Blocking Spam Mails by Network Monitoring Approach

Sharanyaa. S
Department of Information technology
DMI college of Engineering,
Chennai, Tamilnadu, India

Dinesh Kumar. R
Department of Information technology
DMI college of Engineering,
Chennai, Tamilnadu, India

Johnci. J
Department of Information technology
DMI college of Engineering,
Chennai, Tamilnadu, India

Abstract - In today's internet mail server, spam content delivery is the most common issue. Most of the modern spam filtering techniques are deployed only on the receiver side. They are good at filtering spam for end users, but spam messages still keep wasting Internet bandwidth and the storage space of mail server. So to address this problem, we propose a technique to detect the spam mails and block it at the sender side itself. For this, Bro Intrusion Detection System and Adaptive Bloom Filter technique are used. Adaptive Bloom Filter Technique is used to test whether an element in the mail is the member of the set in database that contain related set of words. Bro Intrusion Detection system continuously monitor the network to find out any spam mail that passes through the network. In this way, it detects the spam mail and block the Mail id and the Network Address(IP address and MAC address). Network Monitor blocks the Recipient E-Mail address and Network Address for further activities.

Keywords: Bro Intrusion Detection System, Adaptive bloom filter, Spam filter, Network Address

I. INTRODUCTION

Email delivery has become an indispensable approach to communications in daily life. Due to its popularity and nearly zero cost, it is commonly exploited arryadvertisements, malware, phishing messages, and so on. According to a recent report from [1], around90%of email message sunsolicitemcial ones, namely spams. Email spam, also know as junk email or unsolicited bulk mail (UBE), is a subset of electronic spam involving nearly identical messages sent to numerous recipients by email. Clicking on links in spam email may send users to phishing web sites or sites that are hosting malware. Spam email may also include malware as script or other executable file attacement.Definition of spam usually includes the aspects that email is unsolicited and sent in bulk. Even though modern spam filtering techniques can filter out spam with high accuracy and rare recipients click the links in the spam messages, this problem persists because spammers can still capitalize on spamming due to the huge number of spam messages. Most practices to reduce spam are filtering on the receiver side. Common solutions include cloud-based mail security products such as Symantec MessageLabs and Google Postini, as well as personal security products such as Kaspersky Internet Security and Avast Internet Security. Mail clients such as Microsoft Outlook and Mozilla Thunderbird, as well as mail service providers, also support spam filtering. The solutions receive mail before filtering, so spanuning activities still exist, and spam messages still waste Internet bandwidth and the storage space of mail servers. The objectives of this project is to detecting spamming bots and block that spamming bots in sender side itself with a use of adaptive bloom filter and Bro Intrusion Detection System techniques.

Bro intrusion detection system (IDS) is a device or software application that monitors network or system activities for malicious activities or policy violations and produces reports to a management station. Some systems may attempt to stop an intrusion attempt but this is neither required nor expected of a monitoring system.

II. RELATED WORK

Recent work has leveraged botnet infiltration techniques to track the activities of bots over time, particularly with regard to spam campaigns. Building on our previous success in reverse-engineering C&C protocols, we have conducted a 4-month in-filtration of them egad botnet, beginning in October 2009. Our infiltration provides us with constant feeds on MegaD's com-plex and evolving C&C architecture as well as its spam oper-ations, and provides an opportunity to analyze the botmasters'operations. In particular, we collect significant evidence on theMegaD infrastructure being managed by multiple botmasters.

Further, Fire Eye's attempt to shutdown MegaD on Nov. 6, 2009, which occurred during our infiltration, allows us to gain an in-side view on the takedown and how MegaD not only survived it but bounced back with significantly greater vigor. In addition, we present new techniques for mining informa-tion about botnet C&C architecture: "Google hacking" to dig out egad C&C servers and "milking" C&C servers to extract not only the spectrum of commands sent to bots but the C&C's over-all structure. The resulting overall picture then gives us insight into MegaD's management structure, its complex and evolvingC&C architecture, and its ability to withstand takedown.

