DNS Based Detection of SSH Dictionary Attack in Campus Network

Dennis Arturo Ludeña Romaña,* Yasuo Musashi,** Kazuya Takemori,*
Masaya Kumagai,* Shinichiro Kubota,** Kenichi Sugitani,**
Tsuyoshi Usagawa,* and Toshinori Sueyoshi*

**Graduate School of Science and Technology,*
Kumamoto University, Kumamoto-City, 860-8555, JAPAN
E-mail:dennis@st.cs.kumamoto-u.ac.jp

***Center for Multimedia and Information Technologies,*
Kumamoto University, Kumamoto-City, 860-8555, JAPAN
E-mail:musashi@cc.kumamoto-u.ac.jp

Abstract

We statistically investigated the DNS query access traffic from a university campus network toward the top domain DNS through March 14th, 2009, when the hosts in the campus network were under inbound SSH dictionary brute force attack. The interesting results are obtained, as follows: (1) the several hosts generated the DNS query packet traffic, taking a rate of more than 1,000 hour^{-1}, through 07:30-08:30 in March 14th, 2009, (2) the DNS query packet traffic correlates with the DNS query packet one including more than two specific query keywords (payloads of the packets), and (3) the former keyword is a fully qualified domain name (FQDN) and the latter one is an IP address. Therefore, we can detect inbound SSH dictionary attack by watching frequencies of the FQDNs and the IP addresses as query keywords in the DNS query packets from the hosts in the campus network.

Keywords: Detection, SSH dictionary attack, SSH brute force attack

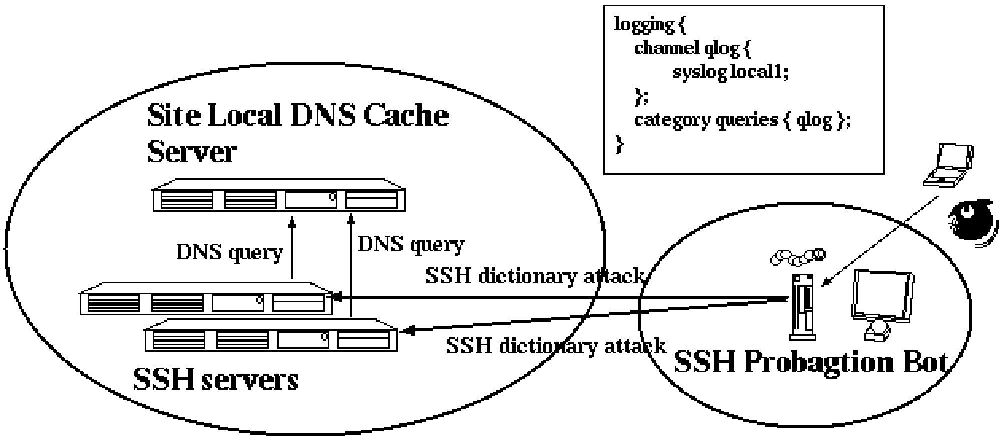
1. Introduction

It is of considerable importance to increase up a detection rate of the SSH dictionary attack bots, since they become components of the bot clustered networks [1-6]. Unfortunately, the SSH dictionary attack (the brute force attack) has been still used to spread out the bots when hijacking the specific vulnerable network servers on the Internet [7, 8]. This is because the network servers can be easily connected with the SSH clients when the attackers know their user IDs and pass phrases, or when, in other words, the account holders use easy breakable pass phrases. Therefore, it is also important to develop detection technologies as countermeasures against the SSH dictionary attack [7, 8].

Recently, several researchers reported prevention technologies for the SSH dictionary attack by employing the distributed and cooperative active response architectures [9, 10]. Currently, we can find the SSH dictionary attack related alert messages from the IDS/IPS or logging agents (sensors) in the network servers, in which these systems, however, observe directly the SSH communication related packets, and they need a cost of installation, update of their security appliances or network configurations.

Previously, on the other hand, we reported that the DNS based detection technologies of the random spam bot activity in the campus network and the host search (HS) activity against the campus top domain name system (**tDNS**) servers [11, 12]. The DNS based detection system has a merit which observes only the DNS query request packet traffic between the DNS server and its clients *i.e.* the DNS resolver has been already installed in almost all the network appliances like PC terminals, routers, switches, servers, etc.

