# DNS-based Countermeasure Technologies for Spam Bot Worm-infected PC terminals in the Campus Network

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

[†]Centre of Multimedia and Information Technologies
Kumamoto University 860-8555 JAPAN
E-mail:musashi@cc.kumamoto-u.ac.jp
Phone +81-96-342-3915 Fax +81-96-342-3829

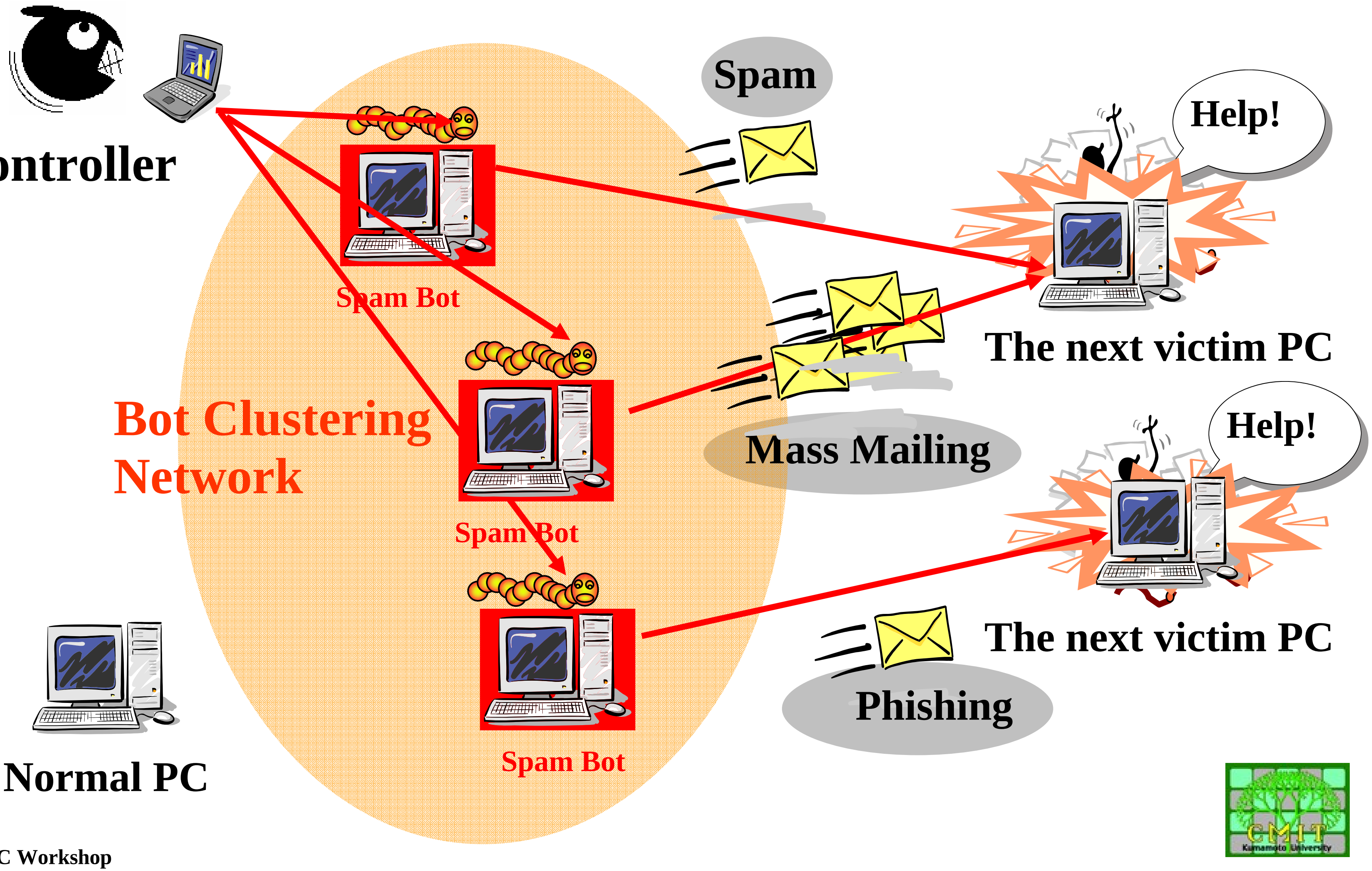
Dennis A. Ludeña R.[††] and Hirofumi Nagatomi[††]

[††]Graduate School of Science and Technology
Kumamoto University, 860-8555, JAPAN
E-mail: {dennis}@st.cs.kumamoto-u.ac.jp
Phone +81-96-342-3013

# Typical Functions of Bot Worm-infected PCs

- Transmitting of the unsolicited E-mails
- A distributed denial of service (DDoS) attack
- Self-Propagation or Launching the other internet worms
- Spying or disclosure a secret (Information Leakage)

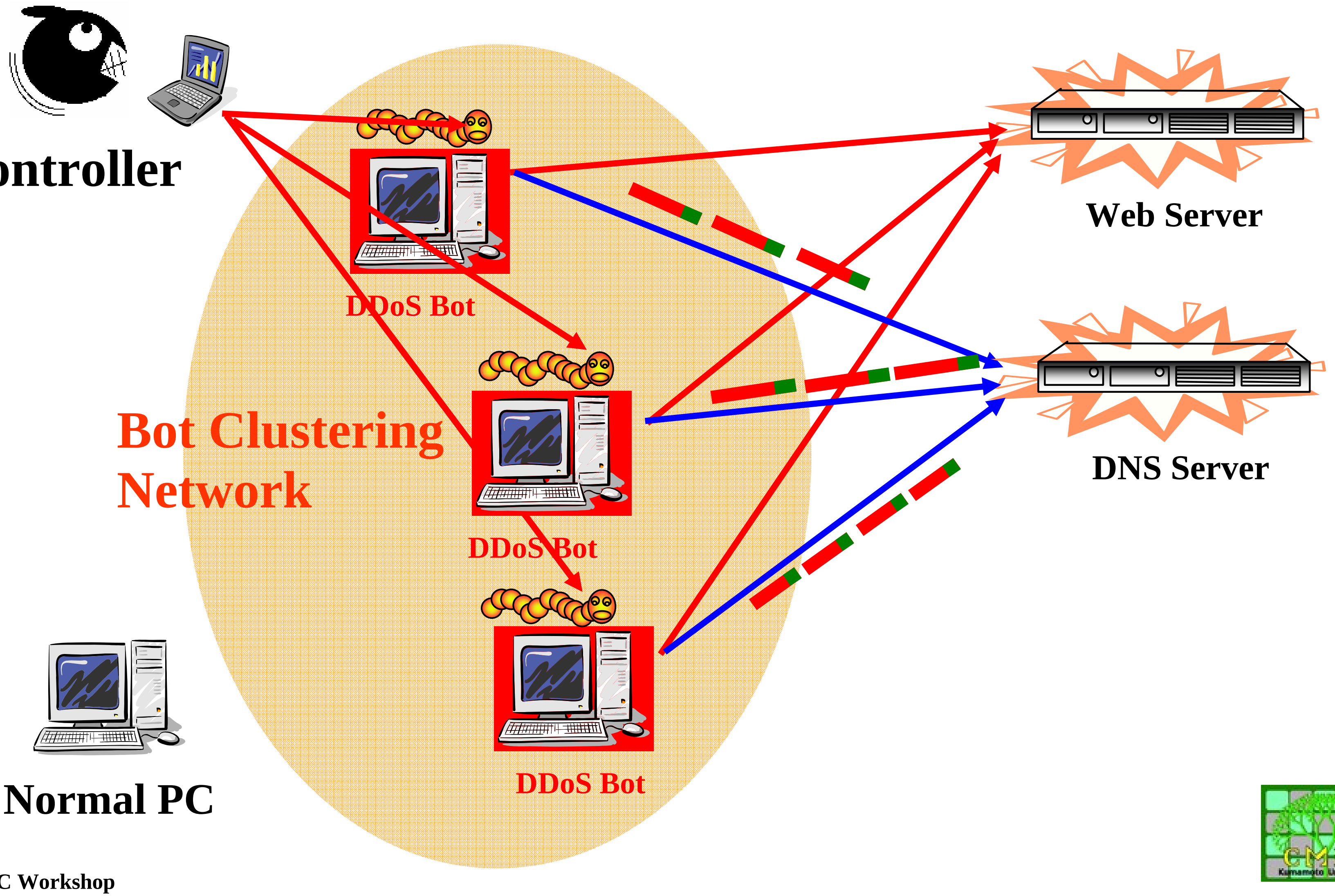
# A Spam Bot as an SMTP proxy



# Typical Functions of Bot Worm-infected PCs

- Transmitting of the unsolicited E-mails
- A distributed denial of service (DDoS) attack
- Self-Propagation or Launching the other internet worms
- Spying or disclosure a secret (Information Leakage)

