A NOVEL APPROACH TO REDUCE FALSE POSITIVES IN SPAM ZOMBIE DETECTION SYSTEMS

[1]*MADHUBABU NALLURI,* [2]*M.VAZRALU*
[1]M.Tech CSC Dept, MRCET, Hyderabad, India,
Email: nallurimadhu90@gmail.com.
[2]Associate Professor, CSE Dept, MRCET,
Email: vazram4u@gmail.com.

ABSTRACT- *Hardware equipments which are also known as external devices are becoming more complicated because they are acting as mediator between threats and our Pc's. Different types of security attacks are cause of these external devices only. These external devices are also referred as compromised machines. Some main attacks they are causing is malware spreading and spamming, using identity and also DDoS. Attackers are capable of getting key economic incentive with the help of spamming. This leads to get huge compromised machines count. Here we are intended to detect the number of compromised machines that are becoming initiative to spamming operations in a network. Such operation is also known as spam zombies. Here we introduced a detection system to find out spam zombie operations known as SPOT. These SPOT continuously used to keep eye on data transactions that the messages which are transmitting outside of a network. Based upon Sequential Probability Ratio Test SPOT was developed. It is a effective statistical tool. The operations of this tool are to find out positive and as well as false negative error rates in a network. Based upon our researches and test study cases proven that SPOT is a perfect and extraordinary in detecting compromised machines which are intended to do zombie operations in a network automatically.*
Moreover SPOT is capable of detecting 95% of compromised machines out of 100. That means accurately it can detect almost every spamming activities compromised machines. When compared to various spam zombie detection algorithms SPOT performance is unique. Because rest of them are based upon number and percentage of messages which are spammed. But SPOT uses a statistical too to detect zombie operations. So out of various spam detecting algorithms SPOT has a unique nature that keeps it out standing in detecting spam zombies. Working and operations wont effect network's efficiency or any security breaches like identity.
Index Terms—*Compromised machines, spam zombies,*

1. INTRUDUCTION

Recently for attackers there is a new concern formed that is compromised machines in the network. These machines are large in count. These machines are becoming initiative for the security attacks. Those are of malware spreading and spamming, using identity and also DDoS. These machines are of two types in nature. These are classified based upon their nature of behaving in the Internet. The first type is sheer volume and next one comes to means second one is widespread. There are so many counter attacks to the actions of these machines but those are less effective so they cannot perform accurate operation on these machines. At the same time identification and as well as removing compromised machines is becoming key challenge to the admin people in a network. These networks may be of any size but it doesn't matters. In this paper, we introduced a detection system to find out spamming operations in a network. Such operation are also known as spam zombies Based on researches one thing was observed that is these compromised machines are making economic incentive which was critical. It was for the controller's cause of this huge number of spamming operations is done cause of these compromised machines. A recent survey finds out that large number of spamming botnets and as well as botnet patterns are from

the largest email service provider. These were detected based upon the spam messages which were used and stored in a network. Detecting compromised machines is better than the detecting spamming botnets in a network while online process is going. Because for these spamming botnets origin is compromised machine so reducing these machines is better. Find outing these compromised machines which were supposed to do spam zombies in a network is a typical procedure. Even though previous detecting systems cannot perform these operations in a network while transmitting data or based on outgoing messages. Our developed detecting system is capable of find outing these spam zombies in a network efficiently and effectively. The previous detecting approaches cannot perform its operations in online. So those are lack of requirement online detection requirement. There may be threat of sequential detection problem. This will rise because of continuously monitoring of the messages which were spammed. In this paper, we introduced a spam zombie detection system, known as SPOT. These SPOT continuously used to keep eye on data transactions that the messages which are transmitting outside of a network. by monitoring outgoing messages. Based upon Sequential Probability Ratio Test SPOT was developed. Wald developed Sequential Probability Ratio Test in his seminal work.