In this paper, (1) we carried out statistical analysis on the total- A- and the PTR-resource records (RRs) based DNS query packet traffic from the SSH dictionary attacked Linux PC network servers through March 14th, 2009, and (2) we assessed the detection rate of the SSH



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

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

dictionary attack based DNS query traffic in the total DNS query packet traffic from the campus network through January 1st to December 31st, 2009.

2. Observations

2.1 Network Systems and DNS Query Request Packet Capturing

We investigated on the DNS query request packet traffic between the top domain (**tDNS**) DNS server and the network servers as DNS clients, corresponding to the Site Local DNS Cache Server and SSH servers, respectively, in Figure 1. The **tDNS** server is one of the top level domain name (kumamoto-u) system (DNS) servers and it plays an important role of domain name resolution including DNS cache function. The operating system is Linux OS (CentOS 4.3 Final) in which the kernel-2.6.9 is currently employed with the Intel Xeon 3.20 GHz Quadruple SMP system, the 2GB core memory, and Intel 1000Mbps EthernetPro Network Interface Cards

In the **tDNS** server, the BIND-9.2.6 program package has been employed as a DNS server daemon [13]. The DNS query request packets and their query keywords have been captured and decoded by a query logging option in the *named.conf* file (see Figure 1 and the *named.conf* manual of the BIND program in more detail).

```
Oct 12 08:38:24 kun named[533]: client 133.95.xxx.yyy#39815: query: 130.13.194.xxx.in-addr.arpa IN PTR
Oct 12 08:38:25 kun named[533]: client 133.95.xxx.yyy#39825: query: dmea.net IN MX
Oct 12 08:38:43 kun named[533]: client 133.95.xxx.yyy#40010: query: mxwall03.hkabc.net IN A
```

Figure 2. Strucure of syslog messages generated by BIND program packages.

The log of DNS query request packet access has been recorded in the system log (syslog) files. All of the syslog files are daily updated by the CRON system. The line of syslog message consists of a time, a source IP address of the DNS client, a fully qualified domain name (A and AAAA resource record (RR) for IPv4 and IPv6 addresses, respectively) type, an IP address (PTR RR) type, or a mail exchange (MX RR) type (see Figure 2).

3. Results and Discussion

3.1 Total, A-, and PTR Resource records (RR) based DNS Query Packet Traffics

Firstly, we illustrate the observed total DNS query packet traffic, and A- and PTR-resource records (RRs) based DNS query request packet traffics from the campus network to the top domain name system (**tDNS**) server in March 14th, 2009, as shown in Figure 3.

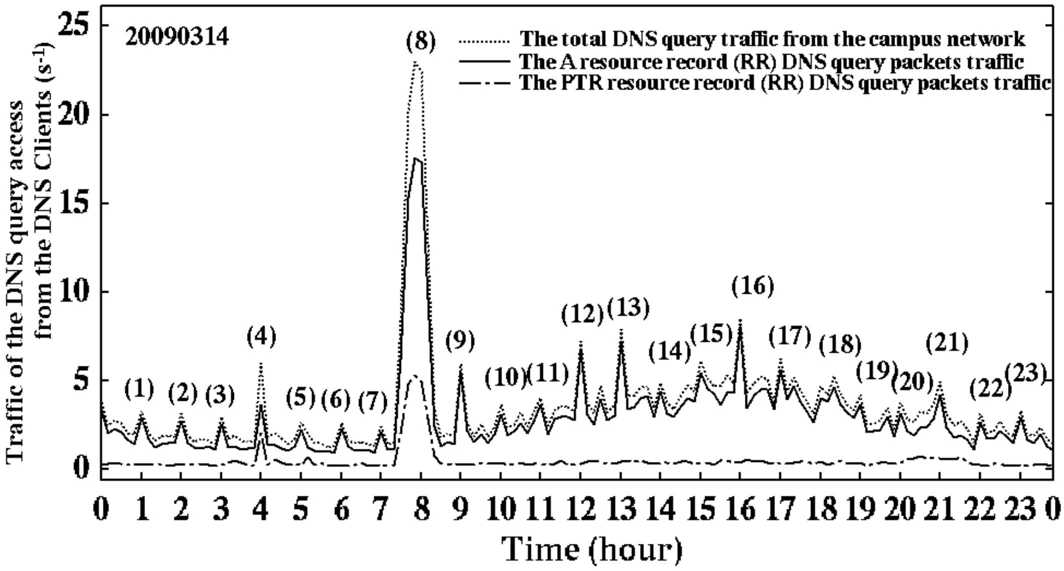
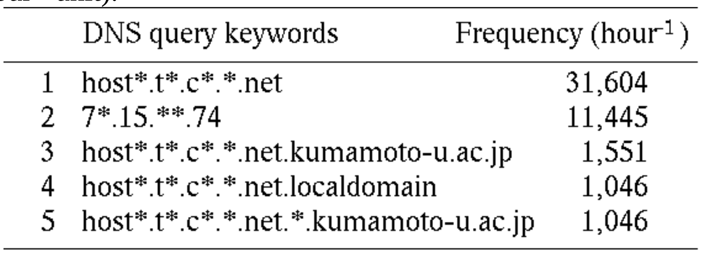


