DNS サーバの syslog 解析による大量メール送信型ワーム感染端末 IP アドレスの特定

松葉 龍一 †、武藏 泰雄 †、杉谷 賢一†

概要: 熊本大学のトップドメイン-セカンダリ DNS サーバの syslog について統計解析を行った。我々の得た興味深い結果は以下の通りである: (1) 大量メール送信型ワームに感染した PC 端末はワーム活動中に A および MX レコードパケットを DNS サーバへ送信する。(2) 乗っとられた UNIX 系の PC 端末等は spam リレー活動中に A、MX および PTR レコードを DNS サーバへ送信する。以上の結果、DNS サーバのログを監視するだけで大量メール送信型ワームに感染した PC 端末を検知可能であることが示された。

Detection of Mass Mailing Worm-infected IP address by Analysis of Syslog for DNS server

Ryuichi Matsuba,† Yasuo Musashi,† and Kenichi Sugitani†

Abstract: The syslog messages of the topdomain-secondary DNS server in Kumamoto University were statistically investigated when infection of mass mailing worm (MMW) like W32/Sobig, W32/Mydoom, and W32/Netsky were increased worldwidely. The interesting results are: (1) The MMW-infected PC terminal sends packets including only both A and MX records to the DNS server when going on MMW-infection. (2) The hijacked/UNIX-like PC terminal transmits packets including A, MX, and PTR records to the DNS server in a spam relay. Therefore, we can detect MMW-infected PC terminals by only monitoring the DNS query traffic from the DNS clients like PC terminals.

1. Introduction

Intrusion detection system (IDS) is one of attractive solutions to keep security of the computer network servers.[1−15] There are two types of IDSs; one is a misuse intrusion detection (MID) type,[4,6] scanning a database of the remote attacking pattern, and the other is an anomaly intrusion detection (AID) type.[4−12] Recent IDS includes both models like the former and the latter. Surely, the IDS provides a plenty of useful messages but it generates too much messages to analyze in a real time.

In order to develop a new useful statistical MID/AID-hybrid IDS against future remote attack on the network servers, it is of considerable importance to get detailed profile/information for traffic of network applications like DNS query packets between a DNS server and a DNS client. We have shown that DNS query packets are predominantly generated from an SMTP engine of an E-mail server, and that the access of the DNS query packets is mainly driven by SMTP accesses.[16−20] Moreover, we have found a relation between the number of the DNS query packets $D_{\rm q}$ and those of the SMTP $N_{\rm SMTP}$ and POP3 $N_{\rm POP3}$ accesses; $D_{\rm q} = m_{\rm SMTP} N_{\rm SMTP} + N_{\rm POP3}$ ($m_{\rm SMTP} \geq 2$).[16]

The present paper is in a series of correlation analyses on DNS query traffic between DNS server and DNS clients that especially include SMTP engines.[16] Particularly, we focus on the case where PC terminals are infected with a mass mailing worm (MMW) and are hijacked UNIX-like PC terminals with a spam relay (SR-embedded). By analyzing syslog messages of the DNS server, we show how to detect IP addresses of the MMW-infected PC terminals and the SR-embedded PC terminals.

†熊本大学総合情報基盤センター・Center for Multimedia and Information Technologies, Kumamoto University.

2. Observations

2.1 Network systems

We investigated traffic of DNS query accesses between the top domain DNS server (**tDNS**) † and DNS clients of A (**cA**), B (**cB**), C (**cC**) and D (**cD**), where **cA**, **cB** and **cC** are W32/Sobig.F, W32/Mydoom.A, and W32/Netsky.C-infected PC terminals, respectively, and **cD** is a Compaq Alpha True64 PC terminal in the laboratory of our university. Figure 1 shows a schematic diagram of a network observed in the present study. **tDNS** is one of the top level secondary domain name (kumamoto-u) server and plays an important role of subdomain delegation and domain name resolution services for many PC terminals. **cA**, **cB** and **cC** are DNS clients of **tDNS** in which the first DNS server is configured to access to **tDNS**.