We also focus on characterizing spamming botnets by Leveraging both spam payload and spam server traffic properties. Towards this goal, we developed a spam signature generation frame-work calledAutoREto detect botnet-based spam emails and botnetmembership. AutoRE does not require pre-classified training data or white lists. Moreover, it outputs high quality regular expression signatures that can detect botnet spam with a low false positive rate. Using a three-month sample of emails from Hotmail, AutoRE successfully identified 7,721 botnet-based spam campaigns togetherwith 340,050 unique botnet host IP addresses. Our in-depth analysis of the identified botnets revealed several interesting findings regarding the degree of email obfuscation, prop-erties of botnet IP addresses, sending patterns, and their correlation with network scanning traffic. We believe these observations are useful information in the design of botnet detection schemes.

III. DETECTION OF SPAMMING BOTS

A. Problem Analysis

Not all bulk email is spam. Some is permission-based, meaning that the recipient has asked to receiver it. This occurs when a user at website voluntarily agrees-for example at the time of making purchase- to receive a newsletter or other email. Unlike spam, opt-in email usually provide a benefit such as free information or sale price. Sending unsolicited email to online customer who have not elected to receive information is considered spam.

Spam is rarely send directly by a company advertising itself. it's usually send by a"spammer",company in the business of distribution unsolicited email. An advertiser enters in to an agreement with a spammer, whose generates email advertisement to a group of unsuspecting recipient. The cost of spam is far less than postal bulk mailings.

Sometimes spamming may by recipient of spam often consider it an unwanted intrusion in their mailbox. Internet service provider(IPs).such as America online, onside spam to be a financial detect and an impediment to internet access because it can clog an ISP's available bandwidth spam has also been linked with fraudulent business schemes, chain letter, an offensive sexual and political messages.

You've probably noticed that much of the spam you receive involves deceptive practices. For example, spam for X-rated sites may be disguised with a personal subject header ("how come you did't write back").And you've no doubt noticed that a lot f the spam(We can help remove you from spam list")that comes you away is attempting to perpetuate some sort of scam-pyramid schemes, bogus stock offerings, pirated software, and quack health remedies.

B. Application of Adaptive Bloom Filter

A Bloom filter is a space-efficient probabilistic data structure, conceived by Burton Howard Bloom in 1970, that is used to test whether an element is a member of a set. False positive matches are possible, but false negatives are not; i.e. a query returns either "possibly in set" or "definitely not in set". Elements can be added to the set, but not removed (though this can be addressed with a "counting" filter). The more elements that are added to the set, the larger the probability of false positivesThe number of REAs in the outgoing mail messages, especially spam

messages, can be large, so an efficient data structure to store them is essential. So Bloom filter is used to maintain the REAs from each individual internal host.

Adaptive Bloom filter consists of an m-bit array to store n objects, which are the REAs of the outgoing mail messages in this work. The bits in the array are all initialized to 0. Each REA Ai of an outgoing message is stored into the Bloom filter for the host sending the message by setting the bits at the positions hl(Ai),h2(Ai), ... ,h_k(Ai) to 1, where hl,h2,'" ,h_k are k independent hash functions. Before storing REA it will check whether the same REA has been in the Bloom filter before storing AiNotice that a Bloom filter may be filled up, if too many REAs appear. An expiration mechanism is therefore necessary. A sensitive idea is marking the storage time of each REA. If the lifetime of an REA is longer than a pre-defined expiration time, the address will be purged. Spam detection process is done by using the list of spam related words stored in the database.

C. Detection of Spam Mails

In this module mailing process had been implemented. Initially the receiver id has to be typed in to the bar. Then the content will be selected from the system or typed on own. In this, a single person can send mail to any number of users of some server specified file content memory. The sender can even send multimedia mail also. Even the files regarding documents, web page links can also be sent. But we can't send the .exe extension file to anyone. Knowingly or unknowingly the sender may send spam mail. Mail can be sent to other mail id.Mail with attachment can be sent to other mail id and user can download these attachments.Received mail will show the sender mail id, time, date, and network id.In this module a single person can send one or more messages to other person. It is open chat, which means all authorized user can send or receive messages to/from others. This messages either spam or no spam. Spam means the Spam status updated in system.

The beginning of spam detection lies in this module. Mostly spam mails are used to attract the user or reader. We analyses the so many spam's mail and prepared the list of spam related words. The mail which is composed to send will be checked for malware before sending it and indicates the sender about the spam mail that they tried to send. Spam detection process is done by using the list of spam related words stored in a data base.