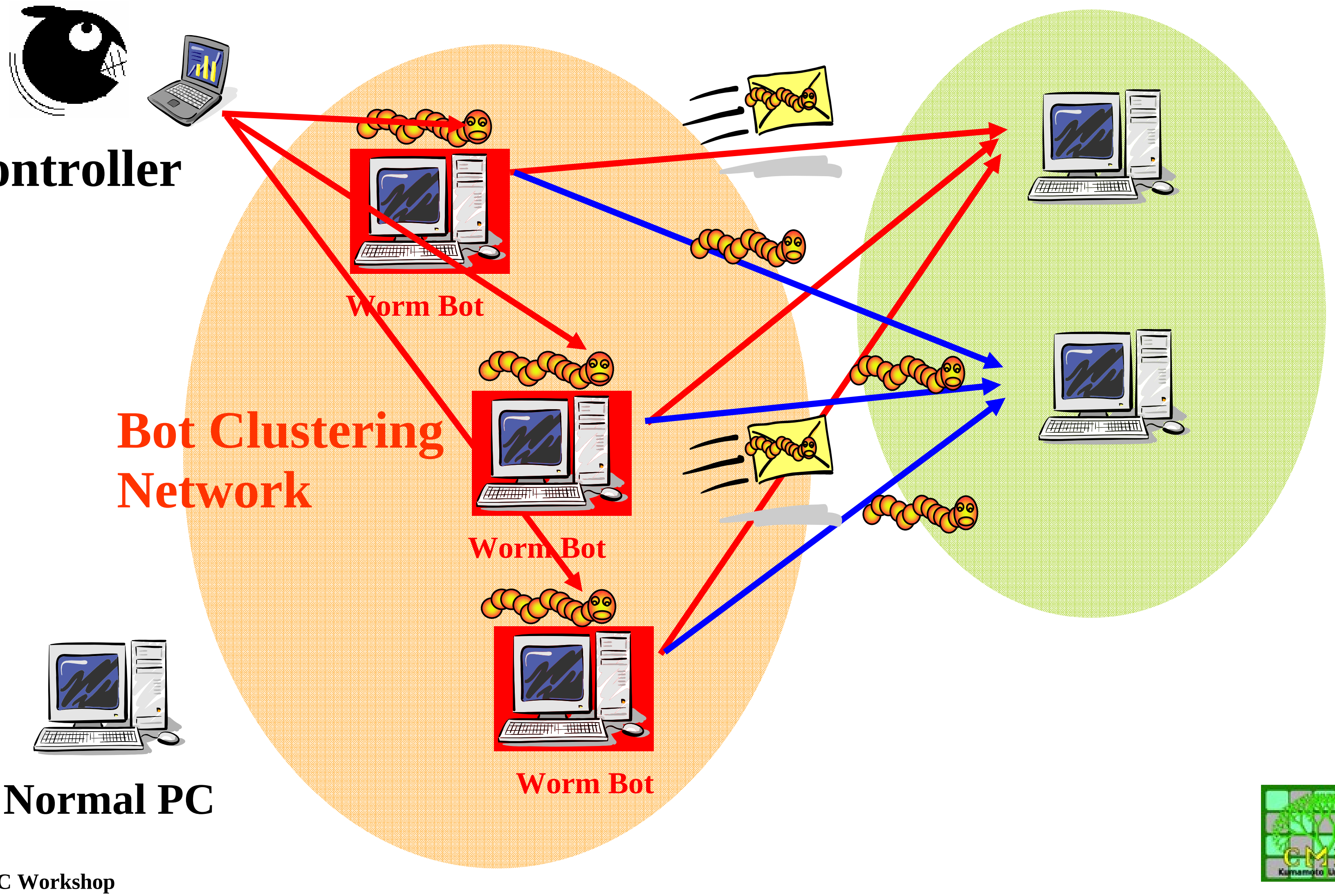
# A Distributed DoS (DDoS) Cyber Attack



# Typical Functions of Bot Worm-infected PCs

- Transmitting of the unsolicited E-mails
- A distributed denial of service (DDoS) attack
- Self-Propagation or Launching the other internet  worms
- Spying or disclosure a secret (Information Leakage)

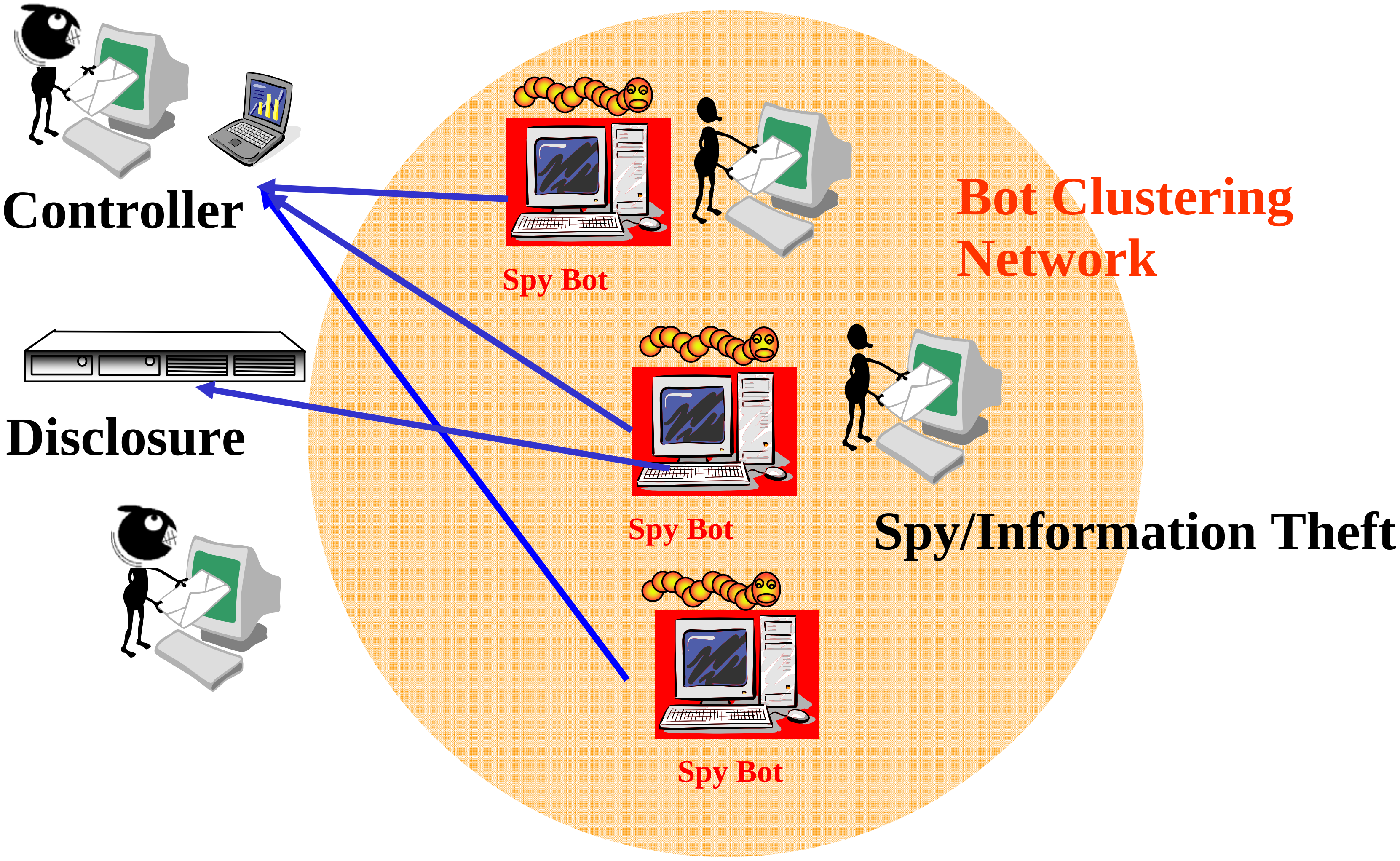
# Bot Propagation/Launching New Internet Worm



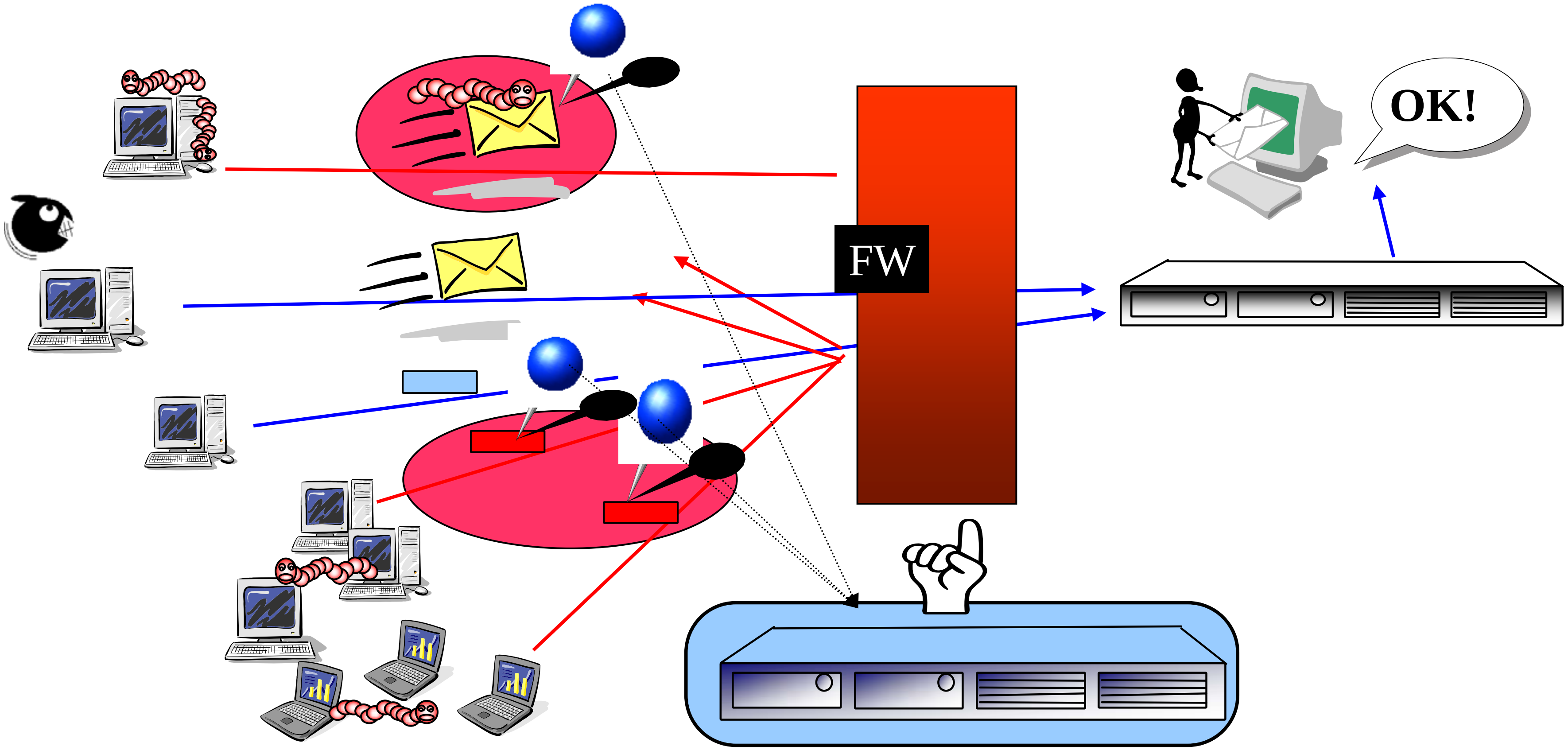
# Typical Functions of Bot Worm-infected PCs

- Transmitting of the unsolicited E-mails
- A distributed denial of service (DDoS) attack
- Self-Propagation or Launching the other internet worms
- Spying or disclosure a secret (Information Leakage)

# Information Leakage



# We need Countermeasures against the Bot Worm



# Countermeasures for Today

- Transmitting of the unsolicited E-mails
- A distributed denial of service (DDoS) attack
- Self-Propagation or Launching the other internet worms
- Spying or disclosure a secret (Information Leakage)

# Conventional Detection Technologies

- Direct Observation/Analysis of the traffic packets

For instance:

| | |
|---|---|
| E-mail exchange | SMTP packets |
| Web access | HTTP packets |

- Week points

(1) Ciphered/Encrypted Data is hardly to decode *i.e.* to hardly find out the security incidents in the encrypted data

(2) Privacy Disclosure

Direct observation of the network traffic always includes much privacy related information.

