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

2005－DSM－38（5）
2005／8／5

プロトコル異常検知による A レコード型 DNS パケット分散サービス妨害攻撃の阻止

武藏 泰雄 †　松葉 龍一 †　杉谷 賢一†

概要: ある大学の DNS サーバに対する大量の DNS アクセスがあり、そのアクセスに含まれる A レコード型 DNS クエリパケットのコンテンツについて調査した。その結果、(1) 大量メール送信型型ワーム (MMW) のワーム活動や spam メール発信活動に関する A レコード型の DNS クエリパケットのコンテンツには、主として “mail”, “smtp”, “mx”, “ns”, “relay” および “gate” 等のキーワードが含まれていることが判明し、また、(2) スパイウェアやボットウィルスに乗っ取られている思われる PC 端末からの A レコード型 DNS クエリパケットのコンテンツには、IP アドレスがキーワードとして直接記述されていることが見い出された。以上結果から、A レコード型 DNS クエリパケットのクエリコンテンツを監視することで、メール型ワームに感染している、また、ボットトネットワークを構成単位としての PC 端末の IP アドレスを検知することが可能であることが判明した。

Prevention of A-record based DNS Query Packets Distributed Denial-of-Service Attack by Protocol Anomaly Detection

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

Abstract: The contents of the A record based DNS query packets between the DNS server and DNS clients in a university were investigated when receiving a large amount of DNS resolution access. The interesting results are: (1) Keywords of “mail”, “smtp”, “mx”, “ns”, “relay”, and “gate” are mainly included in the contents of the A record based DNS query packets from the PC clients that are infected with the Mass Mailing Worm (MMW) and/or hijacked as a spam mail sender, and (2) the IP addresses are directly described in the contents of the A record based DNS query packets from the PC clients that are infected with the spyware or the bot network virus. Therefore, we can clearly detect IP addresses of the PC clients that are the MMW-infected and/or bot network virus by only monitoring the contents of the A record based DNS query packets.

1. Introduction

It is well-known that the DNS server provides very important information such as a fully qualified domain name (an FQDN, an A record, standard access), an IP address (a PTR record, reverse access), and mail exchange (an MX record, mail exchange), to DNS clients like E-mail server (SMTP/POP3) and/or WWW browsing network applications. In other words, these network applications strongly depend on the DNS server. From this point, the DNS query packets provide us very important information at initial stage of the network applications like, for example, the FQDN of the looking Web site by the Web browser, and the IP addresses of the PC clients and the FQDN of the E-mail servers by SMTP server daemon program. These applications are deeply related to many recent security incidents like mass mailing worm-infection and/or pre-scanning for a distributed denial-of-service (DDoS) attack to the network servers such as E-mail and Web servers. Therefore, observing traffic of DNS query packets probably provides us useful information of the security incidents and to develop the intrusion detection/prevention system (IDS/IPS).[1−7] Rikitake *et al.* have reported that the DNS server can be considered to be one of the IDS components.[8] Also,

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

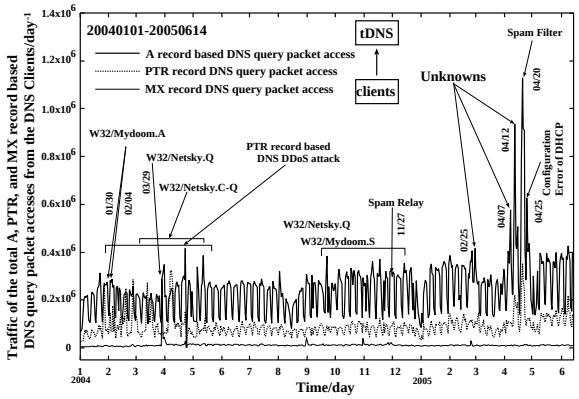

Figure 1. The DNS query traffic between the top domain DNS (**tDNS**) server and the DNS clients through January 1st, 2004 to June 14th, 2005. The thick solid line shows the A record based DNS query packet access, the dotted solid line indicates the PTR record based DNS query packet access, and the thin solid line demonstrates the MX-record based DNS query packet access (day^{-1} unit).

