A Methodology for Analyzing Overall Flow of Spam-Based Attacks

Jungsuk Song[1], Daisuke Inoue[1], Masashi Eto[1], Mio Suzuki[1],
Satoshi Hayashi[2], and Koji Nakao[1]

[1] National Institute of Information and Communications Technology
{song,dai,eto,mio,ko-nakao}@nict.go.jp
[2] Symantec Japan Research Institute
satoshi_hayashi@symantec.com

Abstract. Over the last decade, unsolicited bulk e-mails, i.e., spams, have been dramatically increasing and they have been definitely recognized as a serious Internet threat. Especially, recent spams mostly caused by various malwares (e.g., bots, worms) often contain URLs that navigate spam receivers to malicious Web servers for the purpose of malware infection. In addition, malwares such as bots operate in cooperation with each other, and there are close links between malwares and malicious Web servers. In this paper, considering the need for further studies on the mitigation of recent spam-based attacks, we propose a methodology for analyzing their overall flow in order to investigate the active relationship among spams, malwares and malicious Web servers. Furthermore, we have evaluated our method using double bounce e-mails obtained from our own SMTP server. The experimental results show that the proposed method is highly effective to analyze the correlation between spams' sources and their eventual destinations.

1 Introduction

Over the last decade, unsolicited bulk e-mails, i.e., spams, have been dramatically increasing and they have been definitely recognized as a serious Internet threat. Many researches report that more than 90% of all Internet e-mails today are considered as spam[1], and they are abused for various purposes. Most spams assume the form of advertising or promotional materials, for example, those related to money, debt reduction plans, getting-rich-quick schemes, gambling opportunities, porn, health and so on[2]. Spams lead to many social problems in terms of productivity and IT infrastructure, and as reported in [3,4], the worldwide cost of spams was estimated to be well over US$10 billion in 2005.

On the other hand, recent spams mostly caused by various malwares (e.g., bots, worms) often contain URLs which navigate spam receivers to malicious Web servers for the purpose of malware infection. In addition, many malwares such as bots operate in cooperation with each other, and there are close links between malwares and malicious Web servers. For example, a botnet consists of many distributed bots connected worldwidely and it sends tremendous amount

C.S. Leung, M. Lee, and J.H. Chan (Eds.): ICONIP 2009, Part II, LNCS 5864, pp. 556–564, 2009.

of spams to target e-mail users in large quantities by means of a command from the C&C server remotely controlled by their herder. Further, botnet herders possess their own malicious Web servers, from which they are able to propagate their well-crafted malwares to victims. There are a wide range of ongoing spam sending systems (e.g., unauthorized relay, dedicated MTA, Web mail, and bots) and diverse purposes of spam (e.g., malware infection, phishing, and e-mail address harvesting). Therefore, in order to cope with these complicated and unclear situations, it is strongly expected to investigate spams' sources and their destinations, and the correlation between them all together in an integrated manner.

There are many anti-spam strategies; however, their main goals are limited only to spam detection, or they take into account spams' sources and URLs' destinations separately[1,5,6]. In this paper, we propose a practical methodology for analyzing the overall flow of spam-based attacks in order to investigate the active relationship among spams, malwares and malicious Web servers. To this end, we construct two analyzers, i.e., source analyzer and destination analyzer, which analyze the characteristics of spams' sources and their eventual destinations, respectively. For the correlation analysis between them, we generate spams' profiles which contain the analysis results obtained by two analyzers. Furthermore, we have evaluated our method using double bounce e-mails obtained from our own SMTP server. The experimental results show that spams' sources and malicious Web servers are closely connected to each other, and the proposed method is highly effective to analyze the correlation between spams' sources and their eventual destinations.

The rest of the paper is organized as follows. In section 2, we give a brief description of the existing researches related to our research. In section 3, we present our approach, and the experimental results including their analysis are given in section 4. Finally, we present our concluding remarks.

2 Related Work

In [1], Anderson et al. have focused on the scam infrastructure that is nourished by spam and have described an opportunistic measurement technique called spamscatter that mines e-mails in real-time, follows the embedded link structure, and automatically clusters the destination websites using image shingling to capture graphical similarity between rendered sites.

In [5], Kreibich et al. have presented an inside look at how campaign orchestration takes place by analyzing the raw material used to produce spam, including textual templates employed for generating highly diverse spam instances. Their analysis shows that today's spamming business operates at a frightening scale without requiring sophisticated mechanisms to counter the anti-spam industry's efforts.