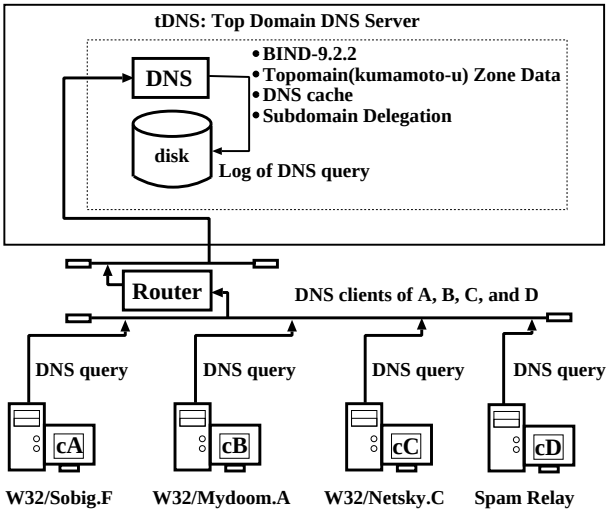

Figure 1. A schematic diagram of a network observed in the present study.

2.2 A Method of Analysis

In **tDNS**, BIND-9.2.3 program package has been employed as DNS server daemon.[23] The DNS query packets and their contents have been recorded by the query logging option (see man named.conf), as follows:

```
logging {
        channel qlog {
                syslog local1;
        };
        category queries { qlog; };
}
```

The log of DNS query access has been recorded in the syslog file.[24] All of the syslog files are daily updated by the crond system. It is known that a DNS server provides mainly a host domain name (A record), an IP address (PTR record), and mail exchange (MX record) to DNS clients.

We extract lines described DNS query accesses only including MX records from the syslog file in **tDNS**. After discarding IP addresses of DNS query accesses from the outside of university and the E-mail servers that are authorized in our university, we sorts the lines to get top IP addresses of DNS query accesses by using "sort -r" and "uniq -c" commands, as two and one times, respectively, and to show a frequency of the DNS access. If the frequency takes over 50 times, we investigate DNS query contents to get how many MX, A, and PTR records. These procedures are properly worked out to a C-Shell "mmwip" script.

Also, we have developed an automated MX record packet detection system (MX-RPDS) that consists of C-shell (mscan) and Perl scripts (mydwat.pl). The "mydwat.pl" script hooks up the "mscan" script in order to scan the syslog file of **tDNS** in a time per 10 seconds. The "mscan" script kicks the "mmwip" script and "smail" command that we also prepared for a direct SMTP transmitting program with the gcc-2.95.3 C compiler and that it sends an E-mail without using a local MTA because of network security. The "mscan" script also scans a database file that includes E-mail addresses of registered subnetwork managers so that the DNS clients that include MX record is automatically E-mailed to subnetwork manager. With use of MX-RPDS, we can easily and rapidly detect abnormality of MX record access in **tDNS**.

†**tDNS** is a primary DNS server in Kumamoto University (kumamoto-u). The OS is Linux OS (kernel-2.4.24), and an Intel Xeon 2.40GHz machine.

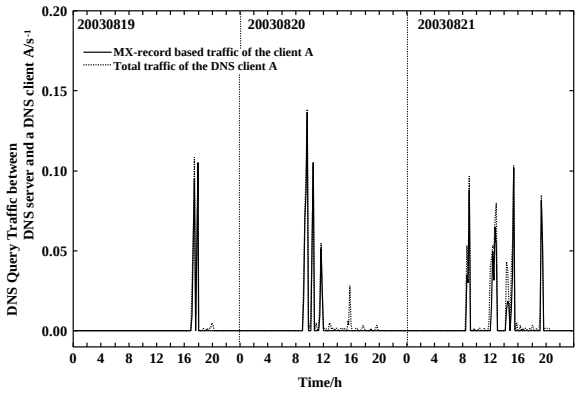

Figure 2. Traffic of the DNS query access between the top domain DNS server and the DNS client A through August 19th to 21st, 2003. The dotted line shows the total DNS traffic and the solid line indicates the MX-record based DNS traffic (s^{-1} unit).