Figure 3. The total, A and PTR resource records (RRs) based DNS query packet traffics between the top domain DNS (**tDNS**) server and the DNS clients on the campus network at March 14th, 2009 (s^{-1} unit).

Table 1. Detected top unique query keywords and their frequencies in the total DNS query packet traffic from the university campus network through 07:30-08:30 March 14th, 2009 ($hour^{-1}$ unit).



	DNS query keywords	Frequency ($hour^{-1}$)
1	host*.t*.c*.*.net	31,604
2	7*.15.**.74	11,445
3	host*.t*.c*.*.net.kumamoto-u.ac.jp	1,551
4	host*.t*.c*.*.net.localdomain	1,046
5	host*.t*.c*.*.net.*.kumamoto-u.ac.jp	1,046

In Figure 3, we can find twenty three peaks and they are categorized into two groups, as: {(1)-(3),(5)-(7),(9)-(23)} and {(4),(8)}. In the former group, the total DNS query packet traffic correlates only with the A RR based DNS query packet traffic, while in the latter one, the total DNS query packet traffic does with the both A- and PTR-RRs based DNS query packet traffics, simultaneously.

We investigated statistics of the DNS query keywords in the total DNS query packet traffic in the peak (8) and the results are shown in Table 1. In Table 1, the top query keyword is a fully qualified domain name (FQDN) of the host on the Internet, and the second top one is an IP address corresponding to the FQDN.

Also, we investigated the syslog files in several network servers on the campus network, and then we observed the same FQDNs and the same IP addresses as the SSH clients on the Internet. Also, we found several signatures that the SSH clients tried repeatedly to login into the network servers in a short period *i.e.* this feature shows that the network servers were under SSH dictionary attack through 07:30 to 08:30 March 14th, 2009. Therefore, it can be concluded that we can detect the SSH dictionary attack based DNS query packet traffic when the pairs of the top FQDNs and the IP addresses are observed in the top DNS query keywords (Table 1).

3.2 A Newly Developed Detection Technology

We show a suggested system as a SSH dictionary attack detector and a newly developed algorithm for detecting the SSH dictionary attacked network servers in Figure 4.

