社団法人 情報処理学会　研究報告
IPSJ SIG Technical Report

2004－CSEC－27　(8)
2004／12／20

DNS クエリアクセス監視による大量メール送信型ワーム感染端末の検知

武藏 泰雄[†]　　Kai Rannenberg [‡]

概要: DNS サーバと PC 端末との DNS クエリアクセスを監視するだけで大量メール送信型ワーム (MMW) に感染した PC 端末の IP アドレスを検知する新しいシステムの開発を行ったので報告する。

Detection of Mass Mailing Worm-infected PC terminals by Observing DNS Query Access

Yasuo Musashi,[†] and Kai Rannenberg[‡]

Abstract: We have developed a new detection system of IP addresses of the mass mailing worm (MMW)-infected PC terminals by only watching the domain name system (DNS) query access between the DNS server and the PC terminals.

1. Introduction

Recent internet worms (IW) are mainly categorized into two types, as follows: one is a mass-mailing-worm (MMW) like W32/Sobig.F MMW[1], W32/Mydoom.A MMW[2], and W32/Netsky.Q MMW[3] which transfers itself by an attachment file of the E-mail and the other is a system-vulnerability-attack-worm (SVAW) like W32/Welchia SVAW[4] and W32/Sasser.D SVAW[5] that transfers itself by attack on vulnerabilities of remote buffer overflow in the operating systems and/or the application softwares. Especially, since March 29th, 2004, the former MMW like W32/Netsky.Q MMW has attacked a lot of Windows PC terminals worldwidely because of quick development of W32/Netsky MMW variants[3] and the delay of delivering a virus pattern for worm detection in PC terminals. The speed of the MMW development is too much fast to fix it so that we have detected a total of 104,010 worm actions of the W32/Netsky.Q MMW-infected PC terminals in our university from 12:50 to 17:44 at March 29th, 2004. Furthermore, our university has 920 PC terminals for learners that have a large scaled potential for MMW breeding; in fact, we have detected 800 IP addresses of W32/Welchia SVAW at August 20th, 2003, which include a wide IP address range of 920 PC terminals for learners *i.e.* the PC terminals are considerably to be a big threat. From these points, it is of considerable importance to detect an IP address of the MMW-infected PC terminal.

One of the attractive solutions to detect of an IP address of the MMW-infected PC terminal is to employ an intrusion detection system (IDS)[6−14]. We know that the Snort[14] is an open-sourced package and a rule-based network based IDS, which has a lot of functions such as packet capture, IP defragmentation, TCP stream reassembling (stateless/stateful), and content matcher (detection engine). However, it is unable to detect packets that are related with the worm action of the virus like MMW without any signature patterns. Therefore, we designed and developed a new indirect virus detection system (MXRPDS) that detects IP addresses of the mass mailing worm (MMW)-infected PC terminals by only watching the domain name system (DNS) query traffic between the DNS

[†]熊本大学総合情報基盤センター・Center for Multimedia and Information Technologies, Kumamoto University.
[‡]Mobile Commerce & Multilateral Security, Goethe University Frankfurt

server and the PC terminals.

The present paper discusses on (1) illegal DNS query MX record packet accesses from the MMW-infected PC terminals in which PC terminals are infected with a mass mailing worm (MMW), (2) how to prevent infection of the MMW, and (3) evaluation of the newly developed system.

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**, **cC**, and **cD** are W32/Sobig.F, W32/Mydoom.A, W32/Netsky.Q, and W32/Mydoom.S MMW-infected PC terminals, respectively. Figure 1 shows a schematic diagram of a network observed in the present study. **tDNS** is one of the top level secondary domain name server and plays an important role of subdomain delegation and domain name resolution services for many PC terminals. In **tDNS**, the operating system (OS) is employed Linux OS (kernel-2.4.27), and an Intel Xeon 2.40GHz Dual CPU machine. The PC terminals, **cA**, **cB**, **cC**, and **cD** are DNS clients of **tDNS** in which the first DNS server is configured to access to **tDNS**.

In **tDNS**, BIND-9.2.3 program package has been employed as DNS server daemon[15]. 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[16]. All of the syslog files are daily updated by the crond system.