3. Results and Discussion

3.1 W32/Sobig.F MMW

We observed DNS query access traffic from a DNS client A (**cA**) to the top domain name server (**tDNS**) for August 19th-21st, 2003.

We shows the observed DNS query access traffic in Figure 2. The abscissa is times in units of hour and the ordinate is access count rates from **cA** to **tDNS**. Since **cA** is an Windows/XP system as used only PC terminal *i.e.* **cA** is not a server, **cA** generates only very small DNS query traffic and the traffic includes only A record packet in usual (see before 17:00 at August 19th, 2003 the dotted line in Figure 2). The **cA** DNS query traffic changes in a large scale manner after 17:00 at August 19th, 2003, and the traffic is continued to 20:30 at August 19th, 2003. The large change in traffic was taken place with an infection of mass mailing worm (MMW) in **cA**. How do we recognize the change as the infection of MMW ?

Table 1 gives the total number of lines described MX, A, and PTR records on **cA** for the observed days. Interestingly, the total traffic consists of MX and A records. No PTR record can be found in the syslog messages for **cA**. Also, the MX record based traffic curve emerges as the solid line in Figure 2. These features provide important information that **cA** has an SMTP engine. We confirm that the

Table 1. The total number of lines for MX, A, and PTR records per a day in the syslog file in **tDNS**, relating to the DNS client access from **cA**.

day	MX	A	PTR
Aug. 19th	190	36	0
Aug. 20th	335	89	0
Aug. 21th	422	201	0

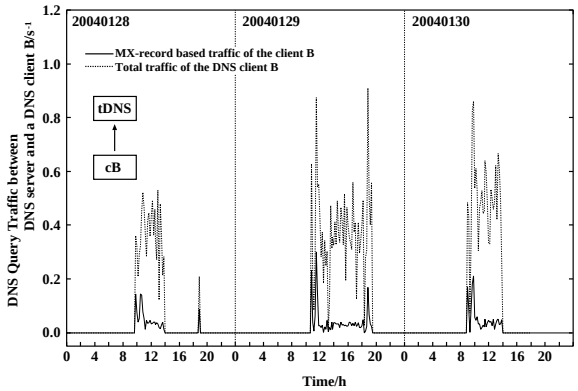

Figure 3. Traffic of the DNS query access between the top domain DNS server and the DNS client B through January 28th to 30th, 2004. The dotted line shows the total DNS traffic and the solid line indicates the MX-record based DNS traffic (s^{-1} unit).

DNS query traffic is dominated by MX records, and that the query drastically increases when **cA** is turned on in the latter two days.

When we see the syslog file for DNS query packets from **cA**, we encounter head lines in which, "A.ROOT-SERVERS.NET", "A.ROOT-SERVERS.NET", "B.ROOT-SERVERS.NET", "B.ROOT-SERVERS.NET",..., etc are written. These head lines are included in the virus database as the W32/Sobig.F mass mailing worm and its infection is detected in public at the August 19th, 2003.[25] Therefore, we can clearly detect that **cA** is surely infected with the W32/Sobig.F. It is noted that the head lines includes two same lines. This is because **cA** is also infected with the W32/Sobig.C.