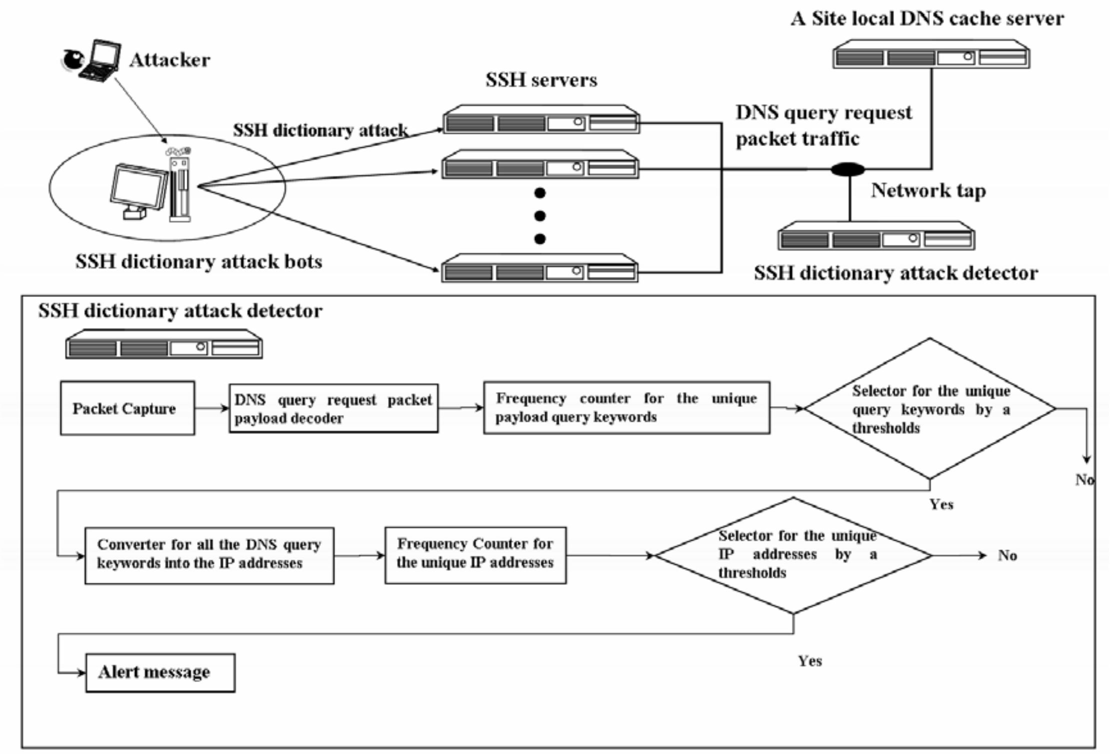
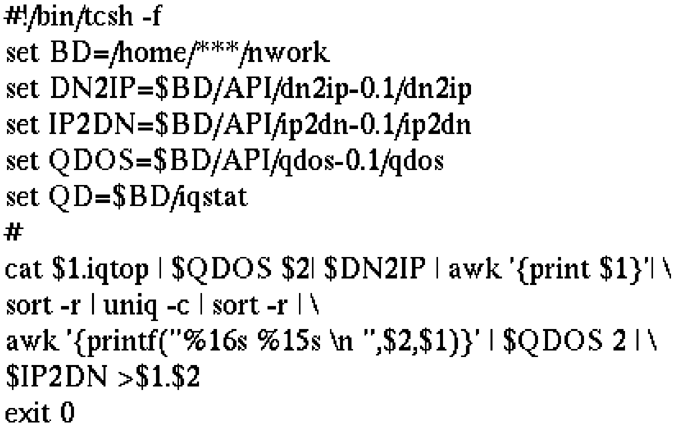

A Site local DNS cache server
Attacker
SSH servers
SSH dictionary attack
DNS query request packet traffic
Network tap
SSH dictionary attack bots
SSH dictionary attack detector
SSH dictionary attack detector
Packet Capture
DNS query request packet payload decoder
Frequency counter for the unique payload query keywords
Selector for the unique query keywords by a thresholds
No
Yes
Converter for all the DNS query keywords into the IP addresses
Frequency Counter for the unique IP addresses
Selector for the unique IP addresses by a thresholds
No
Yes
Alert message

Figure 4. A newly developed algorithm for detecting the SSH dictionary attacked network servers.



```
#!/bin/tcsh -f
set BD=/home/***/nwork
set DN2IP=$BD/API/dn2ip-0.1/dn2ip
set IP2DN=$BD/API/ip2dn-0.1/ip2dn
set QDOS=$BD/API/qdos-0.1/qdos
set QD=$BD/iqstat
#
cat $1.iqtop | $QDOS $2| $DN2IP | awk '{print $1}'| \
sort -r | uniq -c | sort -r | \
awk '{printf("%16s %15s \n ",$2,$1)}' | $QDOS 2 | \
$IP2DN >$1.$2
exit 0
```

Figure 5. The evaluation program for the DNS based detection system of the SSH dictionary attack to the network servers in the campus network.