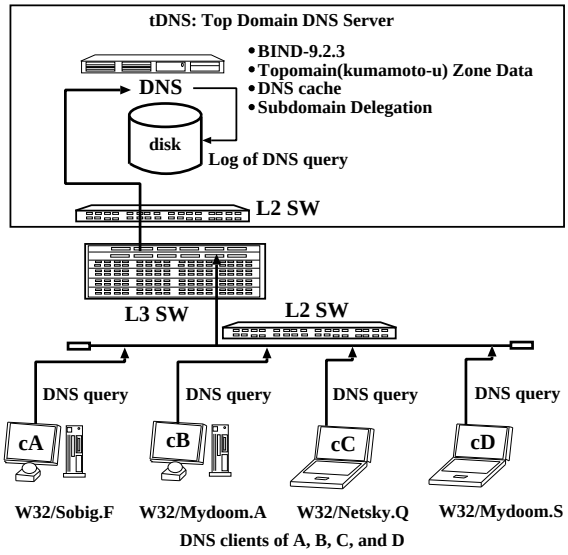

Figure 1. A schematic diagram of a network observed in the present study.

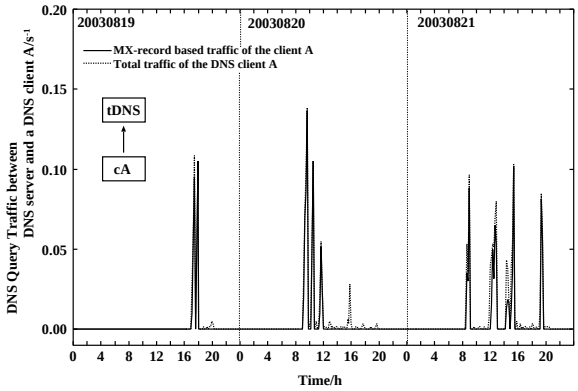

Figure 2. Traffic of the DNS query access between the top domain DNS server and the DNS client A through August 19th to 21st, 2003. The dotted line shows the total DNS traffic and the solid line indicates the MX-record based DNS traffic (s^{-1} unit).

2.2 A method of analysis

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

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

query contents to get how many MX, A, and PTR records. These procedures are properly worked out to an **mmwip** program compiled with gcc-2.95.3 C compiler.

3. Results and Discussion

3.1 Mass Mailing Worm

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 show 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 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[1]. 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.

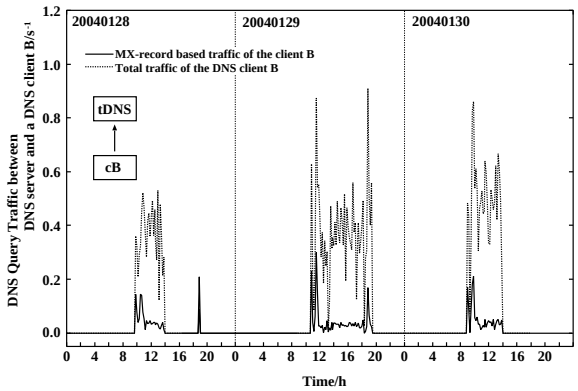

Figure 3. Traffic of the DNS query access between the top domain DNS server and the DNS client B through January 28th to 30th, 2004. The dotted line shows the total DNS traffic and the solid line indicates the MX-record based DNS traffic (s^{-1} unit).

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

When we see the syslog file for DNS query packets from **cB**, we encounter head lines in which, "mx.xxxxx.co.jp", "mail.xxxxx.co.jp", "smtp.xxxxx.co.jp", "mx1.xxxxx.co.jp", "mxs.xxx-xx.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[2]. 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 result 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."[2]. Therefore, this scan generates a lot of A record packets more than MX record ones.

It is clear that the DNS query traffic of the DNS clients like Windows PC terminals includes MX and A record packets without PTR record packet when infected with the MMW like W32/Sobig.F MMW and W32/Mydoom.A MMW.

3.2 Development of MXRPDS

We have designed and developed an automatic MX record packet detection system (MXRPDS) consisting of DNS query MX and PTR record packets capture, PTR-record packet preprocessor "apra", a detection engine "mscan", and alert mailer "smail". The "mxrpds.pl" script hooks up the "mscan" script in order to scan the syslog file of **tDNS** in a time per 10 seconds.