In [6], Kawakoya et al. have analyzed the characteristics of Web-based passive attacks which are driven by the URLs included in spams. They have developed a client-type honeypot to do this. During 15 days observation for a bulk of spam

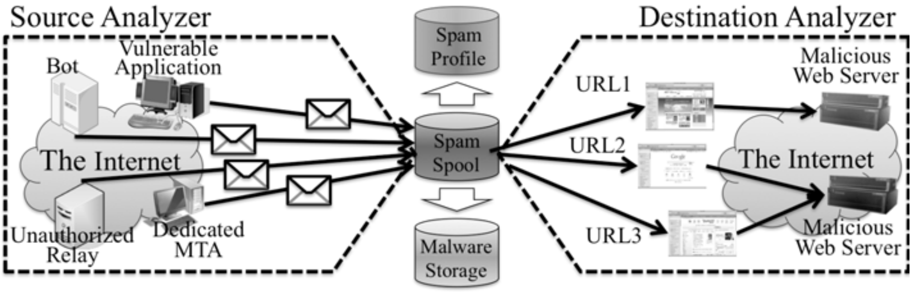

Fig. 1. Overall structure of the proposed method

sent to a mail server of Kyoto University, they found 409 malicious Web sites and 31,618 spam mail messages that had links to malicious websites.

Although these approaches are most closely related to ours, there is a fundamental limitation of their approach is that they have focused on the analysis of either spams' sources or their eventual destinations, so that it is quite difficult to analyze not only the overall flow of spams, but also the correlation between them.

3 Proposed Method

Figure 1 shows the overall structure of the proposed method, which consists of two analyzers (i.e., source analyzer and destination analyzer) and three repositories (i.e., spam spool, spam profile, and malware storage). In our method, we first collect spams from the Internet, and they are stored at a spam spool repository. In order to collect spams, we used double bounce e-mails described in Section 3.1. The source analyzer aims to identify the types of spam sending system and spam herder, and its analysis results are stored at a spam profile repository. The destination analyzer investigates which e-mail has URLs linked to malicious Web pages, and downloads their HTML contents and malwares if possible. Its analysis results are stored at two repositories, i.e., spam profile and malware storage. See Sections 3.2 and 3.3 for more detail.

3.1 Double Bounce E-Mail

During the previous years, many machine learning methods have been proposed for spam detection. However, these methods cannot accurately distinguish spams from normal e-mails. Further, the effective lifetime of the existing techniques is highly short, because spammers frequently change their modus operandi to compromise spam detection techniques. In other words, in order to maintain the effectiveness of spam filters, constant upgrades and new developments are essentially required. Considering the above, we do not capture our analysis data, i.e., spams, using the previous spam detection methods.

Double bounce e-mails indicate that they have no valid recipient address and return-path address. In the case of a normal e-mail, it contains one return-path address at least in its header field, even if a sender mistyped the recipient address to his/her e-mail. In this context, double bounce e-mails can be regarded as pure spam. This double bounce e-mail is quite similar to the concept of darknet which is an area of a routed, and allocated IP space where no active services or servers reside, and thus the incoming packets to it can be treated as abnormal ones. Darknet has been used for analyzing incidents in many researches[7], because many abnormal activities can be easily observed in the darknet. Similarly, we collect double bounce e-mails from the Internet and use them as our analysis data.

3.2 Source Analyzer

The source analyzer identifies the types of spam sending system and its herder. In our method, we classify the types of spam sending system into six categories: unauthorized relay, bot, Web mail, dedicated MTA, vulnerable application, and unidentified source. The identification mechanism is as follows.

- unauthorized relay: if an e-mail arrived at an internal SMTP server from more than two organizations, i.e., two different domain names, it is classified as unauthorized relay. Since in the case of large organizations, e.g., ISP, which often relay e-mails on their internal networks, we do not classify them in this category.
- vulnerable application: if the Received field of the e-mail header contains “localhost” or “127.0.0.1”, and the X-Mailer field represents an e-mail application such as IPB PHP Mailer, operating on the free bulletin board, then the corresponding e-mail falls in this category.
- bot: if an e-mail satisfies the following two conditions, then it is classified in this category. (1) There is no regular SMTP server in the Received field of the e-mail header, and (2) the Return-path field of the e-mail header is empty. If not, the hostname of the Return-path field is unmatched to the IP address which is obtained by reverse lookup of the MX record. Because bots try to directly connect to a target SMTP server, and it tends to inform a target SMTP server that their hostname is unmatched to the IP address.
- dedicated MTA: Dedicated MTA makes a regular form of the e-mail headers. If there are no suspicious fields in the e-mail header, we classify it as dedicated MTA.
- Web mail: if the Received field contains an IP address of Web mail servers, e.g., Hotmail, Yahoo, etc., then the corresponding e-mail is classified as a Web mail.
- unidentified source: if an e-mail does not belong to the above five types, and the Received field contains a suspicious value, e.g., a host connected to several different SMTP servers, then we classify the e-mail as an unidentified source.

