Detecting spam zombies by monitoring outgoing messages

ASSIGN BUSTER

Abstract:

Compromised machines are one of the key security threats on the Internet. They are often used to launch various security attacks such as spamming and spreading malware [15]. Given that spamming provides a key economic incentive for attackers to recruit a large number of compromised machines, we focus on the detection of the compromised machines in a network that are involved in the spamming activities, commonly known as spam zombies [12].

Introduction:

As the use of internet increased in the era of science and technology the problem of spam has also been increased. There are multiple ways in which spam takes place we would like to discuss the spam that is passed through messages specifically through emails. When these spam mails are passed into the system these makes the system compromised and the data in the network can be stolen or lost these kind of spamming is more concern to the industry or any kind of organization where privacy is the key aspect in this competitive world.

Spam:

Spam can be defined as “ Simple Pointless Annoying Messages”. According to US Federal Trade Commission (FTC) spam is defined as, “ any commercial electronic mail message sent, often in bulk, to a consumer without the consumer’s prior request or consent” [1]. A recent study conducted by SMX an email security provider the percentage of spam is about 80% approx. And

the average size of these spam messages in 16 Kb [2]. The results above indicate the seriousness of the problem. There are several techniques proposed and employed in filtering these spam messages such as Machine learning techniques like Neural networks(NN), Support Vector Machines(SVM), Naive Bayes Classifier. Some techniques are based on probability and others on architectural. According to Anil Kumar Gupta along with two others in his research paper stated that training SVM is easy compared to NN because NN takes more time to train than SVM and NN will not offer binary classification mechanism where has SVM does that technique to verify the legitimate of the email [2]. Rafiqul Islam in his research proposed an architecture for spam filtering based on support vector machine [3]. T. Hamsapriya along with three others in 2014 proposed Filtered Bayesian Learning technique to increase the performance of the naÃ¯ve Bayes classifier. These all techniques have contributed in controlling spam to very much extent [4].

Spam Zombies:

A machine is said to be compromised if it is successfully exploited by the attacker. These machines are used to launch various attacks in the network. These compromised machines are called zombies. The machine is made compromised when an attacker sends a spam mail to the targeted system and made a zombie [5].

Spam in Messages:

Today’s communication mostly happening through messages that are sent electronically through email or text messages in mobile. Our main

concentration is confined to messages that are going out through a network and coming into the network that are emails. Body message based spam detection is employed in larger servers but in a research conducted by Shukor Bin Abdul Razak in 2013 showed that the feature can be manipulated and has several issues such as Manipulation of lexical patterns, efficiency, future trends. So he proposed an email header technique that has a potential in filtering spam efficiently [6]. In 2015 Wazir Zada Khan along with three others stated that the detection criterion for web spam is substantially different, so, the email spam coming from botnets cannot be handled by the web spam detection techniques. Then they proposed architecture for email spam botnet detection [7].

Algorithm:

SPOT detection algorithm is used to detect spammers. Before proposing SPOT detection techniques there are few works which happened in detecting spam zombies. S. Yuvaraj in 2013 came up with a four module system which consists of compose mail process, Filter spam detect, IP capture, Extraction of payloads and payload disassembly and this algorithm is called has semantic aware statistical algorithm (SAS) [8]. But this algorithm fails to catch spammers but detects spam zombies. The research also proposed algorithms in the field of botnet which is usually called a group of computers affected with malware and controlled without the notice of administrator. To control these botnets issues Guofei Gu from Georgia institute of technology came up with bot hunter based on correlation between inbound and outbound communication. This system also uses intrusion detection system(IDS) to find out the compromised machines in the network [9]. Later

in 2008 again Guofei Gu along with Wank Lee proposed another technique called botsniffer in which he extended his research in detecting compromised servers depending on the behavioral similarity in a single group of connected computers [10]. After all these works with different techniques people came up with standard algorithm called SPOT applied in detecting spam zombies which functions by monitoring outgoing messages in the network. Z. Duzan in 2009 proposed an algorithm using Sequential Probability Ratio Test(SPRT) depending on the mathematical value of the SPRT the email is as spam or not spam [11]. But he ignored the impact of dynamic IP address on the data which is considered for analysis. His research is as limitations since the algorithm is based on probability analysis and the messages arrived assumed to independent of each other but this may not be the practical scenario. Spam filters are used to detect the spam emails but these filters are not 100 percent efficient. Later in 2012 Pen cheng along with Z. Duzan modified his algorithm they introduced two more terms called count threshold and percentage threshold to calculate the impact of dynamic IP address[12]. In continuation to the work of Z. Duzan, Ar. Arunachalam along with his two students in 2013 added two more modules and applied Z. Durzan techniques in calculating the impact of dynamic IP address to entire system by adding user interface module and spam zombie detection module where he has reset the values of the captured spam emails continuously [13]. Similar work has been done by R. Vasanth Kumar and K. Ravi Kumar in 2013 they modified the existing algorithm using the IP address of the sending machine and introducing a new term called message index[14]. Parvathi Bhadre and Deepthi Gothawal in 2014 proposed a new method using SPOT detection algorithm consisting of four modules namely

virus checks, Spam Checks and Spam filter, blocking of spammers using SPOT and Recovery [15]. But their research does not talk anything about the impact spam mails generated using dynamic IP address. In 2015 Anupsingh Thakur and Prof. Praful Sambhare conducted a survey on spamming and detection control through various methods like SVM, Domain key integrated mail system(DKIMS) and SPOT detection system defined how SPOT is accurate in detecting Spams [16].