# Why do we observe DNS Query Packets? (Low Privacy)

- **A** resource record (RR) type: a fully qualified domain name (**FQDN**) into the IP **address(es)**
- **PTR** RR type: **an IP address** into the **FQDN**
- **MX** RR: a generic **domain name (DN)** into the **FQDN** of an E-mail server

http://www.cc.kumamoto-u.ac.jp/ 133.95.21.16

http://host.domainname/

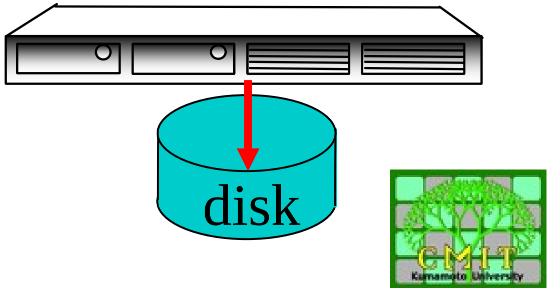
http://FQDN/ http://www.DN/

musashi@cc.kumamoto-u.ac.jp smtp://nyx.cc.kumamoto-u.ac.jp/

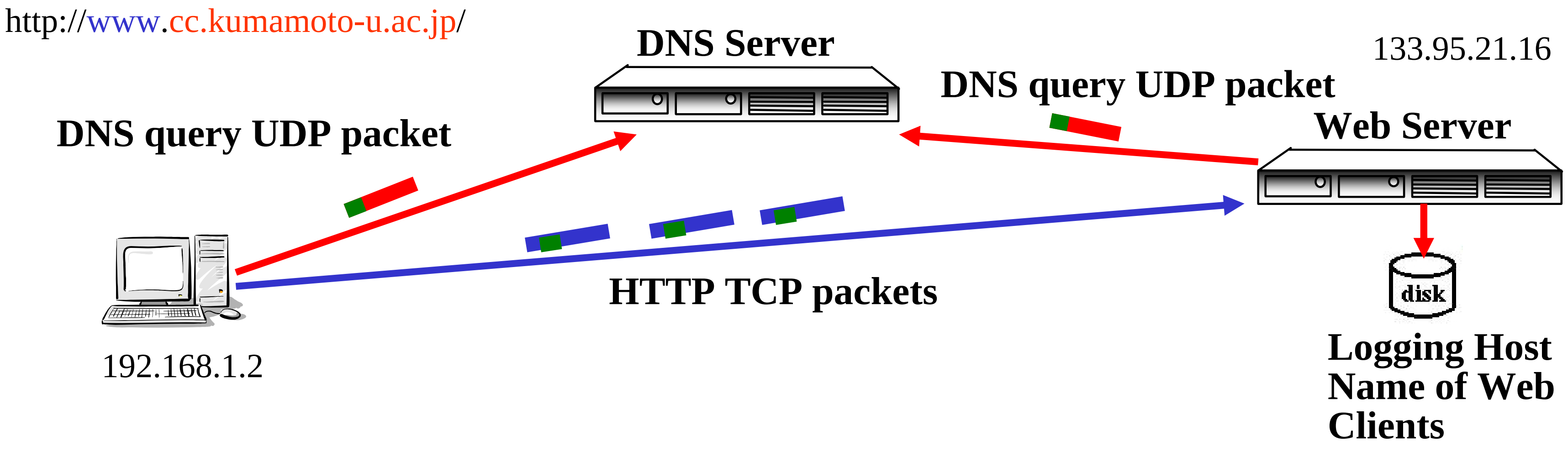
account@domainname smtp://host.domainname/

account@DN smtp://mail.DN/

```
Mar 17 23:45:22 cupid postfix/smtpd[10877]: connect from aaa.sub.kumamoto-u.ac.jp[133.95.x.y]
Mar 17 23:45:22 cupid postfix/smtpd[10877]: 1487B9D5: client=aaa.sub.kumamoto-u.ac.jp[133.95.x.y]
Mar 17 23:45:26 cupid postfix/cleanup[10879]: 1487B9D5: message-id=<2004031*****2.1487B9D5@¥
sub.kumamoto-u.ac.jp>
Mar 17 23:45:26 cupid postfix/smtpd[10877]: disconnect from aaa.sub.kumamoto-u.ac.jp[133.95.x.y]
Mar 17 23:45:26 cupid postfix/qmgr[627]: 1487B9D5: from=<foo@cupid.cc.kumamoto-u.ac.jp>, size=640,¥
 nrcpt=1 (queue active)
Mar 17 23:45:26 cupid postfix/smtp[10880]: 1487B9D5: to=<musashi@sub.kumamoto-u.ac.jp>,
relay=mail.sub.kumamoto-u.ac.jp[133.95.zzz.yyy], delay=4, status=sent (250 Ok: queued as ¥
D48F4C6D4A)
```



# Why do we observing DNS Query Packets



http://www.cc.kumamoto-u.ac.jp/

DNS Server

133.95.21.16

DNS query UDP packet

Web Server

DNS query UDP packet

HTTP TCP packets

192.168.1.2

Logging Host
Name of Web
Clients

192.168.1.2=pcx.aaa.jp

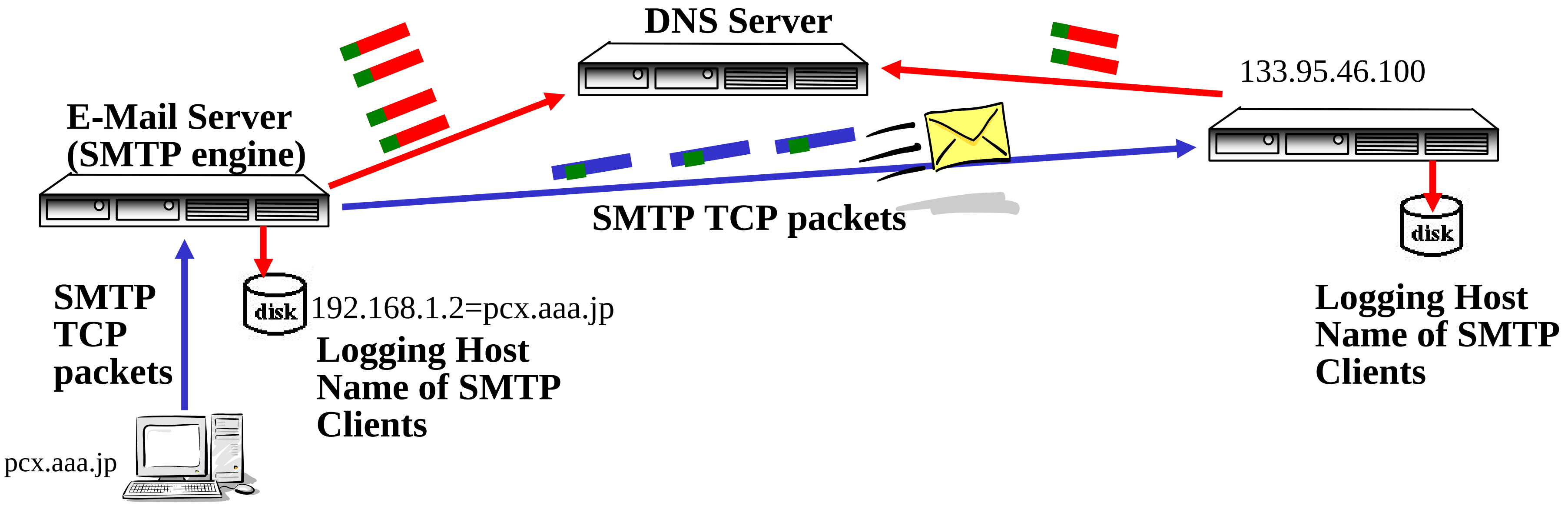
www.cc.kumamoto-u.ac.jp → 133.95.21.16

**The A (standard) resource record (RR) type DNS query packet transferred to the DNS server**

192.168.1.2 → pcx.aaa.jp

**The PTR (pointer)  RR type DNS query packet transferred to the DNS server**

# Why do we observing DNS Query Packets



DNS Server

133.95.46.100

E-Mail Server
(SMTP engine)

SMTP TCP packets

SMTP
TCP
packets

192.168.1.2=pcx.aaa.jp

Logging Host
Name of SMTP
Clients

Logging Host
Name of SMTP
Clients

pcx.aaa.jp

cc.kumamoto-u.ac.jp → nyx.cc.kumamoto-u.ac.jp

**The MX (Mail eXchange RR type DNS query packet transferred to the DNS server**

nyx.cc.kumamoto-u.ac.jp → 133.95.46.100

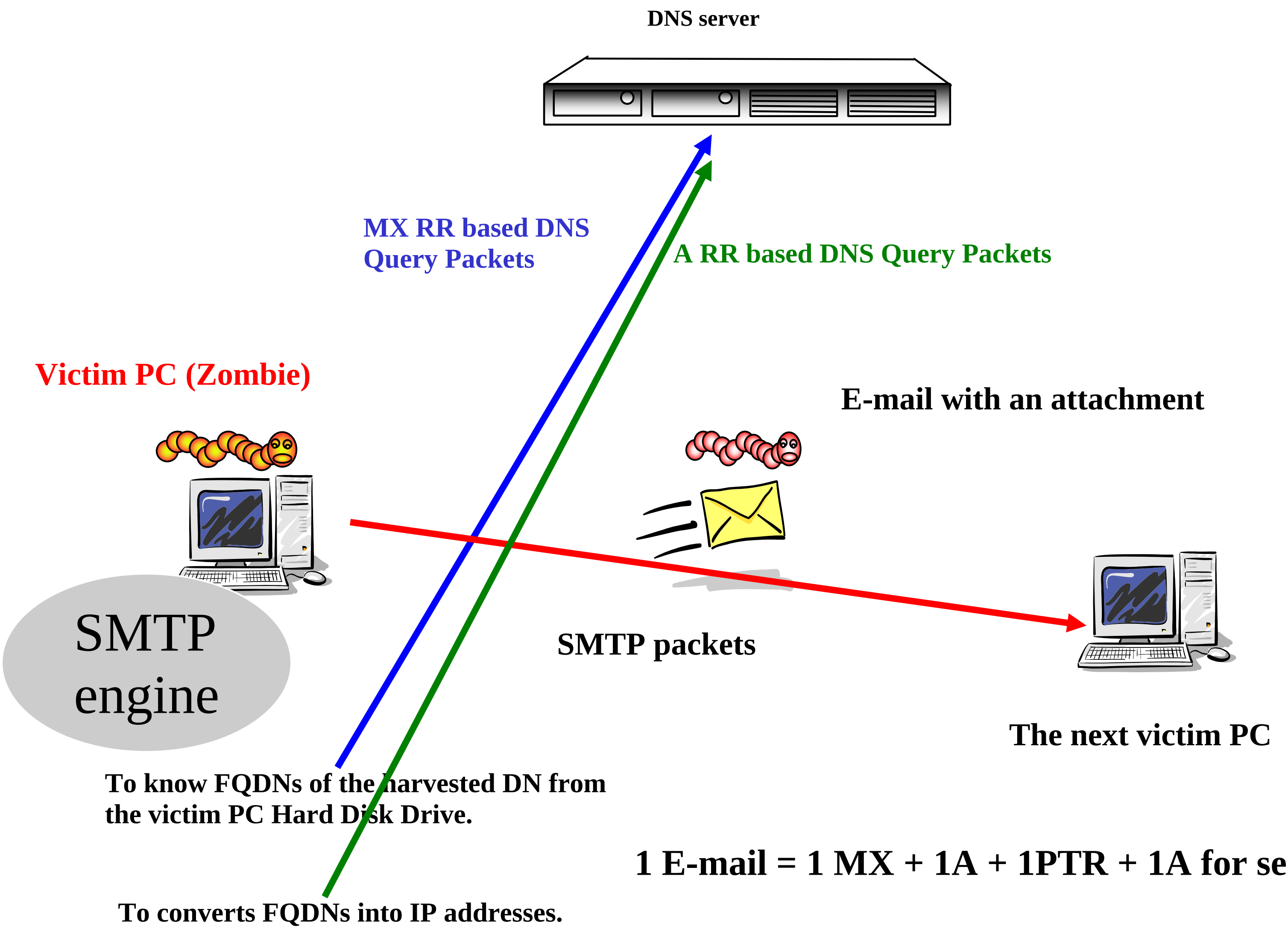
**The A (Standard) RR type DNS query packet transferred to the DNS server**

**1 SMTP = 1 MX + 1A**

**1 E-mail = 1 MX + 1A + 1PTR + 1A for seding**

**1 E-mail = 1PTR + 1A for receiving**

# At 29.03.2004, we reported that…



DNS server

MX RR based DNS Query Packets

A RR based DNS Query Packets

Victim PC (Zombie)

E-mail with an attachment

SMTP engine

SMTP packets

The next victim PC

To know FQDNs of the harvested DN from the victim PC Hard Disk Drive.