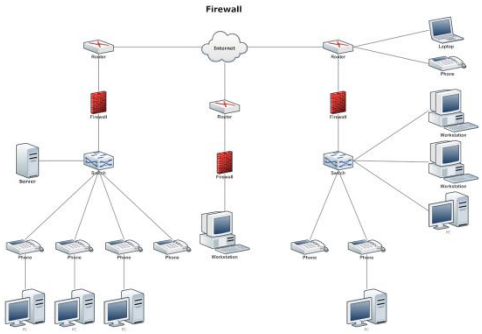

Fig.1. Network model

This Sequential Probability Ratio Test is a efficient statistical tool which is used for couple of things mainly. Those things deals with the difference machine are compromised or not. At the same times as the outgoing messages are going sequentially or not.

Sequential Probability Ratio Test number of is a efficient statistical tool has some features which are desirable. It decreases observations count needed to declare a judgment in the group of all sequential and non sequential tests which are statistical without any error rising. SPOT detection system is time efficient one also because it can detect the compromised machine in less spam when compared to other detecting System. By user-defined thresholds bounds false positive and negative probabilities which are of SPRT. It is the user's choice to find out the desired thresholds which he wants of SPOT system. These are used to keep the false positive and false negative rates controlling in the system. In this paper, accurately SPOT can detect almost every spamming activities compromised machine. When compared to various spam zombie detection algorithms SPOT performance is unique. Because rest of them are based upon number and percentage of messages which are spammed. But SPOT uses a statistical too to detect zombie operations. DBSpam is proxy-based outgoing spam detection system. It works based on the packet symmetry property. It cannot used to identify the large number of internal zombies. The proposed system is to identify all the zombies of the internal network. It can be applied on malware propagation detection. The proposed system is Spam Zombie Detection System (SPOT).

2. SPAM ZOMBIE DETECTION ALGORITHMS

In this section, we will develop three spam zombie detection algorithms. The first one is SPOT, uses SPRT which is known as statistical tool which we have discussed before. The impact on SPOT by SPRT parameters were explored in spam zombie detection concept. The other two spam zombie detection algorithms are developed on some of the features as a origin. Those are spam messages count and the spam messages percentage of sent from an internal machine, respectively.

SPOT Algorithm

SPOT is designed based on the statistical tool SPRT we discussed in the last section. Detecting spam zombies in SPOT, we consider C1 as a detection and C0 as normality. That is, C1 is true if the concerned machine is compromised, and C0 is true if it is not compromised. In addition, we let Xi = 1 if the I th message from the concerned machine in the network is a spam, and Xi = 0 otherwise. Recall that SPRT requires four configurable parameters from users, namely, the desired false positive probability α, the desired false negative probability β, the probability that a message is a spam when C1 is true (θ1), and the probability that a message is a spam when C0 is true (θ0). We discuss how users configure the values of the four parameters after we present the SPOT algorithm. Based on the user-specified values of α and β, the values of the two boundaries A and B of SPRT are computed. In the following, we describe the SPOT detection algorithm. Algorithm 1 outlines the steps of the algorithm. When an outgoing message arrives at the SPOT detection system, the sending machine's IP address is recorded, and the message is classified as either spam or non spam by the spam filter. For each observed IP address, SPOT maintains the logarithm value of the corresponding probability ratio ∧n, whose value is updated as message n arrives from the IP address. Based on the relation between ∧n and A and B, the algorithm determines if the corresponding machine is compromised, normal, or a decision cannot be reached and additional observations are needed.

Algorithm 1. SPOT spam zombie detection system
1: An outgoing message arrives at SPOT
2: Get IP address of sending machine m
3: // all following parameters specific to machine m
4: Let n be the message index
5: Let Xn =1 if message is spam, Xn = 0 otherwise
6: if (Xn == 1) then
7: // spam, 3
8: ∧n+ = lnθ1/θ2
9: else
10: // nonspam
11: ∧n+ = ln 1-θ1/1-θ2
12: end if
13: if (∧n ≥ B) then
14: Machine m is compromised. Test terminates for m.
15: else if (∧n ≤A) then
16: Machine m is normal. Test is reset for m.
17: ∧n =0
18: Test continues with new observations
19: else
20: Test continues with an additional observation
21: end if