In the DNS query packet capture, DNS query MX and PTR record request packets and their contents are recorded and decoded with the query-logging system of BIND-9.2.3[15].

The preprocessor "arpa" changes a description format of an IP address in the content of PTR record packet, like sorting "D.C.B.A.in-addr.arpa" to "A.B.C.D", where A, B, C, and D indicate 8 bits unsigned integer (0-255) values. This is because the described IP address in the content of PTR record is complicated for the detection engine. This "arpa" is a packet filter to be sensitive only for a string that includes a key word as "in-addr.arpa" and it is compiled with the gcc-2.95.3 C compiler. The preprocessor is called in the following detection engine and prints out the filtered contents of the PTR record packets into a "newdb" file and the old "newdb" file is renamed as an "olddb" file.

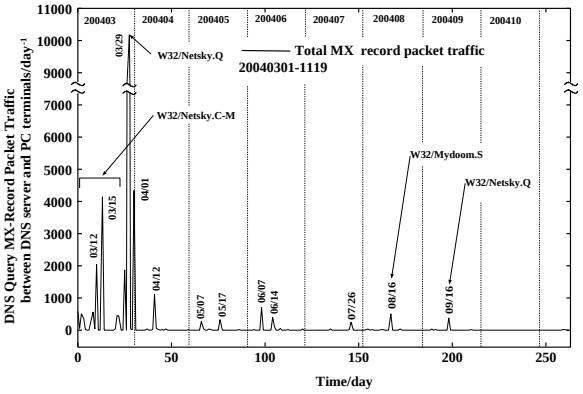

Figure 4. Traffic of the DNS query MX record packets access between the top domain DNS server and the PC terminals through March 1st to November 19th, 2004 (day^{-1} unit).

The detection engine "mscan" is a C-shell script program consisting of four components, an MMW IP filter "mmwip", a difference checker for MX record, a difference checker for PTR record, and an E-mailer without a local MTA "smail". The "mmwip" program compiled by the gcc-2.95.3 C compiler is a filter to discard the registered and/or authorized E-mail servers. The difference checker for MX (PTR) record packet is a "diff" command to check difference between "olddb" ("polddb") and "newdb" ("pnewdb") files. Before the difference checker for PTR record packet, the preprocessor "arpa" is called. After the checker, if the "newdb" file differs from the "olddb" one, the following two results are obtained: (1) If the "pnewdb" file is equal to the "polddb" one, the difference means detecting a MMW-infected PC terminal. (2) If the "pnewdb" file differs the "polddb" one, the difference shows detecting an E-mail server. These results are e-mailed to a network manager by the "smail" program.

3.3 Evaluation

We implemented MXRPDS into the top domain DNS (**tDNS**) server and evaluated detection rate (March 1st, 2004). The machine in the evaluation

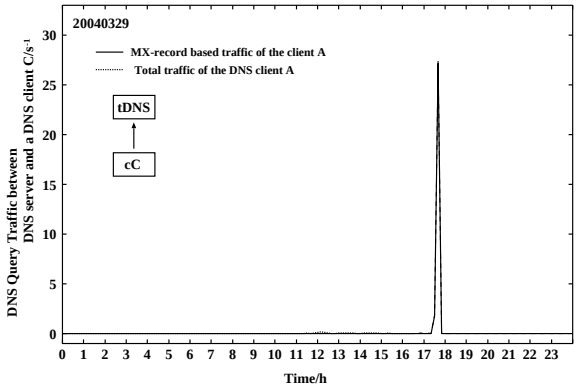

Figure 5. Traffic of the DNS query access between the top domain DNS server and the DNS client C at March 29th, 2004. The dotted line shows the total DNS traffic and the solid line indicates the MX-record based DNS traffic (s^{-1} unit).