To converts FQDNs into IP addresses.

1 E-mail = 1 MX + 1A + 1PTR + 1A for seding

# DNS-based Detection of the Incidents and Related Works

- Musashi, Matsuba, Sugitani, IPSJ-CSEC19(2002)/CSEC20(2003)

  http://www.cc.kumamoto-u.ac.jp/~musashi/200{2,3}p.html
- Rikitake, Nogawa, Tanaka, and Shimojo, IPSJ-CSS2003
- Matsuba, Musashi, and Sugitani, IPSJ-DSM32(Japan) and ICETA2004 (Košice, Slovakia)

  http://www.cc.kumamoto-u.ac.jp/~musashi/2004p.html
- Kristoff, NANOG32, Reston, VA (2004) and Northwestern University

  http://www.nanog.org/mtg-0410/kistoff.html

  http://aharp.ittns.northwestern.edu/talks/bots-dns.pdf
- Whyte, van Oorschot, Kranakis, Carleton Univ., Technical Report

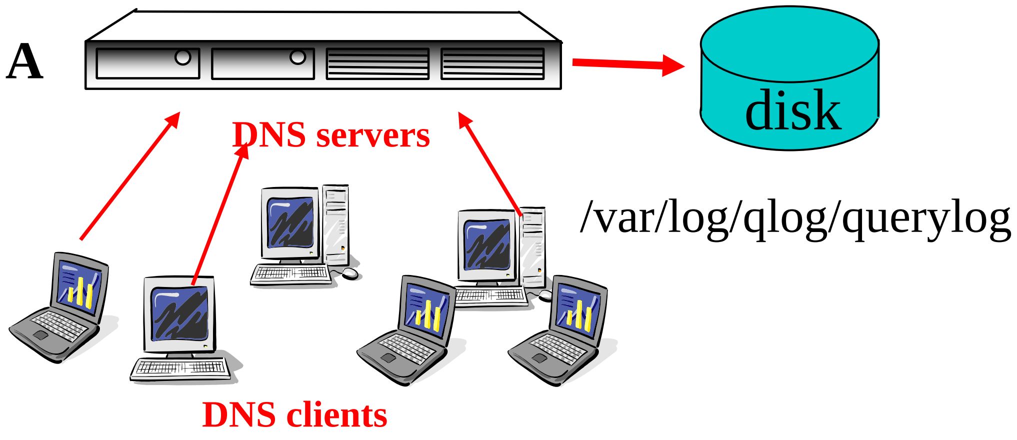
  http://www.scs.carleton.ca/research/tech_reports/2005/download/TR-05-06.pdf
- Ishibashi, Toyono, Toyama, Ishino, Ohshima, and Mizukoshi, ACM SIGCOMM workshop, 2005 http://www.acm.org/sigs/sigcomm/sigcomm2005/paper-IshToy.pdf
- Schonewille and van Helmond, University of Amsterdam, SURFnet, 2006

  http://staff.science.uva.nl/~delaat/snb-2005-2006/p12/report.pdf

# Log Analysis of the DNS Query Contents

Capturing of DNS query packet by the optional configuration of the DNS server daemon program like BIND: /etc/named.conf

Intel Xeon 2.4GHz Dual CPU, 1GB main memory,
Intel 100Mbps NIC, and80GB ATA 133 HDD



B

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

Optional Configuration of BIND-9.2.6

**C** Date h:m:s hostname named[PID]: client IP address#port: query: contents of DNS query packet and IN query type

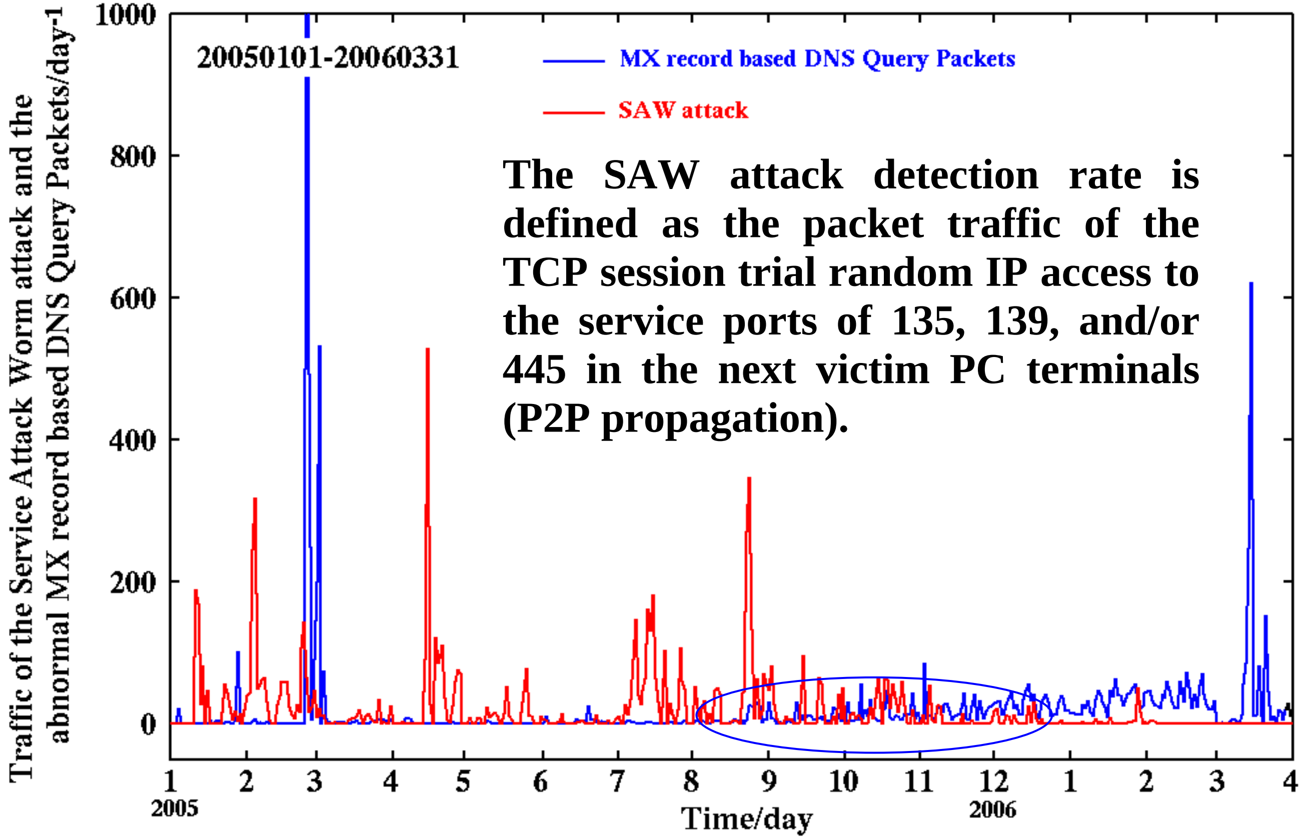
```
Oct 12 08:38:24 kun named[533]: client 133.95.xxx.yyy#39815: query: 130.13.194.xxx.in-addr.arpa IN PTR
Oct 12 08:38:25 kun named[533]: client 133.95.xxx.yyy#39825: query: dmea.net IN MX
Oct 12 08:38:43 kun named[533]: client 133.95.xxx.yyy#40010: query: mxwall03.hkabc.net IN A
```

The well-known three DNS query types are:

**A** resource record (RR) type: a fully qualified domain name (FQDN) into the IP address(es)
**PTR** RR type: an IP address into the FQDN
**MX** RR type: a generic domain name into the FQDN of an E-mail server

**D** The activities of the service attack worms are captured by the iplog-2.2.3 packet logger program package.

# Detection Rate of the clients-based MX RR Query Access and Service Attack Worm-infected PC terminals



The SAW attack detection rate is defined as the packet traffic of the TCP session trial random IP access to the service ports of 135, 139, and/or 445 in the next victim PC terminals (P2P propagation).

**The client-based MX RR DNS traffic synchronizes in almost the same manner with the detection rate of the SAW-infected PCs the late days of August to the middle of December, 2005. After the late days of 2005, it is, however, very difficult to find out the IP addresses of the BW-infected PC terminals by only watching P2P propagation or client MX RR based DNS query traffic (Detection Evasion).**

# Detection Strategies

**Statistical Analysis on:**

**(1) the source IP address (IPv4) based DNS query traffic from the bot worm (BW)-infected PC terminals in the campus network,**

**(2) the IPv6-source IP based DNS query traffic from the bot worm (BW)-infected PC terminals in the campus network, and**

**(3) the query contents based DNS query traffic from detection systems on the internet (the other sites) like IDS/IPS, spam filter, etc.**