3. SYSTEM DEVELOPMENT

Step1: Packet Capturing
Opens the network interface card for incoming packets.
Packet capturing is a continuous process.
Field Extraction (SIP, SP, DIP, DP, Protocol) were gathered for each packet.
Step2: Session Management
Each five session parameters are unique.
Identifying the session relevant packets. Extracts and aggregates to the unique sessions.
Identify the packets with payloads from all the packets of the session.
Step3: Packet Payload Extraction
Opens the outbound packet.
Parses the header by header up to the application payload.
Extracts the payload.
Step4: Signature Verification
Signatures should be written to identify the outgoing spam mails.
Spam signatures on HTTP and SMTP
The signatures are cross verified to identify the outgoing packets.
If the signature matches, The SPOT logs the attack and sends to the administrator.
The SPOT detection algorithm is based signature verification is done.
Step 5: Administrator Interface
Login: Provides the admin to login.
Verification of Logs
The administrator verifies the attacks or spam propagation information.

Adding of Signatures
Administrator can add a signature to the detection system.
Signature Viewing
Administrator is allow to modify the existing signatures.
Logs provides the information about the internal zombies such as (IP, SPAM Instance, SignatureID).

4. RELATED WORK

We explored related work in this section which was about compromised machines detection and reduction. Initially we kept focused on using spamming activities to find out bots. After that we explained a number of trails in normal botnets detecting. From a huge e-mail service provider data base the e-mail messages collected. And there are these features of spamming botnets , botnets size and patterns of botnets. A recent survey finds out that large number of spamming botnets and as well as botnet patterns are from the largest email service provider. These were detected based upon the spam messages which were used and stored in a network. These spamming Zombies are done based on spam messages clustering into spam campaigns present at the provider with the help of near-duplicate content and clustering embedded URLs. However, for large e-mail service these approaches are suited better. And the email providers have to understand the aggregate global features of spamming botnets rather than individual deployment in networks find out compromised machines. The early developed detecting approaches cannot perform its operations in online. So those are lack of requirement online detection requirement We aim to develop a tool that is a perfect and extraordinary in detecting compromised machines which are intended to do zombie operations in a network automatically. There is a effective tool DB Spam which is used to detect activities proxy-based spamming in a network based upon packet symmetry property of such activities. We intend to identify all types of compromised machines involved in spamming, not only the spam proxies that translate and forward upstream non-SMTP into SMTP commands to downstream mail servers. In the following, we discuss a few schemes on detecting general botnets. Worm is a self propagating program. Worm carries the attack programs from one system to the other system in the internet. When a system is affected with worm, the system is called as “Zombie”. Worm can scan the neighbor systems in the network. It can attack the vulnerable systems. The worm can be in the polymorphic nature. It can change its nature to evade the security systems. When security system fails to identify the worm propagation from internet to the internal network, The worm easily spreads in the internal network. When the zombies are created by the worm, they try to open “backdoor” (a tiny server program, to provide the control to the attacker). The worms that spreads through mails, scans the zombie system’s dictionary for the email-ids that used by the system user. It sends worm attached email to the email-ids found on the zombie. These can be identified only through the outgoing messages from the internal network. DBSpam is proxy-based outgoing spam detection system. It works based on the packet symmetry property. It cannot used to identify the large number of internal zombies.

BotHunter, developed is used to detect compromised machines by correlating the IDS dialog trace in a network. It was developed based on the observation that a complete malware infection process has a number of well-defined stages including inbound scanning, exploit usage, egg downloading, outbound bot coordination dialog, and outbound attack propagation. By correlating inbound intrusion alarms with outbound communications patterns, BotHunter can detect the potential infected machines in a network. Unlike BotHunter which relies on the specifics of the malware infection process, SPOT focuses on the economic incentive behind many compromised machines and their involvement in spamming. An anomaly-based detection system named BotSniffer identifies botnets by exploring the spatial-temporal behavioral similarity commonly observed in botnets. It focuses on IRC-based and HTTP-based botnets. In BotSniffer, flows are classified into groups based on the common