we have reported that we can detect IP addresses of the Mass Mailing Worm (MMW)-infected PC clients by observing the MX record based DNS query packet from the PC clients and developed the detection- and prevention-system against the PTR record based DNS query DDoS attack.[9,10]

Recently, the top domain name system (DNS) server of a university has been under several DNS query packet-based DDoS attacks like transmitting various kinds of DNS query packets, probably because in order to crash the DNS server, to search the FQDNs, the FQDNs of E-mail servers, and the IP addresses of the next victim PC clients, and to be a base of DoS attack. Especially, traffic of the A record based DNS query packets to the top domain DNS server of the university was abnormally increased during the early days of January, 2005 to the middle days of June, 2005 (see Figure 1).

The present paper discusses (1) on the investigation of three different kinds of DDoS attacks through February 25th, April 7th, and 12th, 2005, (2) on correlation analysis of the A record based DNS query packets traffic between the DNS server and the DNS clients that especially transmit strange query contents of the A record based DNS query packets including fully qualified domain names (FQDN) of E-mail servers and IP addresses directly, (3) how to implement a DNS query packets-based DDoS attack detection system into the DNS server, and (4) how to prevent the DoS attack, effectively.

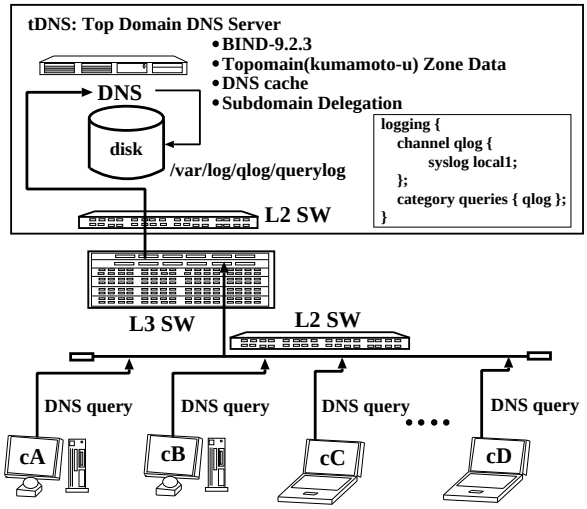

Figure 2. A schematic diagram of a network observed in the present study.

2. Observations

2.1 Network System

We investigated traffic of DNS query accesses between the top domain DNS server (**tDNS**) and the DNS clients. Figure 2 shows an observed network system in the present study and an optional configuration of the BIND-9.2.3 server program daemon[12] of the **tDNS**. The **tDNS** is one of the top level domain name (kumamoto-u) system servers and plays an important role of domain name resolution and subdomain delegation services for many PC clients and the subdomain network servers, respectively, and the operating system is Linux OS in which the kernel-2.4.30 is currently employed with the 1GB core memory and 100Mbps EthernetPro Intel Network Interface Card.

2.2 Capture of DNS Query Packets

In **tDNS**, BIND-9.2.3 program package has been employed as a DNS server daemon.[12] The DNS query packets and their contents have been captured and decoded by a query logging option (Figure 2, see % man named.conf). The log of DNS

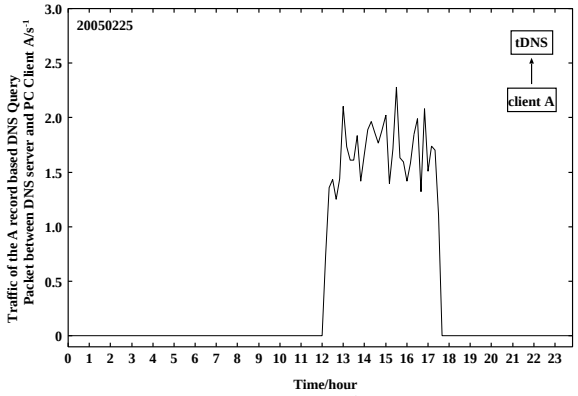

Figure 3. The traffic of the A record based DNS query packet access between the top domain DNS (**tDNS**) server and the DNS client A at February 25th, 2005 (s^{-1} unit).