**IDS/IPS=Intrusion Detection/Prevention System**

# Detection Strategies

**Statistical Analysis on:**

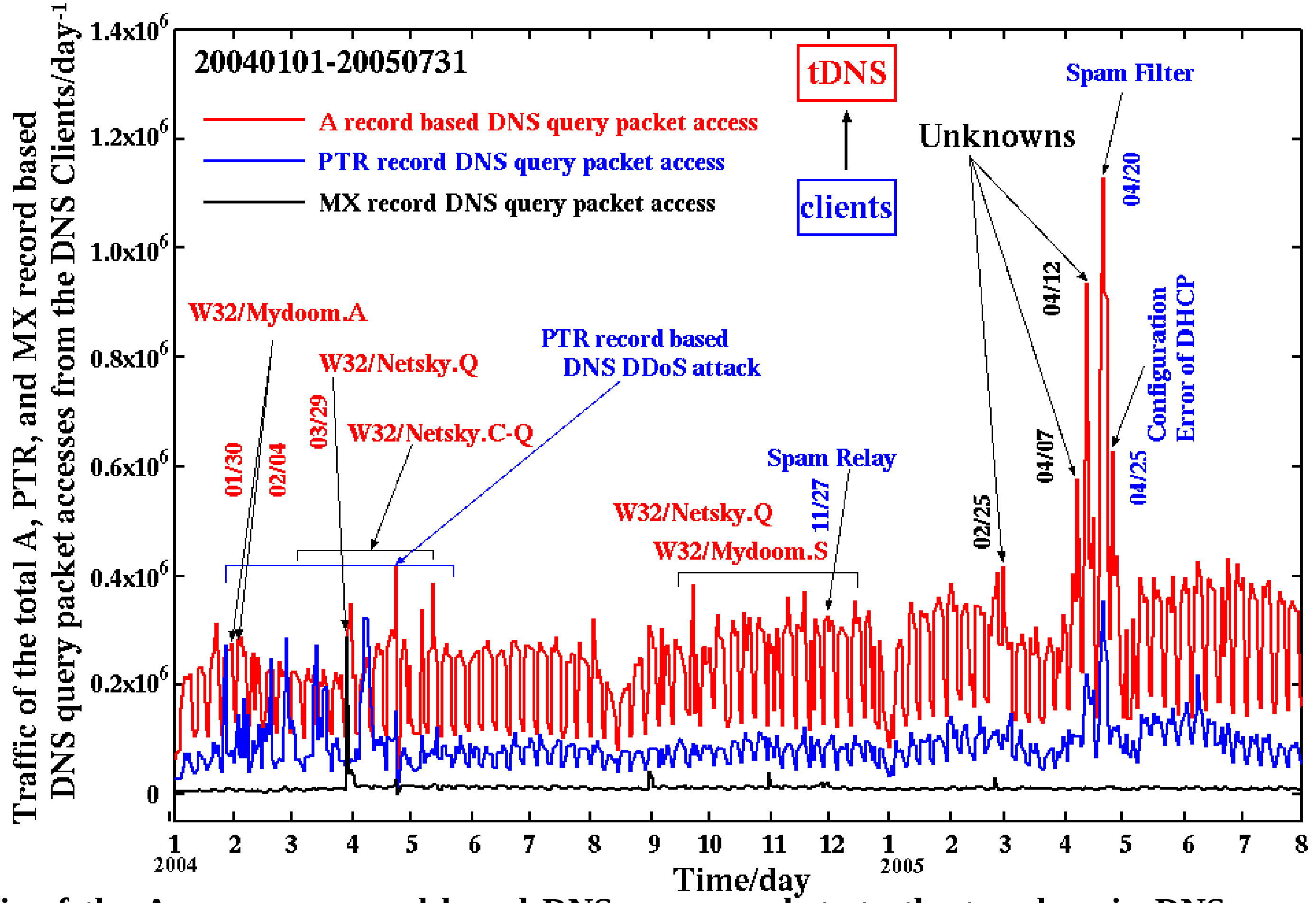
**(1) the source IP address (IPv4) based DNS query traffic from the bot worm (BW)-infected PC terminals in the campus network,**

**(2) the IPv6-source IP based DNS query traffic from the bot worm (BW)-infected PC terminals in the campus network, and**

**(3) the query contents based DNS query traffic from detection systems on the internet (the other sites) like IDS/IPS, spam filter, etc.**

**IDS/IPS=Intrusion Detection/Prevention System**

# DNS Query Traffic includes Worm Information



**Traffic of the A resource record based DNS query packets to the top domain DNS server of the university was abnormally increased through the early days of January to the middle days of June, 2005.**

**Unknowns: 25th February, 7th and 12th April, 2005**

# Example DNS query traffic from the BW-infected PCs

- The PC client A is a top access client in 25th February, 2005

| | |
|---|---|
| Tot: | 32,728/day |
| A: | 32,727/day |
| PTR: | 7/day |

- The PC clients B and C are a top access client in 7th and 12th April, respectively

| | Client B: | Client C |
|---|---|---|
| Tot: | 229,309/day | 400,964/day |
| A: | 229,265/day | 400,964/day |
| PTR: | 34/day | |
| MX: | 1/day | |
| SOA: | 8/day | |
| AAAA: | 1/day | |

# Example DNS query traffic from the BW-infected PCs

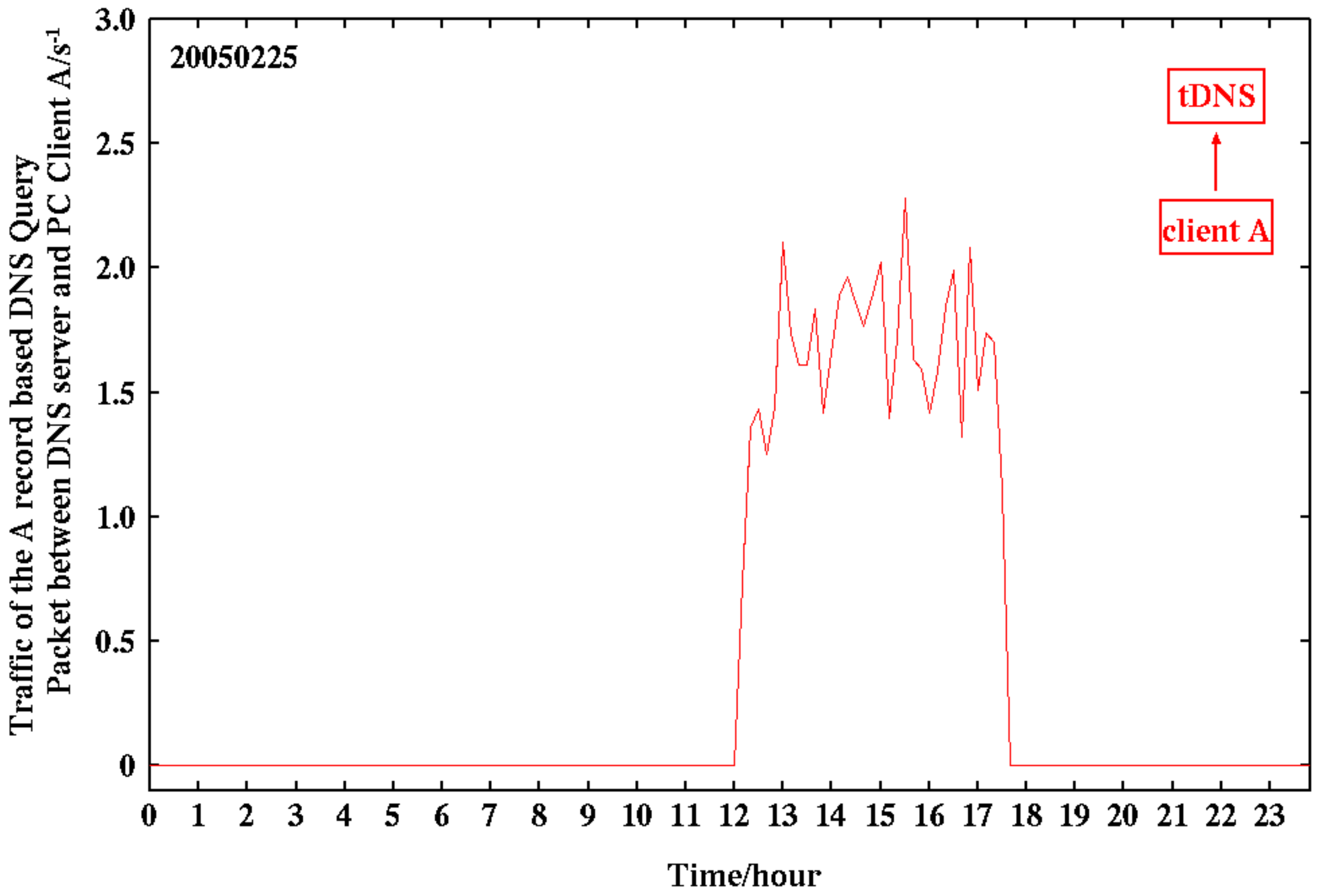
- The PC client A is a top access client in 25th February, 2005

| | |
|---|---|
| Tot: | 32,728/day |
| A: | 32,727/day |
| PTR: | 7/day |

- The PC clients B and C are a top access client in 7th and 12th April, respectively

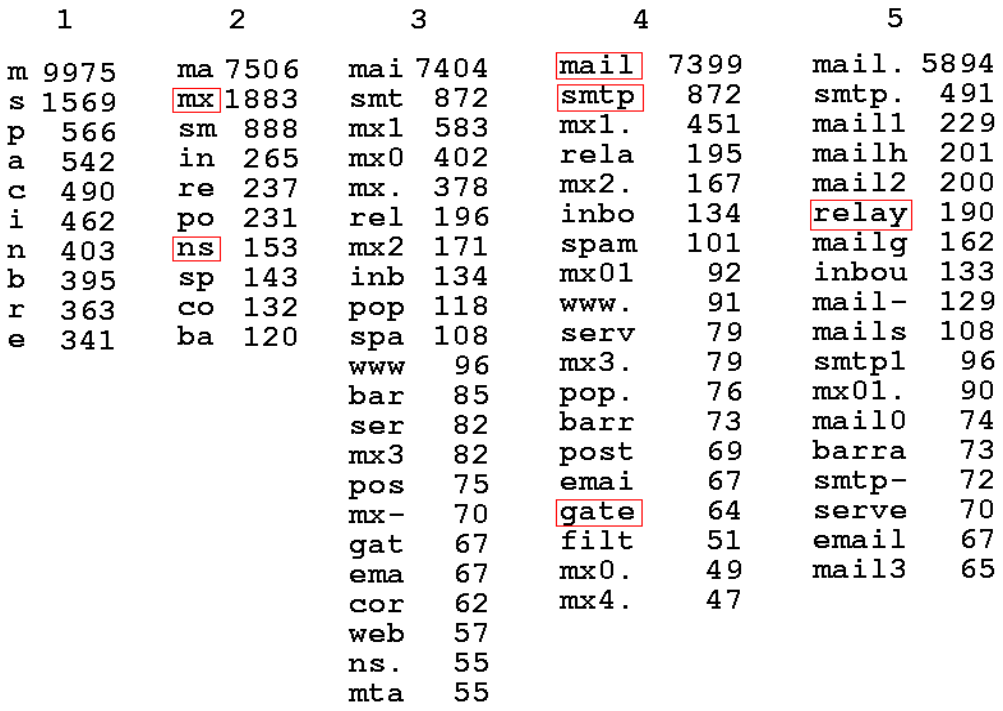
| | Client B: | Client C |
|---|---|---|
| Tot: | 229,309/day | 400,964/day |
| A: | 229,265/day | 400,964/day |
| PTR: | 34/day | |
| MX: | 1/day | |
| SOA: | 8/day | |
| AAAA: | 1/day | |

