SEMINAR AND TECHNICAL COMMUNICATION

Topic: Effective algorithm to detect stepping stone intrusion by removing packet outliers of packet RTTs

NAME: Atharv Rajkumar Nalwade
Roll No: TC146
Class: TE-1

Research guide:
Prof. S. P. Mone

❖ CONTENTS:

❖ Introduction :

A network intrusion is an unauthorized penetration of a computer in your enterprise or an address in your assigned domain. It can be active or passive. Some intrusions are simply meant to let you know the intruder was there, defacing your Web site with various kinds of messages or crude images. Others are more malicious, seeking to extract critical information on either a one-time basis or as an ongoing parasitic relationship that will continue to siphon off data until it's discovered

Attackers on the Internet often launch network intrusions through compromised hosts, called stepping-stones, in order to reduce the chance of being detected. In a stepping-stone attack, an intruder uses a chain of hosts on the Internet as relay machines and remotely log in these hosts using tools such as telnet, rlogin, or SSH. A benefit of using stepping-stones to launch attacks is that intruders can be hidden by a long interactive session and knowing the actual origin of the attack is hard to determine

❖ Motivation:

In the 21st century where we are surrounded with an invisible web of internet care should be taken that one does not access your any sorts of private data without your authorization that's where the concept of security clears its picture in the field of computer networks. Any illicit behaviour on a digital network is known as a network intrusion. Any of the following can be considered an intrusion –

- Malware, sometimes known as ransomware, is a type of computer virus.
- Attempts to obtain unauthorized access to a system
- DDOS (Distributed Denial of Service) attacks
- Destruction of cyber-enabled equipment
- Employee security breaches that are unintentional (like moving a secure file into a shared folder)
- Untrustworthy users, both within and external to your company

- One such intrusion is stepping stone intrusion where the attacker uses a long chain of compromised hosts to gain access to a remote host which can be prevented effectively up to 85% using the algorithm described in the paper
- Phishing campaigns and other methods of deceiving consumers with ostensibly genuine communication are examples of social engineering assaults.

❖ Problem Statement :

To detect steeping-stone intrusion in a computer networks system using a effective algorithms by removing outliers of packet RTT.

❖ Scope:

The algorithms suggested by the author in the research paper titled “Effective Algorithms to Detect Stepping-Stone Intrusion by Removing Outliers of Packet RTTs” by Lixin Wang, Jianhua Yang, Michael Workman, and Pengjun Wan propose an efficient way to eliminate most of the possible RTT outliers of the packets captured in the Internet environment. Then an efficient SSI detection algorithms used by mining network traffic using an improved version of k-Means clustering. The proposed detection algorithm for SSI is accurate, effective, and efficient in the context of the Internet. Effective rate of our proposed SSI detection algorithm is higher than 85.7% in the context of the Internet.

Literature Survey

❖Effective Algorithms to Detect Stepping-Stone Intrusion by Removing Outliers of Packet RTTs by Lixin Wang , Jianhua Yang, Michael Workman, and Pengjun Wan

- An effective method to detect stepping-stone intrusion (SSI) is to estimate the length of a connection chain. This type of detection method is referred to as a network-based detection approach. Existing network-based SSI detection methods are either ineffective in the context of the Internet because of the presence of outliers in the packet round-trip times (RTTs) or inefficient, as many packets must be captured and processed. Because of the high fluctuation caused by the intermediate routers on the Internet
- Advantages:

I. The result of the algorithms makes most of the known network-based detection methods for SSI ineffective in the Internet environment

- Disadvantages:

I. During conduction of the experiment in real time it was found that he RTT dataset collected from the connection chain of length two achieves the smallest standard derivation on six of the seven datasets that were captured (only the output generated on algorithms of the dataset is incorrect)

❖ Matching TCP/IP Packets to Detect Stepping-Stone Intrusion:

- Author suggest using the technique of matching the TCP/IP Protocols using a Step-Function method, to detect network attackers from using a long connection chain to hide their identities when they launch attacks. The objective of the method is to estimate the length of a connection chain based on the changes in packet round trip times. The key point to compute the roundtrip time of a connection chain is to match a Send and its corresponding Echo packet
- ADVANTAGES:

I. The ability to detect intruders in real-time,

II. The ability to handle encrypted terminal sessions,

III. The ability to estimate the length of a chain accurately, and

IV. The ability to tolerate network traffic fluctuation, network load, and workload of chained hosts.

- DISADVANTAGES

  I. We must be able to monitor a packet throughout a connection session in order for this approach to work

  II. If the fluctuation of a connection is higher than the additional time to connect to the next host, we will need a better approach to detect the additional host.