query access has been recorded in the syslog files. All of the syslog files are daily updated by the crond system. The line of syslog message mainly consists of the content of the DNS query packet like a time, a source IP address of the DNS client, a fully qualified domain name (an A record type), an IP address (a PTR record type), and mail exchange (an MX record type).

2.3 Abnormal Traffic of the A Record based DNS Query Access from the Client A

Firstly, we observed traffic of the A record based DNS query packets from a DNS client A to the top domain DNS (**tDNS**) server through the day of February 25th, 2005 (Figure 3), because the client A is one of the top DNS query access clients at the day.

In Figure 3, the traffic starts from 12:00 and ends after 17:30. We noticed this abnormal traffic 17:30 and we filtered this DNS query access. The numbers of the total DNS query packets, the A record based DNS query packets, and the PTR record based ones, are obtained to be 32,728/day, 32,721/day, and 7/day, respectively, and no MX record based packet can be observed. This result shows that the total DNS query access traffic from the client A almost consists of the A record based DNS query access traffic.

We can demonstrate statistics of the contents for the A record based DNS query packets from

1		2		3		4		5	
m	9975	ma	7506	mai	7404	mail	7399	mail.	5894
s	1569	mx	1883	smt	872	smtp	872	smtp.	491
p	566	sm	888	mx1	583	mx1.	451	mail1	229
a	542	in	265	mx0	402	rela	195	mailh	201
c	490	re	237	mx.	378	mx2.	167	mail2	200
i	462	po	231	rel	196	inbo	134	relay	190
n	403	ns	153	mx2	171	spam	101	mailg	162
b	395	sp	143	inb	134	mx01	92	inbou	133
r	363	co	132	pop	118	www.	91	mail-	129
e	341	ba	120	spa	108	serv	79	mails	108
				www	96	mx3.	79	smtp1	96
				bar	85	pop.	76	mx01.	90
				ser	82	barr	73	mail0	74
				mx3	82	post	69	barra	73
				pos	75	emai	67	smtp-	72
				mx-	70	gate	64	serve	70
				gat	67	filt	51	email	67
				ema	67	mx0.	49	mail3	65
				cor	62	mx4.	47		
				web	57				
				ns.	55				
				mta	55				

Figure 4. Statistics of the contents for the A record based DNS query packets from the client A at February 25th, 2005.

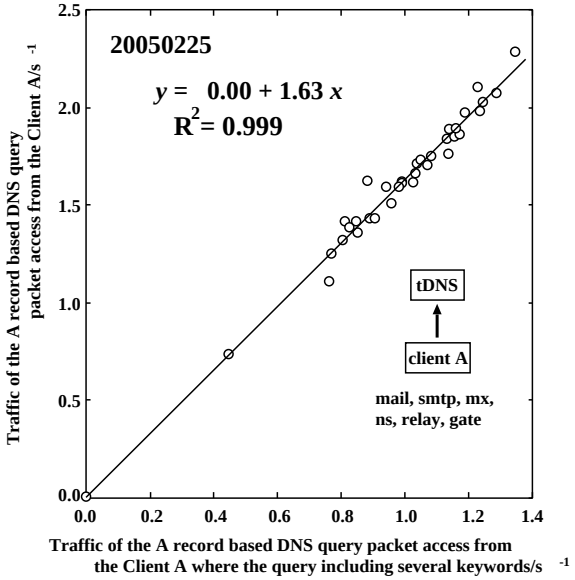

Figure 5. Total traffic of the A record based DNS query packet access from the client A versus traffic of the A record based DNS query packet access from client A where including the six keywords at February 25th, 2005 (s^{-1} unit).