3.2 W32/Mydoom.A MMW

We illustrate the DNS query traffic between **tDNS** and the DNS client B **cB** in Figure 3 through January 28th-30th, 2004. The DNS traffic includes only MX and A records. No PTR record is written in the syslog messages for **cB**. This feature

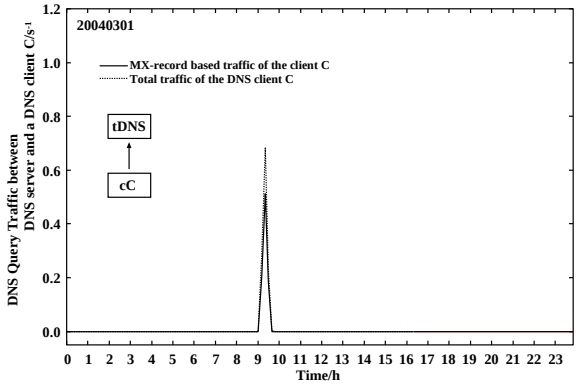

Figure 4. Traffic of the DNS query access between the top domain DNS server and the DNS client C at March 1st, 2004. The dotted line shows the total DNS traffic and the solid line indicates the MX-record based DNS traffic (s^{-1} unit).

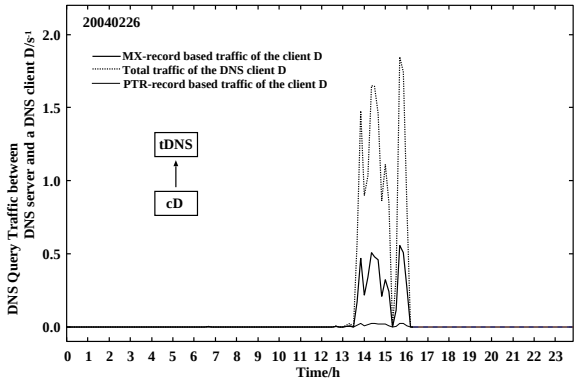

Figure 5. Traffic of the DNS query access between the top domain DNS server and the DNS client D at February 26th, 2004. The dotted line shows the total DNS traffic and the solid line indicates the MX-record based DNS traffic (s^{-1} unit).

is observed in the case of W32/Sobig.F MMW.

When we see the syslog file for DNS query packets from **cA**, we encounter head lines in which, "mx.xxxxx.co.jp", "mail.xxxxx.co.jp", "smtp.xxxxx.co.jp", "mx1.xxxxx.co.jp", "mxs.xxxxx.co.jp", "mail1.xxxxx.co.jp", "relay.xxxxx.co.jp", "ns.xxxxx.co.jp", "gate.xxxxx.co.jp",..., etc are written. These head lines are included in the virus database as the W32/Mydoom.A mass mailing worm and its infection is detected in public at January 28th, 2004.[26] Therefore, we can clearly detect that **cB** is surely infected with the W32/Mydoom.A MMW.

Interestingly, the traffic of MX record packet is totally less than that of total traffic *i.e.* that of A record packet. This results differs from the case of W32/Sobig.F (see Figure 2). It is fact that the total DNS query packets (5630) consist of 807 MX and 4823 A record packets at January 28th, 2004. This feature is interpreted in terms that the W32/Mydoom.A initially searches fully qualified domain name (FQDN) of the next victim PC terminals with the complement of host name keywords as, "mx.", "mail.", "smtp.", "mx1.", "mxs.", "mail1", "relay.", "ns.", and "gate.".[26] Therefore, this scan generates a lot of A record packets more than MX record ones.

3.3 W32/Netsky.C MMW

Figure 4 shows the DNS traffic from the DNS client C **cC**. The DNS traffic includes many MX record packets, a small number of A record packets, and exceptionally only one PTR record packet. The former features resemble well those of the W32/Sobig.F MMW. The last feature is exceptional case because it checks a loop back address (127.0.0.1). We encounter a head MX record line that includes "yahoo.com" when seeing the syslog file for DNS query packets from **cC**. This shows that **cC** is surely infected with the W32/Netsky.C MMW.[27]

3.4 Spam Relay

We illustrate the DNS traffic between **tDNS** and DNS client D **cD** in Figure 5. The DNS query traffic mainly includes MX and A record packets and slightly includes PTR record packets. This traffic including PTR records clearly differs from the traffic from MMW-infected PC terminals that no PTR record is included. This difference is interpreted in terms that **cD** is a True64 UNIX Compaq Alpha PC terminal.