❖ Mining Network Traffic with the k-Means Clustering Algorithm for Stepping-Stone Intrusion Detection:

- In the research paper “Mining Network Traffic with the k-Means Clustering Algorithm for Stepping-Stone Intrusion Detection” by Lixin Wang, Jianhua Yang, Xiaohua Xu, and Peng-Jun author suggest using k-means clustering algorithm which can accurately determine the length of a connection chain without requiring a large number of TCP packets being captured and processed, so it is more efficient.
  This algorithm is also easier to implement than all existing approaches for stepping-stone intrusion detection. The effectiveness, correctness, and efficiency of our proposed detection algorithm are verified through well-designed network experiments
- Advantages:

I. Proposed detection algorithm using the k-means clustering can accurately determine the connection chain length

II. The proposed algorithm does not require capturing and processing a large number of TCP packets. Therefore, our proposed detection algorithm for SSI is efficient

III. Easy to implement

- <u>Disadvantage:</u>

I. The detection methods based on the k-means clustering require that there should be as less outlier values of round-trip times as possible. Therefore, additional algorithms are needed to remove these outlier values of round-trip times from the input file of the detection algorithm

❖ Detecting Stepping Stones:

- We develop an efficient algorithm for detecting stepping stones by monitoring a site's Internet access link. The algorithm is based on the distinctive characteristics (packet size, timing) of interactive traffic, and not on connection contents, and hence can be used to find stepping stones even when the traffic is encrypted.
- Advantages:

I. The algorithm runs on a site's Internet access link. It proves highly accurate, and has the major advantage of ignoring the data contents of the connections, which means both that it works for encrypted traffic such as SSH, and that the packet capture load is greatly diminished since the packet filter need only record packet headers.

- DISADVANTAGES:

I. Algorithm fails when there is the large number of legitimate stepping stones that users routinely traverse for a variety of reasons

II. It was found that the timing-based algorithm missed a stepping stone simply because the connections were exceedingly short

❖ Stepping Stone Detection for Tracing Attack Sources in Software-Defined Networks:

- Study aims to adapt some of the existing stepping stone detection and anti evasion techniques to software-defined networks which use network function virtualization. We have implemented the stepping-stone detection techniques in a simulated environment and use sFlow for the traffic monitoring at the switches. We evaluate the detection algorithms on different network topologies and analyze the results to gain insight on the effectiveness of the detection mechanisms
- ADVANTAGES

I. Applicable to non-interactive network traffic

II. Effective detection of anomaly due to chaff

- DISADVANTAGES

I. Lack of incorporation of specifics of SDN and NFV environment
II. Limited scalability of the solution
III. Use of data store leads to lot of disk accesses
IV. Continuous bandwidth consumption

❖ Software/Hardware Requirements

- A device to access the Internet
- Internet connection

❖ PROPOSED TECHNOLOGY

- In this paper, we first propose an efficient algorithm to eliminate most of the possible outliers of the RTTs of the packets captured in the Internet environment. We then develop an efficient detection algorithm for SSI by mining network traffic using an improved version of k-Means clustering
- The proposed detection algorithm for SSI is accurate, effective, and efficient in the context of the Internet

ALGORITHMS

Algorithm 1: An efficient algorithm for removing the outliers of packet RTTs Input: a TXT file with two columns (including packet timestamps and the packet type) obtained from the packets captured in the Internet environment

- Output: a TXT file output.txt containing the packet RTTs with most of the RTT outliers removed

Step 1) for each of the first five Echo packets, compute the time difference between the Echo packet and its immediate prior Send packet; write this time difference to the file output.txt for each of these five Echo packets

Step 2) compute the average of all the RTTs in the file output.txt and store the value in the variable average

Step 3) for the 6th Echo packet, compute the time differences between this Echo packet and its prior five Send packets; assume these Send packets are represented by Send1, Send2, Send3, Send4, and Send5, respectively

(a) let timediff1 denote the difference between this Echo packet and Send1. Write timediff1 to the file output.txt if timediff1 is less than 10 average; recompute the average of all the RTTs in output.txt and then update the value of average

(b) let timediff2 denote the difference between this Echo packet and Send2. Perform the same process for timediff2 as done in the above