# Abnormal Traffic of the A RR based DNS Query Packets from the Client A



**It took place at February 25th, 2005 12:00-17:30 (Filtered manually).**

# Statistics of the DNS Query Contents in the A RR based DNS query Traffic



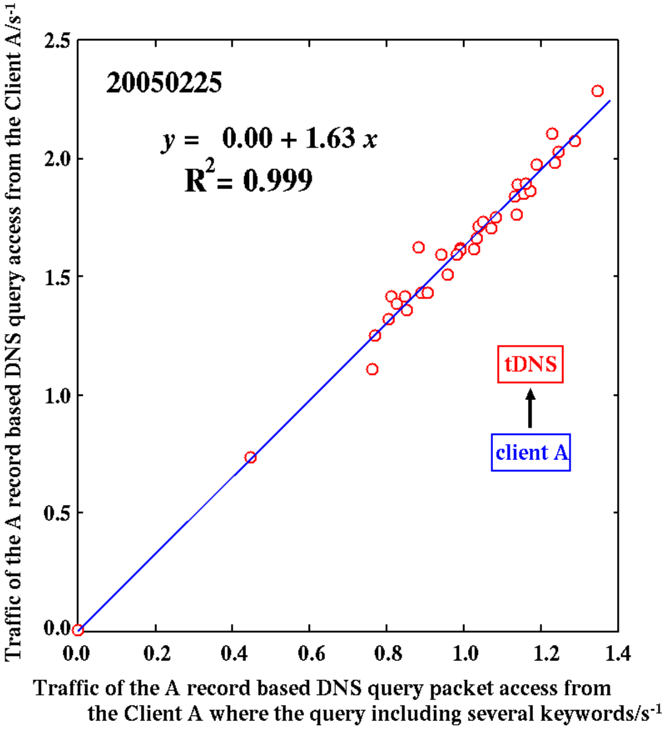
```
   1          2          3            4              5

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
```

**We can see several significant keywords like “mx”, “ns”, “mail”, “smtp”, “gate”, and “relay” in the head words of query contents.**

# Correlation between Total Traffic and Traffic including Several Keywords

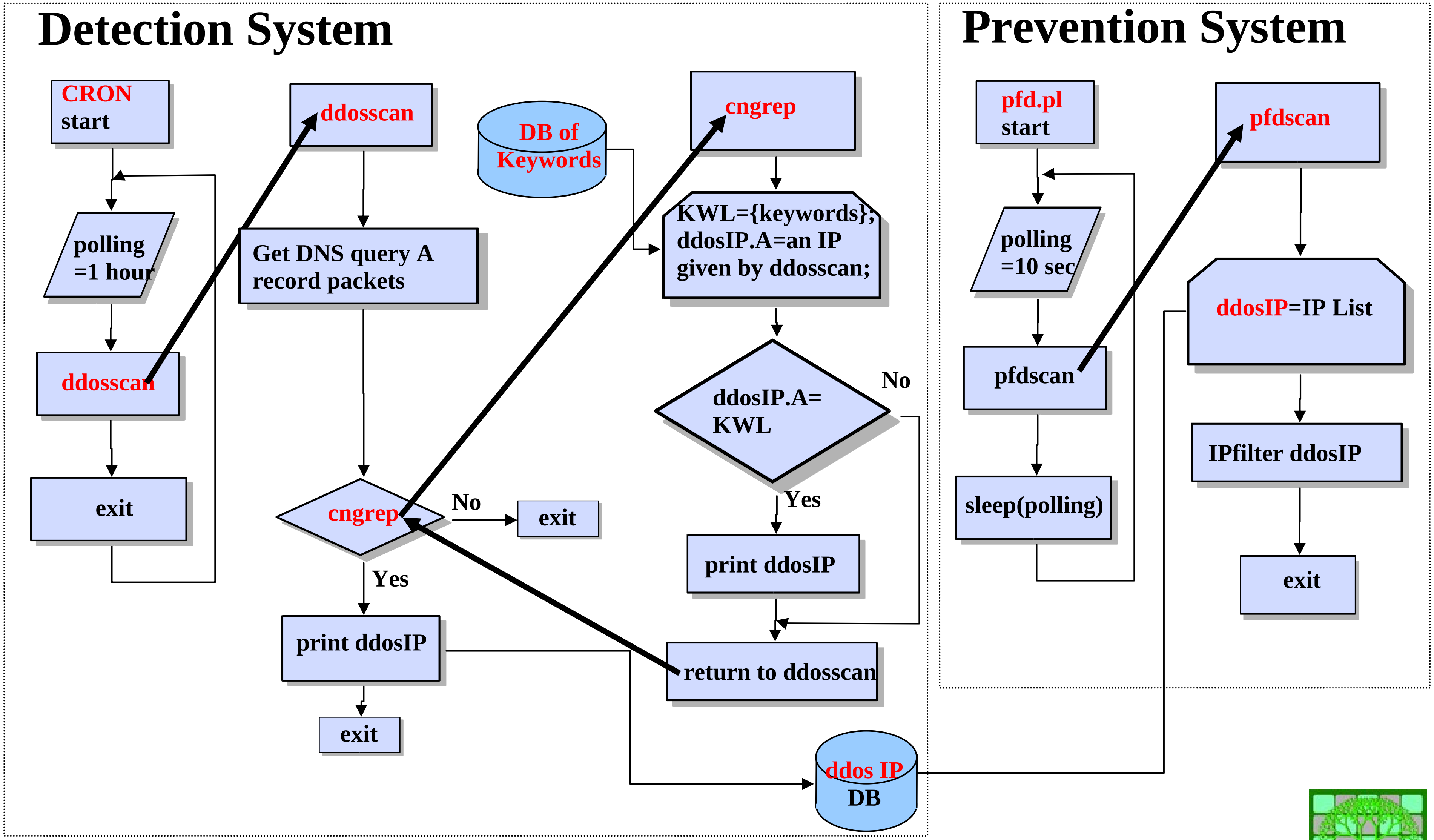
| 1 | | 2 | | 3 | | 4 | | 5 | |
|---|---|---|---|---|---|---|---|---|---|
| m | 9975 | ma | 7506 | mai | 7404 | mail | 7399 | mail. | 5894 |
| s | 1569 | mx | 1883 | smt | 872 | smtp | 872 | smtp. | 491 |
| p | 566 | sm | 888 | mx1 | 583 | mx1. | 451 | mail1 | 229 |
| a | 542 | in | 265 | mx0 | 402 | 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 | | | | |



**(1) This strong correlation is useful to detect the abnormal traffic of the A RR based DNS query packets (IP addresses of BW- or MMW-infected PC terminals ?).**

**(2) The client A is an Windows PC terminal and we cannot find out any MX-record based DNS query packets from the PC terminal.**

# Detection- and Prevetion-System of Abnormal Traffic of the A record based DNS Query Packets from non-MX type BW-infected PC bots



## Detection System

CRON start

polling =1 hour

ddosscan

exit

ddosscan

Get DNS query A record packets

cngrep

No

exit

Yes

print ddosIP

exit

DB of Keywords

cngrep

KWL={keywords}; ddosIP.A=an IP given by ddosscan;

ddosIP.A= KWL

No

Yes

print ddosIP

return to ddosscan

ddos IP DB

## Prevention System

pfd.pl start

polling =10 sec

pfdscan

sleep(polling)

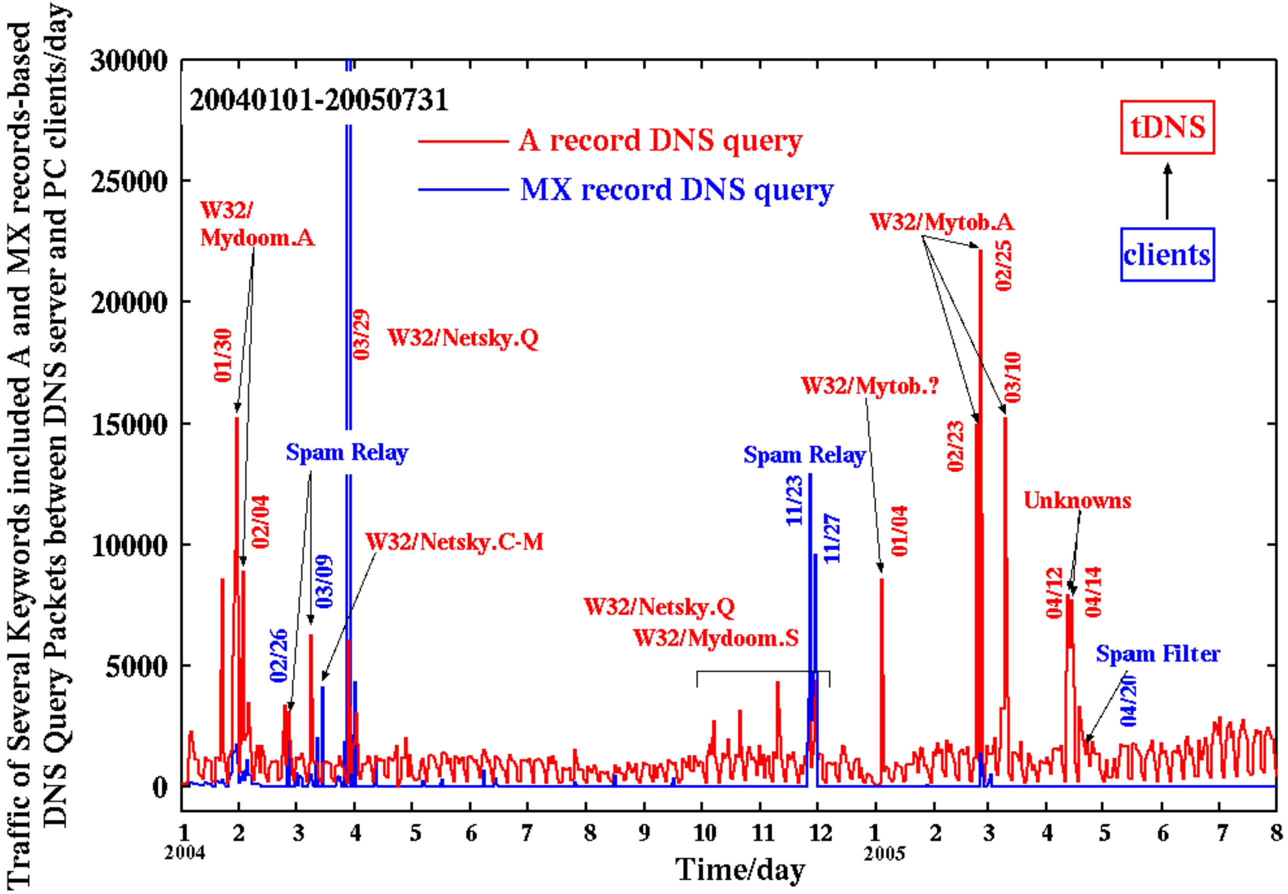
pfdscan

ddosIP=IP List

IPfilter ddosIP

exit

# Evaluation of the Detection and Prevention System: ADPS for non-MX type Mass Mailing Worm-infected PCs



**Mytob.A (non client MX query type MMW or spam bot) were found but the peaks at April 7th and 12th are disappeared or decreased.**