The other purpose of the source analyzer is to classify the given spams into several groups based on similarity of the e-mail header. If the headers of two e-mails have the same pattern, i.e., the order of the header fields, they become the members of the same group. Considering the situation where RFC 2821[8] does not contain the comment about the order of the header fields, if the patterns of two e-mail headers are the same, the patterns can be regarded as being sent by the same SMTP engine. This also means that they originate from the same herder because in many cases SMTP engines basically have different operations.

For each e-mail, an entry including the IP addresses of its sender and receiver, an IP address of a relay SMTP server, a type of spam sending system and a type of spam herder, is inserted to the spam profile repository.

3.3 Destination Analyzer

As shown in Figure 2, the destination analyzer consists of four main parts : URL extractor, database, crawler, and evaluator. The Destination Analyzer first extracts the URLs from each e-mail, and then removes duplicate URLs using the URL extractor. We input each unique URL into the database, and then the crawler reads each of them one by one(①). The crawler attempts to connect to a website directly linked to the corresponding URL, and downloads available data(e.g., HTML contents, malwares) and new URLs, until there are no more URLs to be analyzed(②). The crawling results of ② are inserted into the database(③). The evaluator reads the data to be evaluated from the database, and evaluates whether the data are malicious or not(④). For this investigation, we use a dedicated software, i.e., SPIKE[9] which has three analysis modules:

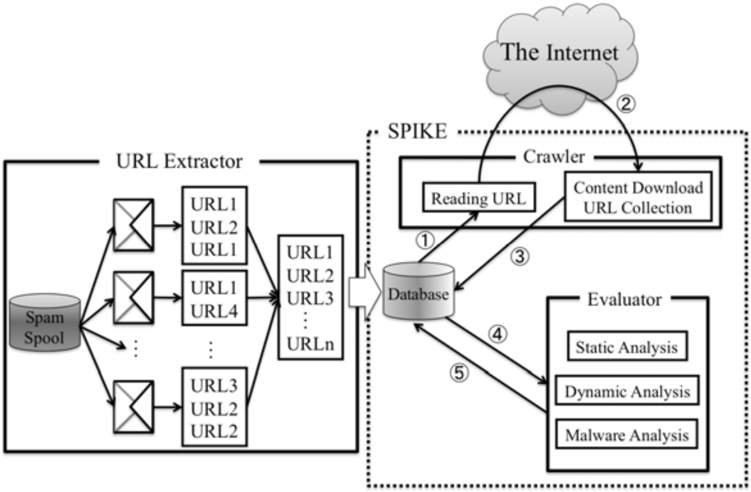

Fig. 2. Process of destination analyzer

static analysis module, dynamic analysis module and malware analysis module. The evaluation results of ④ are stored in the database(⑤).

We add these analysis results, e.g., whether an e-mail contains an URL linked to a malicious Web site or not, and IP addresses related to the URLs to each entry created in the source analyzer. Further, downloaded malwares are stored in the malware storage repository.

4 Experimental Results

4.1 Data Captured for Our Analysis

The experimental data we used were double bounce e-mails that arrived at our own SMTP server. We gathered about four days(April 16, 2009, to April 20, 2009) of double bounce e-mails, and captured 108,839 e-mails. From the original 108,839 e-mails, we filtered out some irregular e-mails from our evaluation data, and consequently we obtained 98,890 regular e-mails. We extracted 159,215 unique URLs from these regular e-mails.