The suggested system can be implemented the line between the SSH and the site local DNS cache servers by employing the network tap device. The suggested system consists of "Packet Capture," "DNS query request packet payload decoder," "Frequency counter for the unique payload query keywords," "Selector for the unique query keywords by a thresholds," "Converter for all the DNS query keywords into the IP addresses," "Frequency Counter for the unique IP addresses," "Selector for the unique IP addresses by a thresholds," and "Alert message."

In Figure 5, we show a script program to evaluate the detection rate for the newly suggested system. In working directory $BD/iqstat/, there are 365 files (20090101.qtop-20091231.qtop), including the unique DNS query keywords and their frequencies. These files are inputs for the evaluation programs and the outputs are several pairs of the specific IP addresses and their FQDNs. The results are written into the standard output and we can get the detection rate by only counting this output lines.

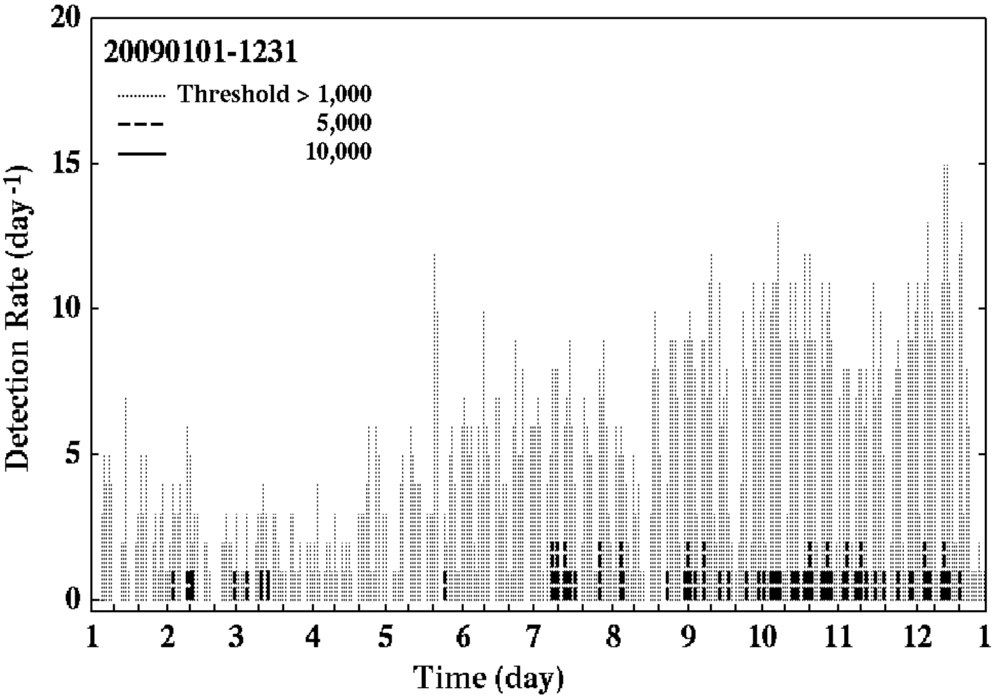
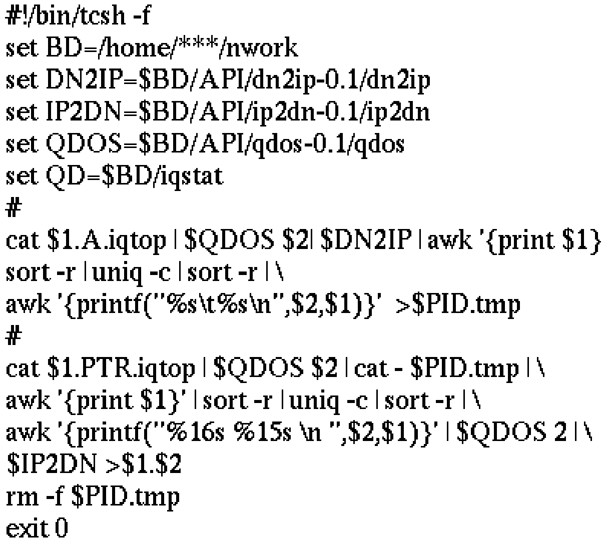


Figure 6. Frequency for the pairs of fully qualified domain names (FQDNs) and IP addresses as query keywords in the DNS query packet traffic from the campus network through January 1st to December 31st, 2009 (day^{-1} unit).