the client A at February 25th, 2005 (Figure 4). In Figure 4, the keywords of "mail", "smtp", "mx", "ns", "relay", and "gate" are used to generate fully qualified domain names of the E-mail servers that have ever been observed when detecting IP addresses of the W32/Mydoom.A or W32/Mydoom.S mass mailing worm (MMW)-infected PC clients,[9,11] *i.e.* the PC client A is probably infected with a new type of mass mailing worm (MMW) which resembles well W32/Mydoom.A or W32/Mydooms.S. This new worm was assigned to be the W32/Mytob.A bot worm (BW) after February 27th, 2005 by several anti-virus vendors.

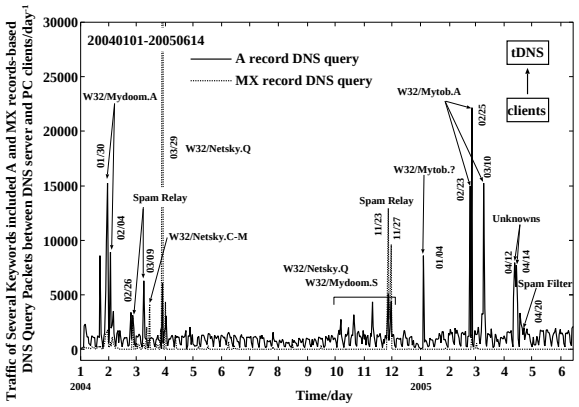

Figure 6. Total traffic of the A record based DNS query packet access including the six keywords ("mail", "smtp", "mx", "ns", "relay", and "gate") in the top domain DNS server (**tDNS**) through January 1st, 2004 to June 14th, 2005. (day^{-1} unit).

Why do the W32/Mydoom.A-S MMWs and the W32/Mytob.A BW decrease sending the MX record based DNS query packet access? This is because SMTP process needs the DNS resolution twice: one is a mail exchange (MX) resolution (sending the MX record based DNS query packet) to get an FQDN of the E-mail server and the other is standard resolution (sending the A record based DNS query packet) to convert the FQDN into an IP address. In order to save the time for the former MX resolution as possible, the W32/Mydoom.A or W32/Mydoom.S MMW is improved to complete or convert the harvested generic domain name from the PC hard disk drive into the FQDN.

Figure 5 shows regression analysis on the total traffic of the A record based DNS query packet access from the client A versus the traffic of the A record DNS query packet access from the client A in which the six keywords of "mail", "smtp", "mx", "ns", "relay", and "gate" are included. The data February 25th, 2005. In Figure 5, the correlation coefficient (R^2) is 0.999. This means that the traffic of the A record based DNS query packet including the six keywords strongly correlates with the abnormal traffic of the A record based DNS query packets from the client A.

From this point, we have further investigated on the total traffic of the A record based DNS query packet access that includes the several keywords consisting of head characters in the FQDNs for E-mail servers.

client B		client C	
0.0.0.0	26	***.***.y****.com	12
.**-u.ac.jp	13	www.*****m.com	7
133.9*.**.192	11	yahoo.co.jp	6
133.9*.**.73	10	www.****.****.co.jp	6
133.9*.**.66	9	mail.****.com	6
133.9*.**.64	9	img.****.co.jp	5
133.9*.**.52	9	i.****.jp	5
133.9*.**.89	6	ai.****.jp	5
mil.***.********-u.ac.jp	5	133.9*.***.194	5
.**.*****-u.ac.jp	5	133.9*.20*.2**	5
2**.*.2**.*8	5	127.0.0.1.***-u.ac.jp	5
133.9*.**.9	5	127.0.0.1	5
133.9*.**.8	5	relay.****.net	4
133.95.**.7	5	rd.*****.co.jp	4

Figure 7. Statistics of the contents for the A record based DNS query packets from the clients B and C at April 7th and 12th, respectively.