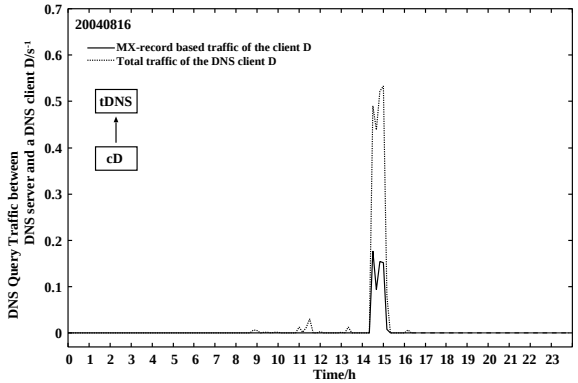

Figure 6. Traffic of the DNS query access between the top domain DNS server and the DNS client D at August 16th, 2004. The dotted line shows the total DNS traffic and the solid line indicates the MX-record based DNS traffic (s^{-1} unit).

has the following configuration: Intel Xeon 2.40GHz Dual CPU, 1GB main memory, Intel 100Mbps Ethernet NIC, and 80 GB ATA133 hard disk drive (The Linux kernel is currently to be a version of 2.4.27).

In Figure 4, the detection rate of the illegal DNS query MX record packet traffic increases when the W32/Netsky.C-Q MMWs are released suddenly. Especially, we can see the most largest peak in March 29th, 2004 when it is successfully detected IP addresses of the W32/Netsky.Q MMW-infected PC terminals. At the day, the W32/Netsky.Q MMW was only unidentified MMW by several anti-virus makers. However, we were able to quickly smash out the new W32/Netsky variant from our private LAN. In Figure 5, the DNS query traffic from the client C (**cC**) of the W32/Netsky.Q PC terminal includes mainly MX record packets. When seeing the initial stage of its log messages, the keyword "yahoo.cz" is written[3](in case of W32/Netsky.C-M: "yahoo.com" is written).

In August 16th, 2004, we have also successfully detected IP addresses of the W32/Mydoom.S (alias: W32/Ratos.A). In Figure 6, the DNS query traffic from the client C (**cD**) of the W32/Mydoom.S MMW-infected PC terminal includes mainly both MX and A record packets and the traffic of A record packets is considerably greater extent than that of MX record ones. This feature is also observed in the DNS query traffic from the W32/Mydoom.A MMW-infected PC terminals.

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: 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), W32/Mydoom.A, W32/Netsky.Q, and W32/Mydoom.S. From this point, we have developed and implemented an indirect virus detection system (MXRPDS) into our top domain DNS server. Successfully, we have been preventing the current MMW-infection.

We continue further investigation to get more detailed information on the automatic prevention system of MMW infection[20−22].

Acknowledgement. All the calculations and investigations were carried out in Center for Multimedia and Information Technologies, Kumamoto University. We specially thank to techni-

cal 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) http://www.trendmicro.com/vinfo/virusencyclo/default5.asp?VName=WORM_SOBIG.F

2) http://www.trendmicro.com/vinfo/virusencyclo/default5.asp?VName=WORM_MYDOOM.A

3) http://www.trendmicro.com/vinfo/virusencyclo/default5.asp?VName=WORM_NETSKY.Q

4) http://www.trendmicro.com/vinfo/virusencyclo/default5.asp?VName=WORM_NACHI.A

5) http://www.trendmicro.com/vinfo/virusencyclo/default5.asp?VName=WORM_SASSER.D

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

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

8) 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.

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

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) 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.

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

14) http://www.snort.org/

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

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

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

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

19) Matsuba, R., Musashi, Y., and Sugitani, K.: Detection of Mass Mailing Worm-infected IP address by Analysis of Syslog for DNS server, *IPSJ SIG Technical Reports, Distributed System and Management 32nd*, Vol. 2004, No.37, pp.67-72 (2004).

20) Musashi, Y., Matsuba, R., and Sugitani, K.: Development of Automatic Detection and Prevention Systems of DNS Query PTR record-based Distributed Denial-of-Service Attack, *IPSJ SIG Technical Reports, Distributed System and Management 34th*, Vol. 2004, No.77, pp.43-48 (2004).

21) Musashi, Y., Matsuba, R, Sugitani, K, and Moriyama, T: Workaround for Welchia and Sasser Internet Worms in Kumamoto University, *Journal for Academic Computing and Networking*, No.8, pp.5-8 (2004).

22) Musashi, Y., Matsuba, R., and Sugitani, K., "Indirect Detection of Mass Mailing Worm-Infected PC terminals for Learners", *Proc. the 3rd International Conference on Emerging Telecommunications Technologies and Applicatications (ICETA2004)*, Košice, Slovakia, pp.233-237 (2004).