4.2 Results and Analysis

We first evaluated 159,215 unique URLs using the destination analyzer from July 16, 2009, to July 29, 2009, and observed that 16 URLs were directly linked to a malicious Web page; however, there were no executable malwares. Further, among the regular 98,890 e-mails, 137 e-mails contained one of those 16 URLs. We divided 137 e-mails into 16 groups according to each URL. We denoted each group as $\{g_0, g_1, \cdots, g_{15}\}$. With respect to these 16 groups, we analyzed them by using the source analyzer. Table 1 shows the analysis results. From Table 1, we can observe that the 16 groups were sent by 6 different spam herders(i.e., a, b, c, d, e, and f), and 3 different spam sending systems(i.e., bot, vulnerable application, and unauthorized relay). We can predict that 6 different types of bots belonged to a single botnet, and they were under a controller (i.e., type

Table 1. Classification result by source analyzer

		Type of spam sending system					
		bot	vulnerable application	unauthorized relay	dedicated MTA	Web mail	unidentified source
Type of spam herder	a	g_0, g_1, g_2, g_3, g_6, g_{12}	-	-	-	-	-
	b	-	g_5, g_{10}, g_{13}	-	-	-	-
	c	-	g_8, g_9, g_{15}	-	-	-	-
	d	-	g_7	-	-	-	-
	e	-	-	g_{11}	-	-	-
	f	-	-	-	-	-	g_4, g_{14}

Table 2. Correlation between spams' sources and URLs' destinations

Group	Sender IP	URL IP	Group	Sender IP	URL IP
g_0	62.h1.h2.206(RU)	216.x1.x2.102(US)	g_5	78.n1.n2.165(TR)	200.v1.v2.78(BR)
g_3	216.i1.i2.102(US)	216.x1.x2.102(US)	g_7	89.o1.o2.97(RO)	200.v1.v2.78(BR)
g_1	189.j1.j2.18(BR)	68.y1.y2.143(US)	g_8	58.p1.p2.204(PK)	200.v1.v2.78(BR)
g_2	78.k1.k2.8(RU)	68.y1.y2.143(US)	g_9	95.q1.q2.158(RO)	200.v1.v2.78(BR)
g_6	201.l1.l2.208(BR)	68.y1.y2.143(US)	g_{10}	200.r1.r2.150(BR)	200.v1.v2.78(BR)
g_{12}	80.m1.m2.94(IR)	68.y1.y2.143(US)	g_{13}	78.s1.s2.76(LT)	200.v1.v2.78(BR)
-	-	-	g_{15}	89.t1.t2.41(RO)	200.v1.v2.78(BR)

"a" of spam herder); further, 3 different types of e-mail sending applications operating on a free bulletin board were abused to send spams.

In order to verify this, we investigated the correlation between spams' sources and URLs' destinations as shown in Table 2. It should be noted that the parenthesis indicates the country code of IP address. From Table 2, we can see that 6 different types of bots are located in 4 different countries (i.e., RU, US, BR and IR), and the spams sent by them induced e-mail receivers to two malicious Web servers in the US. Furthermore, in our investigation, we found that 6 URLs shared a common string, i.e., "http://xxxx.com/resp/xxxx," in which only "xxxx" changes to a different string in each URL.

On the other hand, in the case of 7 groups (i.e., g_5, g_7, g_8, g_9, g_{10}, g_{13}, and g_{15}), we can observe that 7 machines located in 5 different countries (i.e., TR, RO, PK, BR, and LT) are compromised by the vulnerability of e-mail sending application. In fact, we found that 3 different applications, IPB PHP Mailer, vBulletin Mail via PHP and MyBB, were used for sending spams. Further, it is easy to see that 7 URLs included in spams are directly linked to a single malicious Web server located in Brazil. In addition, we observed that they have a quite similar URL pattern, i.e., "http://yyyy.fantasticzoneonline.at/," in which only "yyyy" changes to a different string in each URL. As a result, this means that these spams were sent by only a single herder, and 7 machines may be under his/her control.

In the case of two groups(i.e., g_4 and g_{14}), we classified them into the type of unidentified source because we observed that they tried to connect to two different SMTP servers(including our own SMTP server) from a single host, which is unnatural in a normal case. Finally, it is easy to see that e-mails sent by dedicated MTA were not observed in our experimental data. In general, dedicated MTA sends a large amount of spam to many SMTP servers in a short time, when administrators of dedicated MTA receive a request from their client. As such, their activities can be easily identified, and consequently their domain name is blacklisted. Since a dedicated MTA changes the corresponding domain name to another one so as to evade being blacklisted, the lifetime of a dedicated MTA is considered to be very short. This means that we should analyze spam as soon as possible, so that we can identify which e-mails are sent by the dedicated MTA.