Conclusion:

Brief review on spam, spam zombies, spam in messages, algorithm used and the previous works done are explained. We in our project intending to come up with improved algorithm that could effectively tackle the limitations of the previous works.

References:

1. D. C. Washington, "' Unsolicited commercial e-mail' before the SUBCOMMITTEE ON TELECOMMUNICATIONS, TRADE AND CONSUMER PROTECTION of the COMMITTEE ON COMMERCE UNITED STATES HOUSE OF REPRESENTATIVES," 2013. [Online]. Available: https://www. ftc. gov/sites/default/files/documents/public_statements/prepared-statement-federal-trade-commission-spamming/spamtestimony1103. pdf. Accessed: Mar. 3, 2017.
2. A. G. Kakade, P. K. Kharat, and Anil Kumar Gupta, " Spam filtering techniques and MapReduce with SVM: A study," *2014 Asia-Pacific Conference on Computer Aided System Engineering (APCASE)* , vol. 14666087, pp. 59-64, Feb. 2014.

3. R. I. M, W. Zhou, and M. U. Choudhury, “ Dynamic Feature Selection for Spam Filtering Using Support Vector Machine,” *6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)* , vol. 9864217, Jul. 2007.
4. H. T, L. S. P, K. R. D, and R. C. M, “ SPAM CLASSIFICATION BASED ON SUPERVISED LEARNING USING MACHINE LEARNING TECHNIQUES,” *ICTACT Journal on Communication Technology* , vol. 02, no. 04, pp. 457-462, Dec. 2011.
5. A. Rajagopal and A. P. P, “ SPOT- e-mail Spam zombie detection system,” *International Journal of Innovative Research in Computer and Communication Engineering* , vol. 2, no. 1, pp. 664-669, Jan. 2012. [Online]. Available: https://www. rroij. com/open-access/spot-email-spam-zombie-detection-system. php? aid= 48276. Accessed: Mar. 3, 2017.
6. S. Bin Abd Razak and A. F. Bin Mohamad, “ Identification of spam email based on information from email header,” *2013 13th International Conference on Intellient Systems Design and Applications* , pp. 347-353, Oct. 2014.
7. W. Z. Khan, M. K. Khan, F. T. Bin Muhaya, M. Y. Aalsalem, and H.-C. Chao, “ A comprehensive study of Email Spam Botnet detection,” *IEEE Communications Surveys & Tutorials* , vol. 17, no. 4, pp. 2271-2295, Jul. 2015.
8. Y. M. S. S., “ An effective defense against compromised machines by sas worm detection,” *International Journal of Computer Science and Management Research* , pp. 33-37, 2013.

9. G. Gu, P. Porras, V. Yegneswaran, M. Fong, and W. Lee, “ BotHunter: Detecting Malware Infection Through IDS-Driven Dialog Correlation,” *16th USENIX Security Symposium* , pp. 167-182, 2007.

10. G. Gu, W. Lee, and J. Zhang, “ Botsniffer: Detecting botnet command and control channels in network traffic,” *Proceedings of The 15th Annual Network and Distributed System Security Symposium (NDSS 2008)* , Feb. 2008.

11. Z. Duan, P. Chen, F. Sanchez, Y. Dong, M. Stephenson, and J. Barker, “ Detecting Spam Zombies by Monitoring Outgoing Messages,” *IEEE INFOCOM 2009* , 2009.

12. Z. Duan, P. Chen, F. Sanchez, Y. Dong, M. Stephenson, and J. M. Barker, “ Detecting Spam zombies by monitoring outgoing messages,” *IEEE Transactions on Dependable and Secure Computing* , vol. 9, no. 2, pp. 198-210, Mar. 2012.

13. A. Ar, V. V, and Y. V, “ Detecting Spam Zombies Using Spot Tool by Monitoring Outgoing Messages,” *International Journal of Advanced Research in Computer Science and Software Engineering* , vol. 3, no. 4, pp. 400-402, Apr. 2013.

14. V. kumar R and R. K. K, “ Recognizing Spam Zombies by Monitoring leaving Messages,” *International Journal of Engineering and Computer Science* , vol. 2, no. 11, pp. 3213-3216, Nov. 2013.

15. P. Bhadre and D. Gothawal, “ Detection and blocking of spammers using SPOT detection algorithm,” *2014 First International Conference on Networks & Soft Computing (ICNSC2014)* , pp. 97-101, Aug. 2014.

16. A. Thakur and P. Sambhare, “ Spamming and Detection Control: A Survey,” *INTERNATIONAL JOURNAL FOR RESEARCH IN EMERGING SCIENCE AND TECHNOLOGY* , vol. 2, no. 5, pp. 155-157, May 2015.