Initially the header is checked for spam;if spam is not found then the body of the mail is checked. If any malware is found then that particular mail will be blocked from sending. The report will be sent to the sender id about the spam detected in their mail and mentions the details why that mail is not allowed to send. If no spam is found in the mail, then the mail will be sent to the mentioned receiver in normal way. Pattern matching method is used to identify the meaningless mails. It is used to reduce the internet bandwidth and server memory.

D. Analysis Of Mail Id and Network Address

In this module, the network monitor named bro IDS continuously monitors the network to find whether there is any spam mail is being sent on the network. Unwanted spam mails are detected based on the content present in the subject, message body. When the sender's blocking system block the spam mail, then this monitoring system perform its work. The spam messages can identify according to sender's network (i.e.) if sender send unwanted messages from certain network that can be monitored in sender's mail system. While monitoring the network, when the bro IDS detects any spam mail crossing, immediately it tracks the network id of that sender. And the details of that sender will be sent to the Admin of the server. Later the server decides whether to block or not to block the sender's presence in the network based on the sender's further activities.These monitoring process are not known to user until admin take any action on network id.

The spam messages can identify according to Recipient Email Address (REA).(i.e.)if sender sends unwanted messages to recipient that can be monitored in sender's mail system.The network monitor named bro IDS continuously monitors the network to find whether there is any spam mail is being sent on the network. The spam messages can identify according to sender's mail id (i.e.) if sender send unwanted messages from certain mail id that can be monitored in sender's mail system.

While monitoring the network, when the bro IDS detects any spam mail crossing, immediately it tracks the mail id of the sender. And the details of that sender will be sent to the Admin of the server. Those doubted sender's details will be stored in bloom filter for keen monitoring. Later the server decides whether to block or not to block the sender's account based on the sender's further activities (i.e.)if senders continuously try to send spam mails that will be recorded. It is very use full in decision making process (whether block or not block).These monitoring process are not known to user until admin take any action on spammer mail id.

E. Blocking of Spamming Bots

The blocking of the spam mail sender is implemented in this module. The details that have been stored in the Admin of the server are used for blocking the spam sender. The admin maintains two major lists that are the IP addresses from which the spam mail is tried to send and the mail id details which are cased for same spamming activity. These lists are compared with a list which holds the details about the number of times the sender has tried to send spam mail. If the count seems smaller, then the blocking of the system in the network or the sender's account is not considered but when the count seems larger, then the mail id of the sender or the entire system's access over the network will be blocked. Administrator can black the mail id alone or block the network id alone or black the both id.Thereby the receiver is safe-guarded from receiving spam mail. Since the source system which sent spam mail had been blocked completely, and then there will be less chance of sending spam mail.

A 'bot', short for robot, is a type of software application or script that performs tasks on command like indexing a search engine, and they are really good at performing repetitive tasks. Bad bots perform malicious tasks allowing an attacker to take complete control over an affected computer for the criminal to control remotely. Once infected, these machines may also be referred to as 'zombies'. Taking over one computer is useful, but the real value to a criminal comes from collecting huge numbers of computers and networking these (a botnet) so they can all be controlled at once and perform large scale malicious acts. It only takes minutes for an unprotected, internet connected computer to be infected with malicious software and turned into a bot, underscoring the critical need for every computer and Smartphone user to have up-to-date security software on all their devices.

F. System Flow

Architecture diagram shown in the below figure 1is used to study the working of the entire system. It is a technique to map out the structure of the system to be modeled.

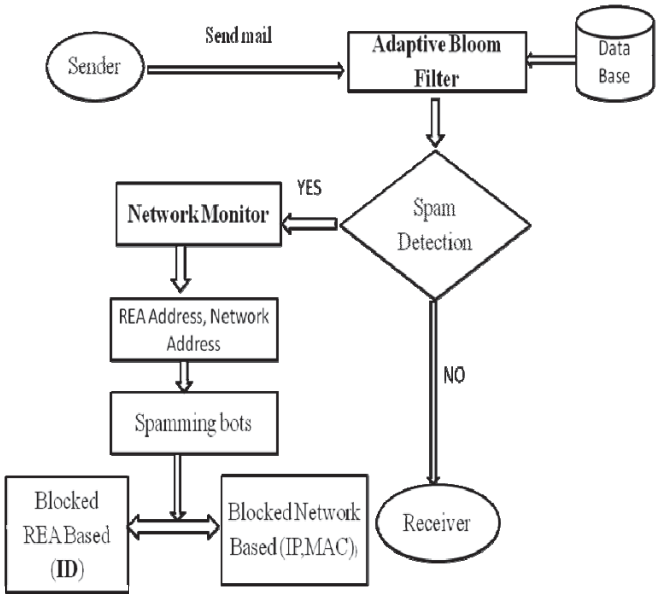

Fig 1, Architecture of Spam Mail Detection and Blocking using Adaptive Bloom filter