3.3 Evaluation

In order to confirm the possibility of the newly proposed detection system, we illustrate calculated frequencies for pairs of the FQDNs and the IP addresses as query keywords in the DNS query traffic from the campus network through January 1st to December 31st, 2009, by configuring three threshold values of 1,000, 5,000, and 10,000 day^{-1}, shown in Figure 6.

In Figure 6, we can observe a lot of peaks in a threshold value of 1,000 day^{-1}, however, these peaks can include much false positives, while in threshold values of 5,000 and 10,000 day^{-1}, we can observe not only several significant peaks, but they mean increase of false negatives.



```
#!/bin/tcsh -f
set BD=/home/***/nwork
set DN2IP=$BD/API/dn2ip-0.1/dn2ip
set IP2DN=$BD/API/ip2dn-0.1/ip2dn
set QDOS=$BD/API/qdos-0.1/qdos
set QD=$BD/iqstat
#
cat $1.A.iqtop | $QDOS $2| $DN2IP | awk '{print $1}'| \
sort -r | uniq -c | sort -r | \
awk '{printf("%s\t%s\n",$2,$1)}'  >$PID.tmp
#
cat $1.PTR.iqtop | $QDOS $2 | cat - $PID.tmp | \
awk '{print $1}' | sort -r | uniq -c | sort -r | \
awk '{printf("%16s %15s \n ",$2,$1)}' | $QDOS 2 | \
$IP2DN >$1.$2
rm -f $PID.tmp
exit 0
```

Figure 7. The improved evaluation program for the DNS based detection system of the SSH dictionary attack to the network servers in the campus network.

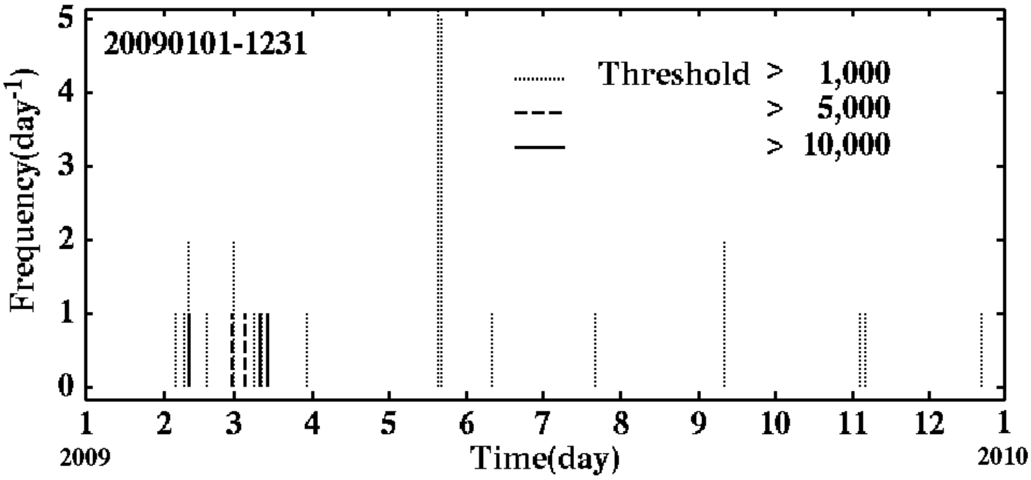


Figure 8. Frequency (improved) for the pairs of fully qualified domain names (FQDNs) and IP addresses as query keywords in the DNS query packet traffic from the campus network through January 1st to December 31st, 2009 (day^{-1} unit).

Fortunately, we found that several FQDNs were converted into an IP address. Usually, we can find it in the Internet that the IP address was assigned to several FQDNs or cites. To avoid this, we separated the A and PTR RRs based DNS query keywords from the input files ($BD/iqstat/{20090101-20091231}.iqstat). The improved evaluation program is shown in Figure 7 and the evaluated results are demonstrated in Figure 8.

In Figure 8, we can observe only nineteen peaks. This feature means that we successfully decreased the false positives like those in Figure 6 *i.e.* we can detect precisely the SSH dictionary attack to the network servers in the campus network by only observing the DNS query request packet traffic from the network servers as the DNS clients.