(c) let timediff3 denote the difference between this Echo packet and Send3. Perform the same process for timediff3 as done in the above

(d) let timediff4 denote the difference between this Echo packet and Send4. Perform the same process for timediff4 as done in the above

(e) let timediff5 denote the difference between this Echo packet and Send5. Perform the same process for timediff5 as done in the above

Step 4) repeat Step 3 for each of the remaining Echo packet until all processed

Algorithm 2 An effective algorithm for SSI detection:

Input: dataset-1, dataset-2, and dataset-3

- Output: The connection chain(s) that are sessions possibly manipulated by malicious hackers

Step 1) call the 2-Means clustering algorithm on dataset-1. Assume1 represents the standard derivation outputted based on Eq. (2) using the two clusters obtained at the end of the 2-Means clustering algorithm execution

Step 2) Call the 2-Means clustering algorithm on dataset-2. Assume2 represents the standard derivation outputted based on Eq. (2) using the two clusters obtained at the end of the 2-Means clustering algorithm execution

Step 3) Call the 2-Means clustering algorithm on dataset-3. Assume 3 represents the standard derivation outputted based on Eq. (2) using the two clusters obtained at the end of the 2-Means clustering algorithm execution

❖ Patrial Results

- The network experiment conducted by the author shows that the effective rate of our proposed detection algorithm for SSI is 6/7$\approx$ 85.7% in the Internet environment

❖ CONCLUSION

- The proposed algorithm is highly effective in the internet environment
- Future research in this direction, we plan to improve our SSI detection algorithms so that they will resist session manipulation by intruders using hacking tools, such as time-jittering and/or meaningless chaff perturbation, in the Internet environment.

❖ REFRENCES:

I. D. Bhattacharjee, A. Gurtov and T. Aura, "Watch Your Step! Detecting Stepping Stones in Programmable Networks," ICC 2019 - 2019 IEEE International Conference on Communications (ICC), 2019.

II. J. Yang, Y. Zhang, R. King and T. Tolbert, "Sniffing and Chaffing Network Traffic in Stepping-Stone Intrusion Detection," 2018 32nd International Conference on Advanced Information Networking and Applications Workshops (WAINA), 2018.

III. M. A. Gamarra, S. Shetty, D. M. Nicol, L. Njilla and O. R. Gonzalez, "Modelling Stepping Stone Attacks with Constraints in Cyber Infrastructure," 2019 IEEE Global Communications Conference (GLOBECOM), 2019.

IV. Marco Gamarra; Sachin Shetty; Oscar Gonzalez; David M. Nicol; Charles A. Kamhoua; Laurent L. Njilla, "Analysis of Stepping-Stone Attacks in Internet of Things Using Dynamic Vulnerability Graphs," in Modelling and Design of Secure Internet of Things, IEEE, 2020.

V. J. Yang, B. Lee, S. S.-H. Huang, Monitoring network traffic to detect stepping-stone intrusion, in Proc. of the 22nd IEEE International Conference on Advanced Information Networking and Applications, Okinawa, Japan, 2008.

VI. Wang, L., Yang, J. “A research survey in stepping-stone intrusion detection”. J Wireless Communication Marathwada Mitra Mandal’s College of Engineering 2022-23 Network 2018, 276 (2018).

VII. L. Wang, J. Yang, M. Mccormick, P. -J. Wan and X. Xu, "Detect Stepping-stone Intrusion by Mining Network Traffic using k-Means Clustering," 2020 IEEE 39th International Performance Computing and Communications Conference (IPCCC), 2020.

VIII. Zhang and V. Paxson, Detecting stepping-stones, in Proc. of the 9th USENIX Security Symposium,Denver, CO, USA, 2000.

IX. Zhang, Yin and Vern Paxson. “Detecting Stepping Stones.” USENIX Security Symposium (2000).

X. Blum, Avrim & Song, Dawn & Venkataraman, Shobha. “Detection of Interactive Stepping Stones: Algorithms and Confidence Bounds”. (2004)

Step 4) If б1=min (б1, б2, б3); dataset-2 and dataset-3 are sessions possibly manipulated by malicious hackers

Step 5) If б2 = min (б1, б2, б3); dataset-1 and dataset-3 are sessions possibly manipulated by malicious hackers

Step 6) If б3 = min (б1, б2, б3), dataset-1 and dataset-2 are sessions possibly manipulated by malicious hacker