When a user who holds email id of a particular server can compose and send mail to any other users who use any server id. The mail without spam will reach the receiver without any interruption. When the sender tries to send a mail which holds spam, there arises interruption i.e., the mail will be blocked in sender area itself. The spam is found by the spam detector which works on every sender block. When the user sends a mail without spam will be received by the receiver. If the sender keep on trying to send the spam mail, the Bro IDS (Intrusion Detection System) running on networking monitoring system will detect that sender and register the information in the Admin about the sender mail id and the number of times the sender had tried to send the spam mail. So thereby that particular sender's mail account will be blocked.If the same sender tries to send spam mail from another account, the verification is made by authorizing the IP address of the sender system. Then the IP address of the sender is

identified by Bro IDS. When many spam mails are tried to send from that IP address will be blocked by the Admin of the Server.

IV. CONCLUSION

Since the spam content mail is considered as the most dreadful thing to damage the system, so it becomes very essential to protect the system from any no expectable damages. This system does the favorable one. Even though the sender's spam blocking system has failed to block the spam mail, the network monitor bro IDS is always active to detect the spam mail and the details about its sender and finally complaining about the sender to the Admin of that particular account server. Later, the admin decides whether to block the mail account or the entire access of the sender over the network which means blocking the IP address. These decisions are taken based on the observation of sender's further activity. Instead of tracking the network for finding the sender's IP address and email id and detecting who had tried to send the spam mail using BRO IDS, it seems more convenient and accurate when we go for finding the probability that the email is spam or not spam. Among many classifiers, Bayesian Spam filter is considered to be more efficient for finding the probability of spam mail tried to send over the network. It is because the probability that is found by Bayesian filter is more accurate.

REFERENCES

[1] Po-Ching Lin,Ping-Hai lin, PinRenChiou, "Detecting Spamming activities by network monitoring with Bloom Filters", IEEE.Trans January 2013.

[2] Y. Xie, F. Yu, K. Achan, R. Panigrahy, G. Hulten and I. Osipkov,"Spamming botnets: signatures and characteristics," In Proceedings of ACM SIGCOMM, Aug. 2008.

[3] C. Y. Cho, J. Caballero, C. Grier, Y. Paxson and D. Song, "Insights from the inside: a view of botnet management from infiltration," In Proceedings of USENIX Workshop on Large-scale Exploits and Emergent Threats (LEET), Apr. 2010.

[4] W. K. Ehrlich, A. Karasaridis, D. Liu and D. Hoeflin, "Detection of spam hosts and spam bots using network flow traffic modeling," In Proceedings of the 3rd USENIX Conference on Large-scale Exploits and Emergent Threats: botnets, spyware, worms, and more (LEET), Apr. 2010.

[5] G. Stringhini, T. Holz, B. Stone-Gross, C. Kruegel, G. Vigna, Botmagnifier: locating spam bots on the internet, in: Proceedings of USENIX Security Symposium, 2011.

[6] Z. Duan, P. Chen, F. Sanchez, Y. Dong, M. Stephenson, J. M. Barker,Detecting spam zombies by monitoring outgoing messages, IEEE Trans.Dependable and Secure Computing 9 (2) (2012) 198-210.

[7] F. Sanchez, Z. Duan and Y. Dong, "Blocking spam by separating end user machines from legitimate mail server machines," In Proceedings of the 8th Annual Collaboration, Electronic messaging, Anti-abuse and Spam Conference (CEAS), Sept. 2011.

[8] S. D. Paola and D. Lombardo,"Protecting against DNS reflection attacks with Bloom filters," In Proceedings of Conference on Detection of Intrusions and Malware & Vulnerability Assessment (DIMVA), July 2011.