing traffic through intermediate hosts that repacketize flows,” in *Proceedings of 26th IEEE International Conference on Computer Communications*, 2007, pp. 634–642.

[23] X. Wang and D. S. Reeves, “Robust correlation of encrypted attack traffic through stepping stones by flow watermarking,” *IEEE Trans. Dependable Secur. Comput.*, vol. 8, no. 3, pp. 434–449, 2011.

[24] X. Wang, J. Luo, and M. Yang, “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*, 2012, pp. 236–243.

[25] K. H. Yung, “Detecting long connection chains of interactive terminal sessions,” in *Proceedings of the 5th International Conference on Recent Advances in Intrusion Detection*, 2002, pp. 1–16.

[26] J. Yang and S. S. Huang, “A real-time algorithm to detect long connection chains of interactive terminal sessions,” in *Proceedings of the 3rd international conference on Information security*, 2004, pp. 198–203.

[27] J. Yang and S. S. Huang, “Matching TCP / IP Packets to Detect Stepping-Stone,” *IJCNCS Int. J. Comput. Sci. Netw. Secur.*, vol. 6, no. 10, pp. 269–277, 2006.

[28] J. Yang, S. H. S. Huang, and M. D. Wan, “A clustering-partitioning algorithm to find TCP packet round-trip time for intrusion detection,” *20th Int. Conf. Adv. Inf. Netw. Appl. AINA*, vol. 1, pp. 231–236, 2006.

[29] Y. Sheng, Y. Zhang, and J. Yang, “Mining Network Traffic Efficiently to Detect Stepping-Stone Intrusion,” in *2012 IEEE 26th International Conference on Advanced Information Networking and Applications*, 2012, pp. 862–867.

[30] W. Ding and S. H. S. Huang, “Detecting intruders using a long connection chain to connect to a host,” in *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, 2011, pp. 121–128.

[31] D. E. Denning and P. G. Neumann, “Requirements and model for IDES—a real-time intrusion detection expert system,” 1985.

[32] N. Ye, S. M. Emran, Q. Chen, and S. Vilbert, “Multivariate statistical analysis of audit trails for host-based intrusion detection,” *IEEE Trans. Comput.*, vol. 51, no. 7, pp. 810–820, 2002.

[33] S. C. Chin, A. Ray, and V. Rajagopalan, “Symbolic time series analysis for anomaly detection: A comparative evaluation,” *Signal Processing*, vol. 85, no. 9, pp. 1859–1868, 2005.

[34] C. Kruegel, D. Mutz, W. Robertson, and F. Valeur, “Bayesian event classification for intrusion detection,” in *Proceedings of 19th Annual Computer Security Applications Conference, 2003*, 2003, pp. 14–23.

[35] M. V Mahoney and P. K. Chan, “Learning nonstationary models of normal network traffic for detecting novel attacks,” in *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining KDD 02*, 2002, pp. 376–385.

[36] J. M. Estévez-Tapiador, P. García-Teodoro, and J. E. Díaz-Verdejo, “Detection of web-based attacks through Markovian protocol parsing,” in *Proceedings - IEEE Symposium on Computers and Communications*, 2005, no. Iscc, pp. 457–462.

[37] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, 2009.

[38] H. Debar, M. Becker, and D. Siboni, “A neural network component for an intrusion detection system,” *Proc. 1992 IEEE Comput. Soc. Symp. Res. Secur. Priv.*, 1992.

[39] S. Mukkamala, G. Janoski, and a. Sung, “Intrusion detection using neural networks and support vector machines,” *Proc. 2002 Int. Jt. Conf. Neural Networks. IJCNN’02*, pp. 1702–1707, 2002.

[40] P. García-Teodoro, J. Díaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, “Anomaly-based network intrusion detection: Techniques, systems and challenges,” *Comput. Secur.*, vol. 28, no. 1–2, pp. 18–28, 2009.

[41] S. M. Bridges and R. B. Vaughn, “Fuzzy Data Mining and Genetic Algorithms Applied To Intrusion Detection,” in *23rd National Information Systems Security Conference*, 2000, pp. 13–31.

[42] J. E. Dickerson and J. a. Dickerson, “Fuzzy network profiling for intrusion detection,” *PeachFuzz 2000. 19th Int. Conf. North Am. Fuzzy Inf. Process. Soc. - NAFIPS (Cat. No.00TH8500)*, pp. 301–306, 2000.

[43] W. Li, “Using genetic algorithm for network intrusion detection,” in *Proceedings of the United States Department of Energy Cyber Security Group 2004 Training Conference, Kansas City, Kansas*, 2004, pp. 24–27.

[44] L. Portnoy, E. Eskin, and S. Stolfo, “Intrusion detection with unlabeled data using clustering,” in *Proceedings of ACM CSS Workshop on Data Mining Applied to Security Philadelphia PA*, 2001, pp. 1–25.

[45] K. Sequeira and M. Zaki, “ADMIT: anomaly-based data mining for intrusions,” in *ACM SIGKDD international conference on Knowledge discovery and data mining*, 2002, pp. 386–395.

[46] J. M. Estevez-Tapiador, P. Garcia-Teodoro, and J. E. Diaz-Verdejo, “Stochastic protocol modeling for anomaly based network intrusion detection,” in *First IEEE International Workshop on Information Assurance, 2003. IWIAS 2003. Proceedings.*, 2003, pp. 3–12.

[47] W. Wang and R. Battiti, “Identifying intrusions in computer networks with principal component analysis,” in *Proceedings - First International Conference on Availability, Reliability and Security, ARES 2006*, 2006, vol. 2006, pp. 270–277.

[48] W. W. Cohen, “Fast Effective Rule Induction,” in *Proceedings of the 12th international conference on machine learning*, 1995, pp. 115–123.

[49] A. Almulhem and I. Traore, “A survey of connection-chains detection techniques,” in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, 2007. PacRim 2007.*, 2007, pp. 219–222.

[50] 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.

[51] S. S. Huang and Y. Kuo, “Detecting chaff perturbation on stepping-stone connection,” *2011 IEEE 17th Int. Conf. Parallel Distrib. Syst.*, pp. 660–667, Dec. 2011.

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

[53] Mohd Nizam Omar, “Intelligent host-based stepping stone detection (I-HSSD): Comparison between self-organization map and data mining

approach," *Int. J. Intell. Inf. Technol. Appl.*, vol. 2, no. 6, pp. 256–263, 2009.

[54] Mohd Nizam Omar and Rahmat Budiarto, "Stepping stone detection (SSD): Towards to provide future SSD-based research," *MASAUM J. Basic Appl. Sci.*, vol. 1, no. 2, 2009.

[55] Mohd Nizam Omar, Lelyzar Siregar, and Rahmat Budiarto, "Hybrid stepping stone detection method," *2008 First Int. Conf. Distrib. Framew. Appl.*, pp. 134–138, Oct. 2008.

[56] J. I. Gilbert, "Scalable wavelet-based active network stepping stone detection," (Master's thesis, Air Force Institute of Technology, Air University, USA, 2012.

[57] J. I. Gilbert, D. J. Robinson, J. W. Butts, and T. H. Lacey, "Scalable wavelet-based active network detection of stepping stones," in *SPIE Defense, Security, and Sensing*, 2012, p. 84080I--84080I.