4. Conclusions

We performed traffic analysis on the DNS query packet access from the campus network to the top domain name system (**tDNS**) server through March 14th, 2009, when the campus network servers were under inbound SSH dictionary brute force attack and we obtained the following results, as: (1) The A resource record (RR) based DNS query packet traffic strongly correlates with the PTR RR based DNS query packet one, in 07:30-08:30, March 14th, 2009. (2) Also, we can observe the specific pairs of the top unique FQDNs and unique IP addresses in the total DNS query packet traffic *i.e*. we can detect the IP addresses of the network servers under inbound SSH dictionary attack. (3) We assessed the calculated detection rate for the proposed system and we can observe the SSH dictionary attack to the network servers in the campus network.

We further continue to develop new and low-costly detection technologies in the near future works, for example, we started to employ the clustering based detection models [14].

Acknowledgements

This study is supported by the Grant aid of Graduate School Action Scheme for Internationalization of University Students (GRASIUS) No. 165240040213 in Kumamoto University.

References

[1] P. Barford and V. Yegneswaran: An Inside Look at Botnets, Special Workshop on Malware Detection, *Advances in Information Security*, Springer Verlag, 2006.

[2] J. Nazario: Defense and Detection Strategies against Internet Worms, I Edition; *Computer Security Series*, Artech House, 2004.

[3] (a) J. Kristoff: Botnets, detection and mitigation: DNS-based techniques, *Northwestern Univerisity*, 2005, http://www.it.northwestern.edu/bin/docs/bots_kristoff_jul05.ppt. (b) J. Kristoff: Botnets, *North American Network Operators Group (NANOG32)*, Reston, Virginia (2004), http://www.nanog.org/mtg-0410/kristoff.html

[4] D. David, C. Zou, and W. Lee: Model Botnet Propagation Using Time Zones, *Proceeding of the Network and Distributed System Security (NDSS) Symposium 2006*; http://www.isc.org/isoc/conferences/ndss/06/proceedings/html/2006/

[5] A. Schonewille and D. v. –J. Helmond: The Domain Name Service as an IDS. How DNS can be used for detecting and monitoring badware in a network, 2006; http://staff.sciece.uva.nl/~delaat/snb-2006/p12/report.pdf

[6] B. McCarty: Botnets: Big and Bigger, *IEEE Security and Privacy*, No.1, pp.87-90 (2003).

[7] C. Seifert: Analyzing Malicious SSH Login Attempts, *Technical Report*, 2006 http://www.securityfocus.com/infocus/1876.

[8] D. Ramsbrock, R. Berthier, and M. Cukier: Profiling Attacker Behavior Following SSH Compromises, *Proceedings of the 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN07)*, Washington D.C., USA, IEEE Computer Society, 2007, p. 119-124.

[9] Y. Oosumi and N. Yamai: Technique of the countermeasure for brute force attack which can cooperate between the hosts, *IPSJ SIG Technical Reports, Distributed System and Management 47th (DSM47)*, Vol. 2007, No. 93, 2007, p.49-54.

[10] J. L. Thames, R. Abler, and D. Keeling: A distributed active response architecture for preventing SSH dictionary attacks, *Proceedings of the Southeastcon*, 2008, IEEE, Huntsville, AL, USA, 2008, pp. 84-89.

[11] D. A. Ludeña Romaña, Y. Musashi, R. Matusba, and K. Sugitani: Detection of Bot Worm-Infected PC Terminals, *Information*, Vol. 10, No. 5, 2007, pp. 673-686.

[12] D. A. Ludeña Romaña, S. Kubota, K. Sugitani and Y. Musashi: DNS Based Detection of Spam Bots and Host Search Activity, *IPSJ SIG Technical Reports, Internet Operation and Technology 3rd (IOT03)*, Vol. 2008, No. 87, 2008, pp.1-6.

[13] BIND-9.2.6: http://www.isc.org/products/BIND/

[14] M. Lei, Y. Musashi, D. A. Ludeña Romaña, K. Takemori, S. Kubota, and K. Sugitani: Detection of Host Search Activity in Domain Name Reverse Resolution Traffic, *IPSJ Symposium Series (IOTS2009)*, Vol. 2009, No. 15, 2009, pp.91-94.