Table 3. Evaluation result according to links-depth by destination analyzer

	Links-depth		
	2	3	4
# of input URLs	4,629	4,629	4,629
# of accessible URLs	4,701	5,843	325,596
# of malicious URLs	9(0)	6(0)	1,181(12)

The above results were obtained under the condition where the depth of links to follow Web pages is only 1. In other words, for each given URL, SPIKE connects to only the corresponding Web page, and never jumps to another Web page linked on that Web page. However, in many cases, Web-based attacks tend to be carried out through several steps, i.e., several Web pages. Because attackers have to lure victims, who are browsing on a famous website such as Google or Yahoo, to their malicious website, in which well-crafted malwares are embedded. Thus, we reevaluated 4,629 URLs that were accessible to connect, and the experimental results are shown in Table 3. From Table 3, we can understand that both the number of accessible URLs and the number of malicious URLs are rapidly increasing when the links-depth is 4. It is to be noted that the parenthesis indicates the number of URLs where executable malware was downloadable. These results signify that it is highly important to trace the overall flow of the Web-based attacks, and our approach provides the basic and expansive methodology which enables one to do that.

5 Conclusion and Future Work

In this paper, we have proposed a methodology for analyzing the overall flow of spam-based attacks in order to investigate useful correlations among spams, malwares and malicious Web servers. To this end, we have constructed spams' profiles, which enable us to identify not only the overall flow of spams but also the characteristics of spam sending system and spam herder and to identify which e-mail is connected to a malicious Web page; further, we have successfully provided the correlation between spams' sources and their eventual destinations.

We have evaluated the proposed methodology using double bounce e-mails obtained from our own SMTP server and have showed its effectiveness to analyze the correlation between spams' sources and their eventual destinations. Our key findings can be summarized as follows: (1) 16 URLs among 159,215 unique URLs extracted from 98,890 regular e-mails were directly linked to a malicious Web page, (2) 6 different types of bots belonging to a single botnet were located in 4 different countries, and spams sent by them lured e-mail recipients to only two malicious Web servers in the US whose URLs were quite similar. (3) 7 machines located in 5 different countries sent spams by abusing the vulnerability of an e-mail sending application operating on the free bulletin board, and spams' URLs were directly linked to a single malicious Web server located in Brazil, and (4) the number of accessible URLs and the number of malicious URLs were rapidly increasing when the links-depth was 4.

For the future work, we need to analyze further the correlation between URLs, i.e., Web pages and their sources with respect to the links-depth larger than 1. In addition, we will evaluate our methodology under a real-time operational environment. In order to analyze spam-based attacks more effectively, we have a plan to visualize their overall flow utilizing the IP address information of the spam profile repository, and perform correlation analysis between malwares stored at the malware storage repository and other type of malwares, e.g., malwares gathered by nicter(Network Incident analysis Center for Tactical Emergency Response)[7].

References

1. Anderson, D.S., Fleizach, C., Savage, S., Voelker, G.M.: Spamscatter: characterizing Internet scam hosting infrastructure. In: Proceedings of the USENIX Security Symposium, Boston (2007)
2. Li, F., Hsieh, M.: An empirical study of clustering behavior of spammers and group based anti-spam strategies. In: Conference on Email and Anti-Spam 2006 (CEAS 2006), pp. 21–28 (2006)
3. Jennings, R.: The global economic impact of spam, 2005 report. Technical report, Ferris Research (2005)
4. Spira, J.: Spam e-mail and its impact on it spending and productivity. Technical report, Basex Inc. (2003)
5. Kreibich, C., Kanich, C., Levchenko, K., Enright, B., Voelker, G., Paxson, V., Savage, S.: Spamcraft: an inside look at spam campaign orchestration. In: Proceedings of the 2nd USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET 2009), Boston (2009)
6. Kawakoya, Y., Akiyama, M., Aoki, K., Itoh, M., Takakura, H.: Investigation of spam mail driven Web-based passive attack. IEICE Technical Report, ICSS2009-5, May 2008, 21–26 (2009)
7. Nakao, K., Inoue, D., Eto, M., Yoshioka, K.: Practical correlation analysis between scan and malware profiles against zero-day attacks based on darknet monitoring. IEICE Transactions on Information and Systems E92D(5), 787–798 (2009)
8. RFC 2821, `http://www.ietf.org/rfc/rfc2821.txt`
9. Hosihzawa, Y., Kawamorita, K., Tachikawa, T., Kamizono, M.: A Proposal for autonomous crawling client honeypot. IEICE Technical Report, IA2009-3, ICSS2009-11, June 2009, 13–17 (2009)