server that they connect to. If the flows within a group exhibit behavioral similarity, the corresponding hosts involved are detected as being compromised. BotMiner is one of the first botnet detection systems that are both protocol and structure independent. In BotMiner, flows are classified into groups based on similar communication patterns and similar malicious activity patterns, respectively. The intersection of the two groups is considered to be compromised machines. Compared to general botnet detection systems such as BotHunter, BotSniffer, and BotMiner, SPOT is a lightweight compromised machine detection scheme, by exploring the economic incentives for attackers to recruit the large number of compromised machines. As a simple and powerful statistical method, Sequential Probability Ratio Test has been successfully applied in many areas. In the area of networking security, SPRT has been used to detect port scan activities, proxy-based spamming activities, anomaly-based botnet detection, and MAC protocol misbehavior in wireless networks.

5. CONCLUSION

In this paper, we introduced a detection system to find out spam zombie operations known as SPOT. These SPOT continuously used to keep eye on data transactions that the messages which are transmitting outside of a network. Based upon Sequential Probability Ratio Test SPOT was developed. It is a effective statistical tool. The operations of this tool are to find out positive and as well as false negative error rates in a network. SPOT is effective cause of it decreases the observations count to find out spam zombie operations. That means accurately it can detect almost every spamming activities compromised machines. When compared to various spam zombie detection algorithms SPOT performance is unique. Because rest of them are based upon number and percentage of messages which are spammed. The SPOT is an affective systems to detect the more number of internal zombies. Identifying the internal zombies is the major task. The SPOT is used to identify the spam propagation and malware propagations.

6. REFERENCES

[1] P. Bacher, T. Holz, M. Kotter, and G. Wicherski, "Know YouREnemy:TrackingBotnets," http://www.honeynet.org/papers/bots, 2011.

[2] Z. Chen, C. Chen, and C. Ji, "Understanding Localized-Scanning Worms," Proc. IEEE Int'l Performance, Computing, and Comm. Conf. (IPCCC '07), 2007.

[3] R. Droms, "Dynamic Host Configuration Protocol," IETF RFC 2131, Mar. 1997.

[4] Z. Duan, Y. Dong, and K. Gopalan, "DMTP: Controlling Spam through Message Delivery Differentiation," Computer Networks, vol. 51, pp. 2616-2630, July 2007.

[5] Z. Duan, K. Gopalan, and X. Yuan, "Behavioral Characteristics of Spammers and Their Network Reachability Properties," Technical Report TR-060602, Dept. of Computer Science, Florida State Univ., June 2006.

[6] Z. Duan, K. Gopalan, and X. Yuan, "Behavioral Characteristics of Spammers and Their Network Reachability Properties," Proc. IEEE Int'l Conf. Comm. (ICC '07), June 2007.

[7] G. Gu, R. Perdisci, J. Zhang, and W. Lee, "BotMiner: Clustering Analysis of Network Traffic for Protocol- and Structure-Independent Botnet Detection," Proc. 17th USENIX Security Symp., July 2008.

[8] G. Gu, P. Porras, V. Yegneswaran, M. Fong, and W. Lee,
"BotHunter: Detecting Malware Infection through Ids-Driven
Dialog Correlation," Proc. 16th USENIX Security Symp., Aug. 2007.

[9] G. Gu, J. Zhang, and W. Lee, "BotSniffer: Detecting Botnet Command and Control Channels in Network Traffic," Proc. 15th Ann.

Network and Distributed System Security Symp. (NDSS '08), Feb. 2008.

[10] N. Ianelli and A. Hackworth, "Botnets as a Vehicle for Online Crime," Proc. First Int'l Conf. Forensic Computer Science, 2006.

Authors:

MADHUBABU NALLURI is an M.Tech student in MRCET, Hyderabad. He received B.Tech Degree in Computer Science and Engineering in 2010 from JNTU, Hyderabad. He is interested in the field of Computer Networks and Network Security.

M.VAZRALU having *7+* years of teaching experience, currently he is working as Associate Professor in CSE Department in MRCET, Hyderabad. He has completed M.Tech in Computer Science and Engineering, currently he is doing PhD. His interested research areas are Software engineering, Network Security and operating system.