3. Results and Discussion

3.1 Query Content-based Scanning

We illustrate the observed total traffic of the A record based DNS query packet access including the six keywords ("mail", "smtp", "mx", "ns", "relay", and "gate") and the MX record based DNS query packet access from the PC clients without any Web/E-mail servers through January 1st, 2004 to June 14th, 2005, as shown in Figure 6. Interestingly, we can find several new peaks of, for instance, January 22nd, (Hijacked PC), February 23rd (W32/Mytob.A), March 9th (Spam Relay), and November 23rd (Spam Relay), 2004, and January 4th (MMW like a W32/Mytob.A?), February 23rd (W32/Mytob.A), and April 14th (Unknown), 2005. Surprisingly, the largest peak for traffic curve of the A record DNS query packets access at April 20th, 2005 (Figure 1), disappears considerably (Figure 6), probably because in the day, a lot of DNS resolving access from the internet in order to get a FQDN/DN or their IP addresses of the E-mail server as a spam relay in the university. Also, the peak at April 12th, 2005, are unexpectedly smaller than that in Figure 1 and the peak at April 7th, 2005, disappears. From these results, we need, therefore, to investigate further on the unknown two peaks at April 7th and 12th, 2005.

The clients B and C are the top DNS query access clients (229,309/day and 400,964/day) that belong to each of peaks at April 7th and 12th, respectively, and statistics for their query contents are shown in Figure 7. In Figure 7, we can clearly

notice that IP addresses are directly included in the query contents in spite of the A record DNS query packets. Usually, query contents of the A record based DNS query packets only include fully qualified domain names (FQDNs). The number of IP addresses are calculated to be 161,329 (70.4%) and 200,645 (50.0%) for the days of April 7th and 12th, respectively.

Figure 8 demonstrates regression analysis on the total traffic of the A record based DNS query packet access from the clients B and C versus traffic of the A record based DNS query packet access from clients B and C including the IP addresses. The data are April 7th and 12th, 2005 and the correlation coefficients (R^2) are 0.994 and 0.999 for clients B and C, respectively. This means that the total traffic of the A record based DNS query packet access from the clients strongly correlates with the traffic of ones that include directly IP addresses. In other words, this feature is useful to detect the abnormal traffic of the A record based DNS query packet access from the PC clients.

We illustrate the observed traffic of the A record based DNS query packets including IP addresses directly from the PC clients of the university, as shown in Figure 9. In Figure 9, we can reproduce the abnormal traffic of the A record DNS query packets from the PC clients that include IP addresses directly at April 7th and 12th, and we can also find new peaks at April 11th and 15th.

As a result, it is clear that (1) the query contents for the abnormal traffic of the A record based DNS query packets mainly include the six keywords ("mail", "smtp", "mx", "ns", "relay", and "gate") and/or directly IP addresses, (2) the abnormal traffics strongly correlate with the total traffic of the A record based DNS query packets, and (3) this feature can be useful for detecting abnormal traffic of the A record based DNS query packets from the PC clients that are infected with or hijacked by the bot worm like W32/Mytob.A.