# Example DNS query traffic from the BW-infected PCs

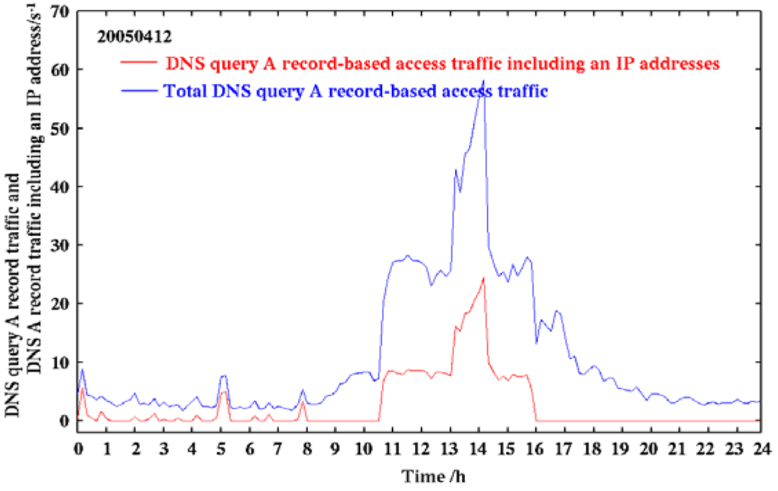
- **The PC client A is a top access client in 25th February, 2005**

| | |
|---|---|
| **Tot:** | **32,728/day** |
| **A:** | **32,727/day** |
| **PTR:** | **7/day** |

- **The PC clients B and C are a top access client in 7th and 12th April, 2005, respectively**

| | **Client B:** | **Client C** |
|---|---|---|
| **Tot:** | **229,309/day** | **400,964/day** |
| **A:** | **229,265/day** | **400,964/day** |
| **PTR:** | **34/day** | |
| **MX:** | **1/day** | |
| **SOA:** | **8/day** | |
| **AAAA:** | **1/day** | |

# Detection of Unusual Traffic of the A RR based DNS Query Traffic



client B

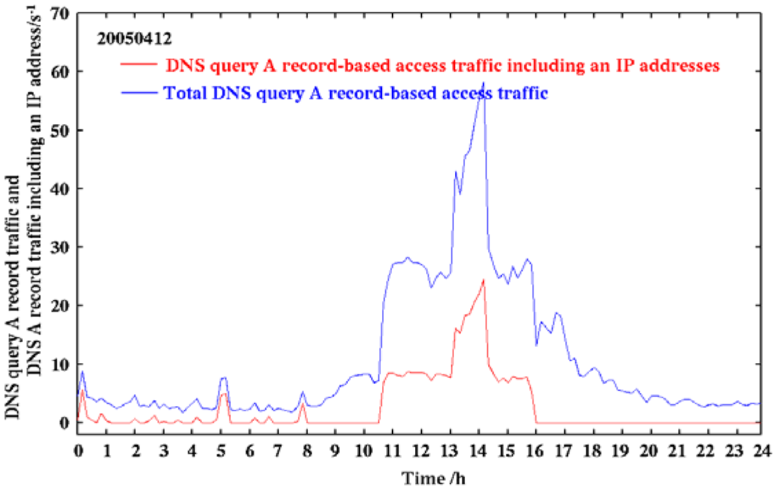
| | |
|---|---|
| 0.0.0.0 | 26 |
| ***.*****-u.ac.jp | 13 |
| 133.9*.**.192 | 11 |
| 133.9*.**.73 | 10 |
| 133.9*.**.66 | 9 |
| 133.9*.**.64 | 9 |
| 133.9*.**.52 | 9 |
| 133.9*.**.89 | 6 |
| mil.***.********-u.ac.jp | 5 |
| ***.**.********-u.ac.jp | 5 |
| 2**.*.2**.*8 | 5 |
| 133.9*.**.9 | 5 |
| 133.9*.**.8 | 5 |
| 133.95.**.7 | 5 |

client C

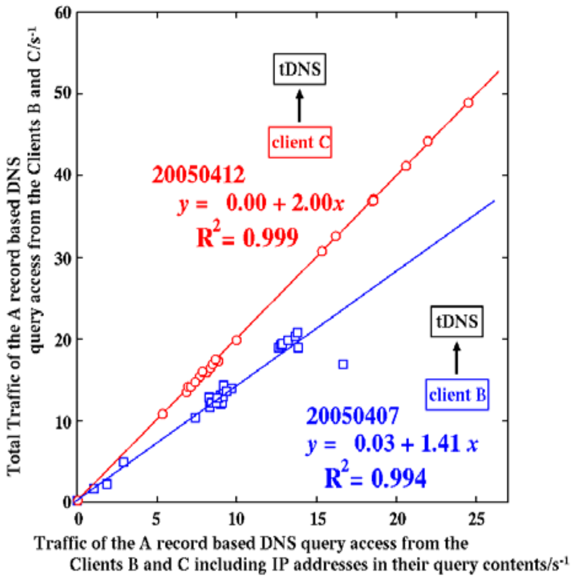
| | |
|---|---|
| ***.***.y****.com | 12 |
| www.*****m.com | 7 |
| yahoo.co.jp | 6 |
| www.****.****.co.jp | 6 |
| mail.****.com | 6 |
| img.****.co.jp | 5 |
| i.****.jp | 5 |
| ai.****.jp | 5 |
| 133.9*.***.194 | 5 |
| 133.9*.20*.2** | 5 |
| 127.0.0.1.***-u.ac.jp | 5 |
| 127.0.0.1 | 5 |
| relay.****.net | 4 |
| rd.*****.co.jp | 4 |

**The query contents of the DNS query access packets in the former four peaks, the IP address is directly included. Normally, only FQDN should be included in the contents of the A record based DNS query packets, howerver, the DNS query packets of the former peaks have IP addresses themselves as their contents. This feature is useful for detection of abnormal traffic of the A RR based DNS query packets.**

# Detection of Unusual Traffic of the A RR based DNS Query Traffic

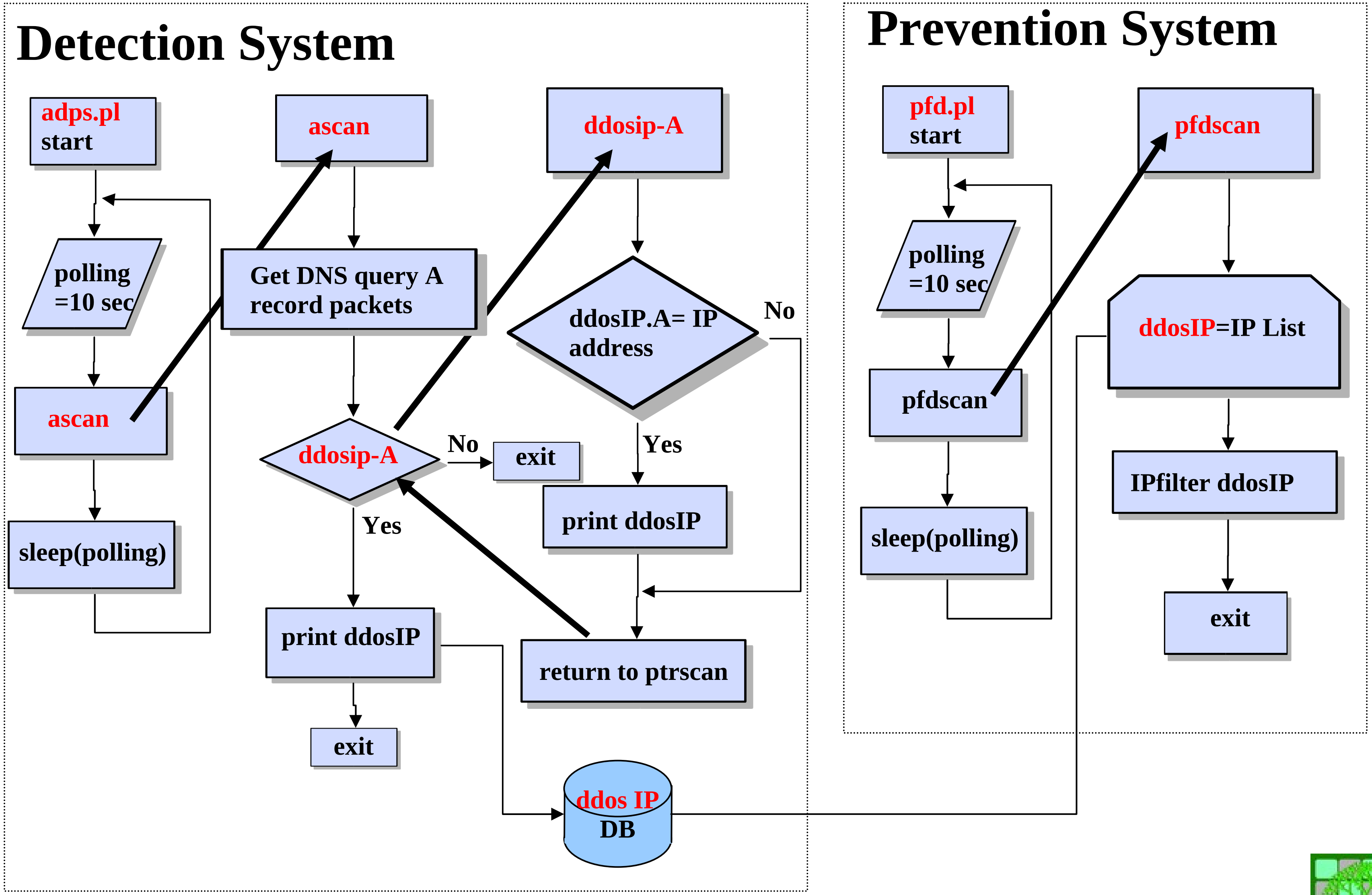


| client B | | client C | |
|---|---|---|---|
| 0.0.0.0 | 26 | ***.***.y****.com | 12 |
| ***.*****-u.ac.jp | 13 | www.*****m.com | 7 |
| 133.9*.**.192 | 11 | yahoo.co.jp | 6 |
| 133.9*.**.73 | 10 | www.****.****.co.jp | 6 |
| 133.9*.**.66 | 9 | mail.****.com | 6 |
| 133.9*.**.64 | 9 | img.****.co.jp | 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 |