Since **cD** is left to be a default setting, **cD** becomes easily to be a spam relay by hijacking. In **cD**, a default SMTP server program (MTA; sendmail old version) is running. The SMTP server

program like sendmail[21] or postfix[22] initially generates two DNS packets that consist of PTR and A records to check the SMTP whether or not clients is authorized.[16] Furthermore, the SMTP server program request at least a couple of packets that consist of MX and A records to get FQDN of domain name E-mail address and to get an IP address of the FQDN that manages an E-mailing destination address.[16] It is fact that the total DNS query packets (9816) consist of 2922 MX, 6755 A, and 139 PTR record packets at February 26th, 2004. Therefore, we can detect a spam relay embedded UNIX-like PC terminal by presence of PTR records in its DNS query packets.

4. Concluding Remarks

We statistically investigated system log (syslog) files in the top domain DNS server (**tDNS**) when several PC terminals were infected by mass mailing worm (MMW). By monitoring the DNS query accesses on **tDNS**, we have found information about detection of an IP address of a MMW-infected PC terminal: (1) Usually, the DNS query traffic of the DNS clients like Windows PC terminals includes an A (Address) record only, but it contains MX (Mail Exchange) and A records without PTR (Pointer/Reverse) record when the DNS clients are infected with the MMW like W32/Sobig.F(with W32/Sobig.C) and W32/Mydoom.A. Exceptionally, the W32/Netsky.C MMW-infected DNS clients send DNS query packets including only one PTR record. (2) When the DNS clients like UNIX/UNIX-like PC terminals is unauthorized as a network sever in our university, the DNS query traffic is usually small. However, the DNS query traffic increases to a greater extent than that in the usual when the DNS client becomes a spam relay and/or an hijacking target and it includes MX, A and PTR records.

From these results, it can be reasonably concluded that we can detect the MMW-infected PC terminals in a high probability because we can discriminate between MMW-infection and spam relay with checking existence of PTR record in the DNS query traffic.

We continue further investigation in order to get more information that improves the automated system detecting the MMW-infected PC terminals with a spam relay.

Acknowledgement. All the calculations and investigations were carried out in Center for Multimedia and Information Technologies, Kumamoto University. We specially thank to technical officers, K. Tsuji, M. Shimamoto and T. Kida, and K. Makino who is a system engineer of MQS (Kumamoto) for daily supports and constructive cooperations.

References and Notes

1) Northcutt, S. and Novak, J., *Network Intrusion Detection*, 2nd ed; New Riders Publishing: Indianapolis (2001).

2) Sato, I., Okazaki, Y., and Goto, S.: An Improved Intrusion Detecting Method Based on Process Profiling, *IPSJ Journal*, Vol. 43, No.11, pp.3316-3326 (2002).

3) Jones, D.: Building an E-mail Virus Detection System for Your Network, *LINUX Journal*, No.92, pp.56-65 (2001).

4) Denning, D. E.: An Intrusion-detection model, *IEEE Trans. Soft. Eng.*, Vol. SE-13, No.2, pp.222-232 (1987).

5) Cisco Systems: The Science of Intrusion Detection System Attack Identification, http://www.cisco.com/warp/public/cc/pd/-sqsw/sqidsz/prodlit/idssa_wp.htm, 2002.

6) Laing, B.: How To Guide-Implementing a Network Based Intrusion Detection System, http://www.snort.org/docs/iss-placement.pdf, ISS, 2000.

7) Mukherjee, B., Todd, L., and Heberlein, K. N.: Network Intrusion Detection, *IEEE Network*, Vol. 8, No.3, pp.26-41 (1994).