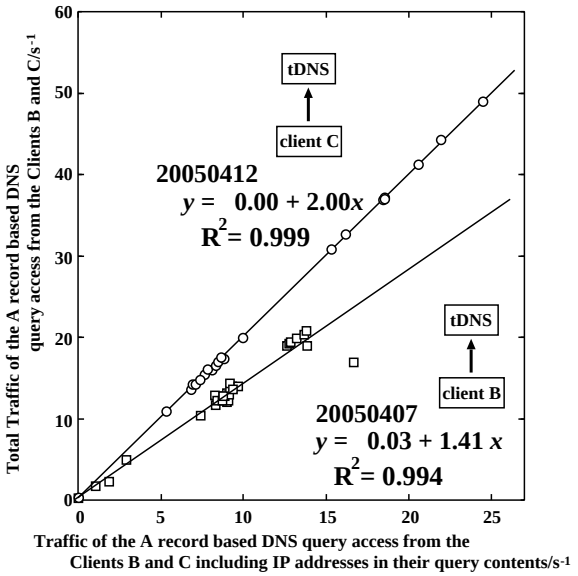

Figure 8. Total traffic of the A record based DNS query packet access from the Clients B and C versus traffic of the A record based DNS query packet access from clients B and C including the IP addresses at April 7th and 12th, 2005 (s^{-1} unit).

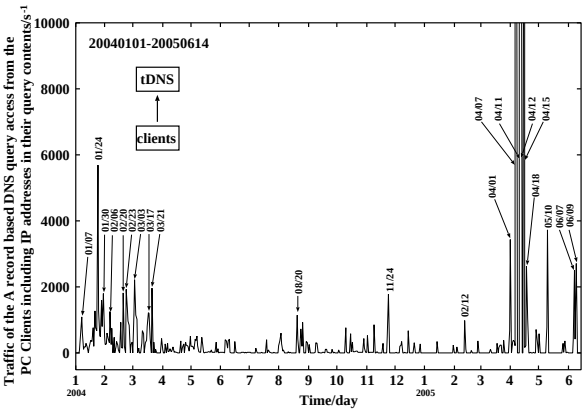

Figure 9. Traffic of the A record based DNS query packet access including IP addresses to the top domain DNS server (**tDNS**) through January 1st, 2004 to June 14th, 2005. (day^{-1} unit).

3.2 ADPS

We designed and developed new detection- and prevention-system against the abnormal traffic of the A record based DNS query packet access that include the six keywords and IP addresses directly as their contents (ADPS). The prevention part of the new system (ADPS) is the same as the previously reported system of PTRDPS.[10] The detection part of the ADPS checks the syslog messages (including client source IP addresses and their query contents by an optional configuration of BIND-9.2.3, see Figure 2) of the DNS server program daemon with the six keyword in the head

word of an FQDN or an IP address as "A.B.C.D" (A,B,C, and D are a digit number:0-255) in the query contents of the A record based DNS query packets. The sampling rate is arbitrarily fixed in a time per 1h. And if the traffic becomes greater than the threshold (=500/hour), it can be concluded that IP addresses of the suspicious traffic is detected. Then, the IP addresses are sent to the prevention part of the ADPS. This ADPS has been installed into the top domain DNS (**tDNS**) server of the university after the day of April 25th, 2005 and the abnormal traffic of the A record DNS query packets has been decreased after the day (see Figures 6 and 9).

4. Concluding Remarks

We statistically investigated syslog files of the top domain DNS server (**tDNS**) in a university when observing abnormal traffic of the A record based DNS query packets and we have concluded that the abnormal traffic are detectable because the six keywords and IP addresses themselves are included in the query contents of the A record DNS query packets. From these results, we have developed and installed the detection- and prevention-system (ADPS) into the **tDNS**, we are currently testing it, and we start to investigate why the total traffic of the A record based DNS query packets are still increasing (Figure 1).

Acknowledgement. All the studies were carried out in CMIT of Kumamoto University and M-Lehrstuhl in Johann Wonfgang Goethe University of Frankfurt am Main. We gratefully thank to all the CMIT staffs and this study is a grant aid of Promoting Advanced Educational Program (2004) for Overseas Dispatch by the Ministry of Education, Culture, Science and Sport.

References and Notes

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

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

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

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

5) 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).

6) 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.

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

8) Rikitake, K., Nogawa, H., Tanaka, T., and Shimojo, S.: Behavioral Analysis of DNS and TCP Connections, *Computer Security Symposium 2003 (CSS2003), IPSJ Symposium Series*, Vol. 2003, No.15, pp.521-526 (2003).

9) 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).

10) 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).

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

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