8) Barbará, D., Wu, S., and Jajodia, S.: Experience with EMERALD to DATE", Proceedings 1st USENIX Workshop on Intrusion Detection and Network Monitoring, Santa Clara, California, April 1999, pp.73-80, http://www.csl.sri.com/neumann/det99.html

9) Neumann, P. and Porras, P.: Detecting Novel Network Intrusions using Bayes Estimators", First SIAM International Conference on Data Mining, 2001, http://www.siam.org/meetings/sdm01/pdf/-sdm01_29.pdf

10) Warrender, C., Forrest, S., and Pearlmutter, B.: Detecting Intrusions Using System Calls: Alternative Data Models, *Proc. IEEE Symposium on Security and Privacy*, No.1, pp.133-145 (1999).

11) Hofmeyr, S. A., Somayaji, A., and Forrest, S.: Intrusion Detection Using Sequences of System Calls, *Computer Security*, Vol. 6, No.1, pp.151-180 (1998).

12) Ptacek, T. H. and Newsham, T. N.: Insertion, Evasion, and Denial os Service: Eluding Network Detection, January, 1998, http://www.robertgraham.com/mirror/Ptacek-Newsham-Evasion-98.html

13) Anderson, D., Lunt, T. F., Javitz, H., Tamaru, A., and Valdes, A.: Detecting unusual program behavior using statistical component of the Next-generation Intrusion Detection Expert System (NIDES), *Computer Science Laboratory SRI-CSL-95-06*,1995.

14) Symantec: ManHunt, http://enterprisesecurity.symantec.com/products/products.cfm?ProductID=156&EID=0

15) Yamamori, K.: An Improvement of Network Security Using an Intrusion Detection Software, *Journal for Academic Computing and Networking*, No.4, pp.3-13 (2000).

16) Musashi, Y., Matsuba, R., and Sugitani, K.: Traffic Analysis on a Domain Name System Server. SMTP Access Generates Many Name-Resolving Packets to a Greater Extent than Does POP3 Access, *Journal for Academic Computing and Networking*, No.6, pp.21-28 (2002).

17) Musashi, Y., Sugitani, K., and Matsuba, R.: Traffic Analysis on Mass Mailing Worm and DNS/SMTP, *IPSJ SIG Notes, Computer Security 19th*, Vol. 2002, No.122, pp.19-24 (2002).

18) Musashi, Y., Matsuba, R., and Sugitani, K.: Statistical Analysis in Logs of DNS Traffic and E-mail Server, *IPSJ SIG Notes, Computer Security 20th*, Vol. 2003, No.18, pp.185-189 (2003).

19) Musashi, Y., Matsuba, R., and Sugitani, K.: Statistical Analysis in Log Files of Electronic-Mail Server and Domain Name System Server. SPAM Mail Generates Many DNS Query Packets Traffic Analysis on a Domain Name System Server, *Journal for Academic Computing and Networking*, No.7, pp.5-11 (2003).

20) Matsuba, R., Musashi, Y., and Sugitani, K.: Statistical Analysis in Syslog Log Files ins DNS and Spam SMTP Relay Servers, *IPSJ Symposium Series*, No.2004, pp.31-36 (2004).

21) http://www.sendmail.org/

22) http://www.postfix.org/

23) http://www.isc.org/products/BIND/

24) Bauer, M.: syslog Configuration, *LINUX Journal*, No.92, pp.32-39 (2001).

25) http://www.trendmicro.co.jp/vinfo/virusencyclo/-default5.asp?VName=WORM_SOBIG.F

26) http://www.trendmicro.co.jp/vinfo/virusencyclo/-default5.asp?VName=WORM_MYDOOM.A&V-Sect=T

27) http://www.trendmicro.co.jp/vinfo/virusencyclo/-default5.asp?VName=WORM_NETSKY.C