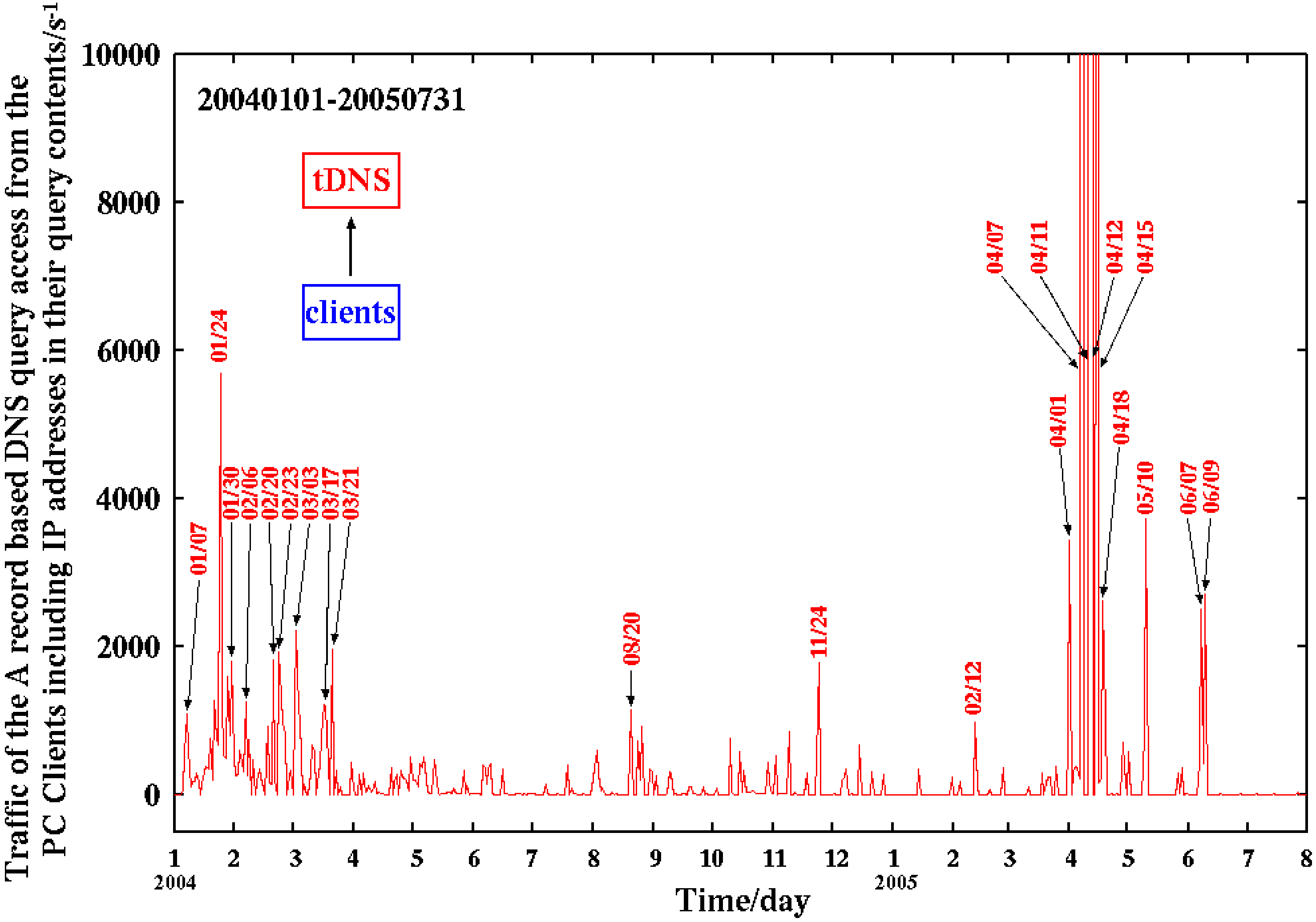


The query contents of the DNS query access packets in the former four peaks, the IP address is directly included. Normally, only FQDN should be included in the contents of the A record based DNS query packets, howerver, the DNS query packets of the former peaks have IP addresses themselves as their contents. This feature is useful for detection of abnormal traffic of the A record based DNS query packets.

# Detection- and Prevetion-System of Abnormal Traffic of the A RR based DNS Query Packets: ADPS for Direct IP



# Evaluation of the Detection and Prevention System: ADPS for Direct IP address included A RR based DNS query packets



# Detection Strategies

**Statistical Analysis on:**

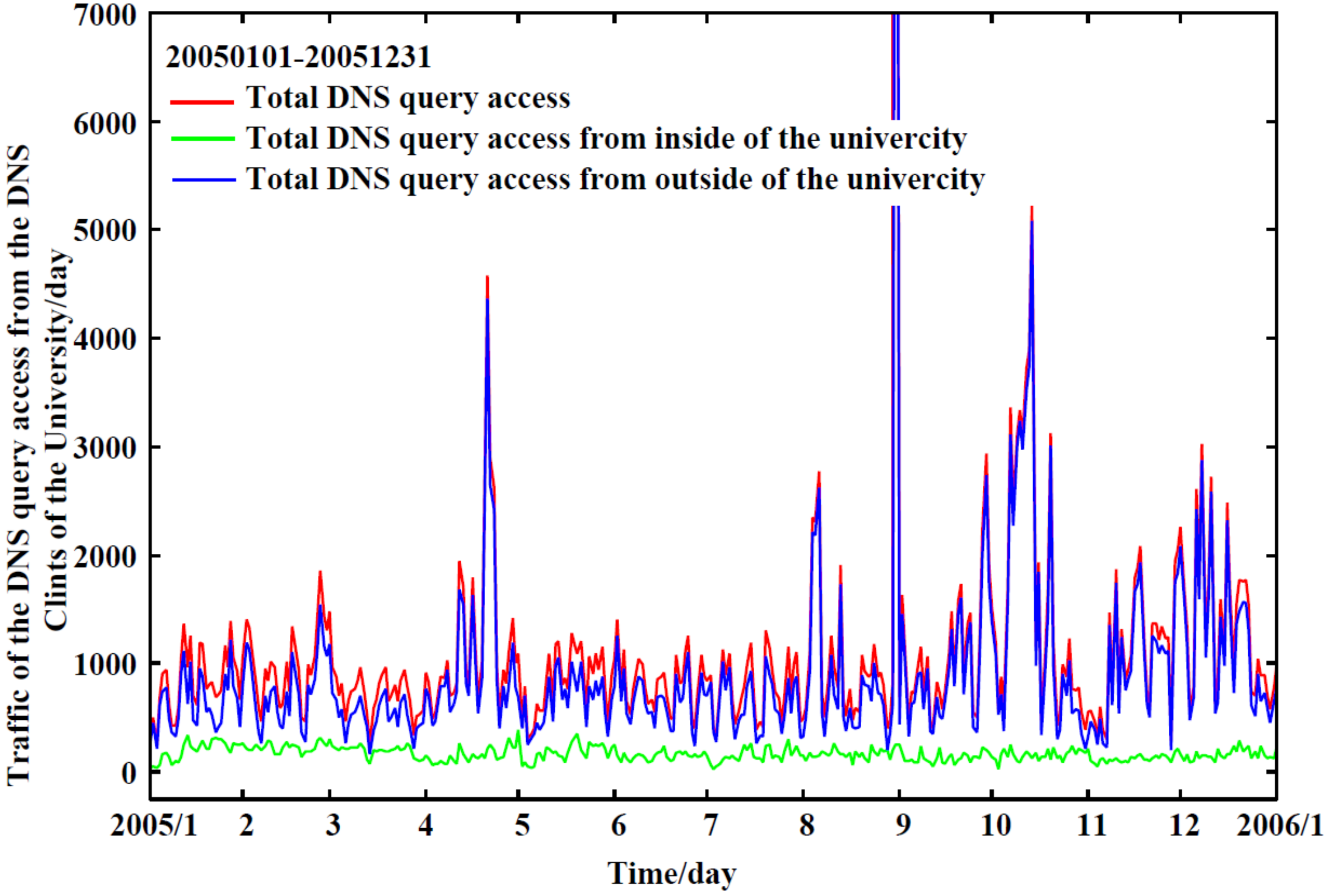
**(1) the source IP address based DNS query traffic from the bot worm (BW)-infected PC terminals in the campus network,**

**(2) the IPv6-source IP based DNS query traffic from the bot worm (BW)-infected PC terminals in the campus network, and**

**(3) the query contents based DNS query traffic from detection systems on the internet (the other sites) like IDS/IPS, spam filter, etc.**

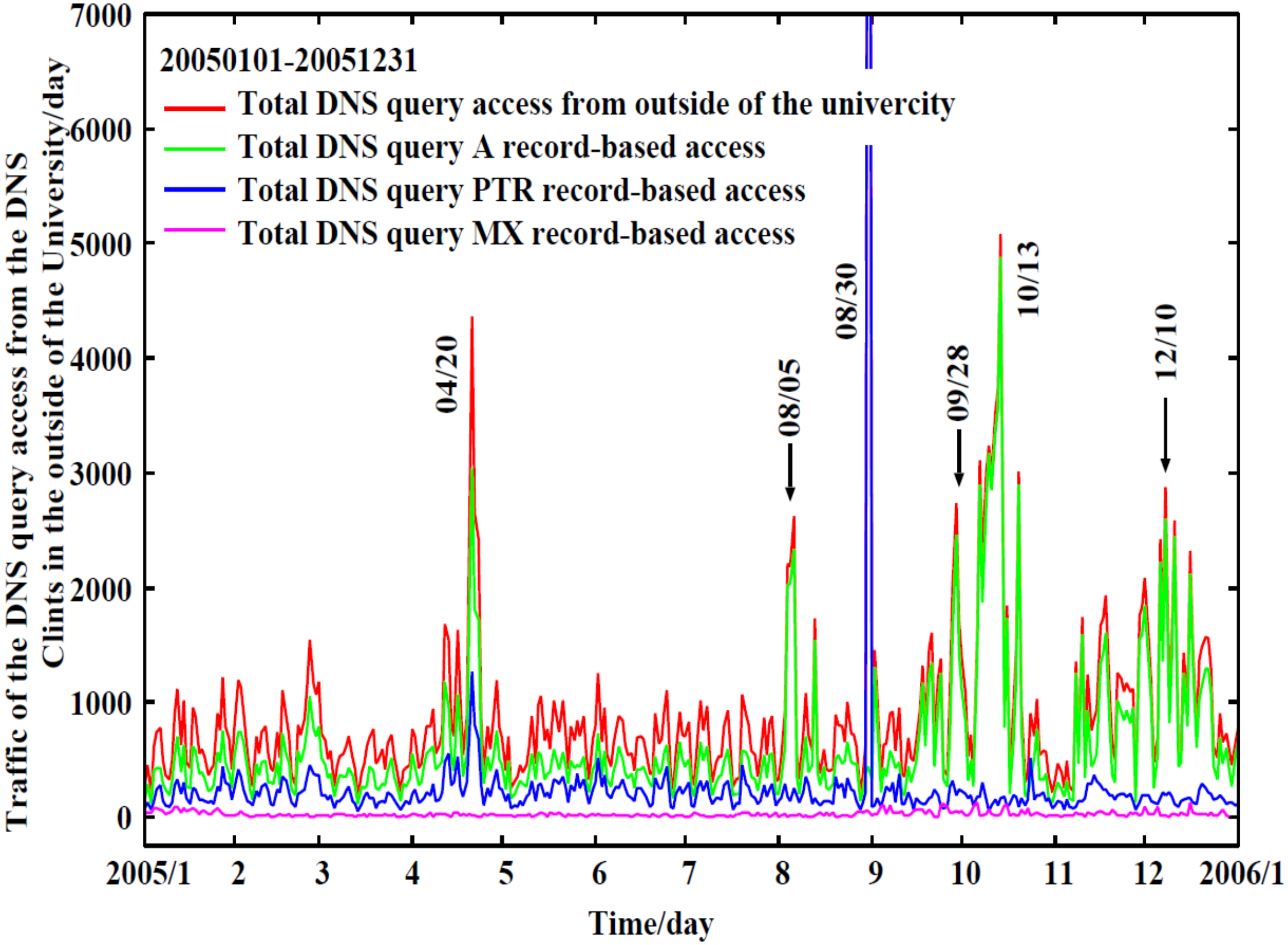
**IDS/IPS=Intrusion Detection/Prevention System**

# Total IPv6 based DNS query traffic



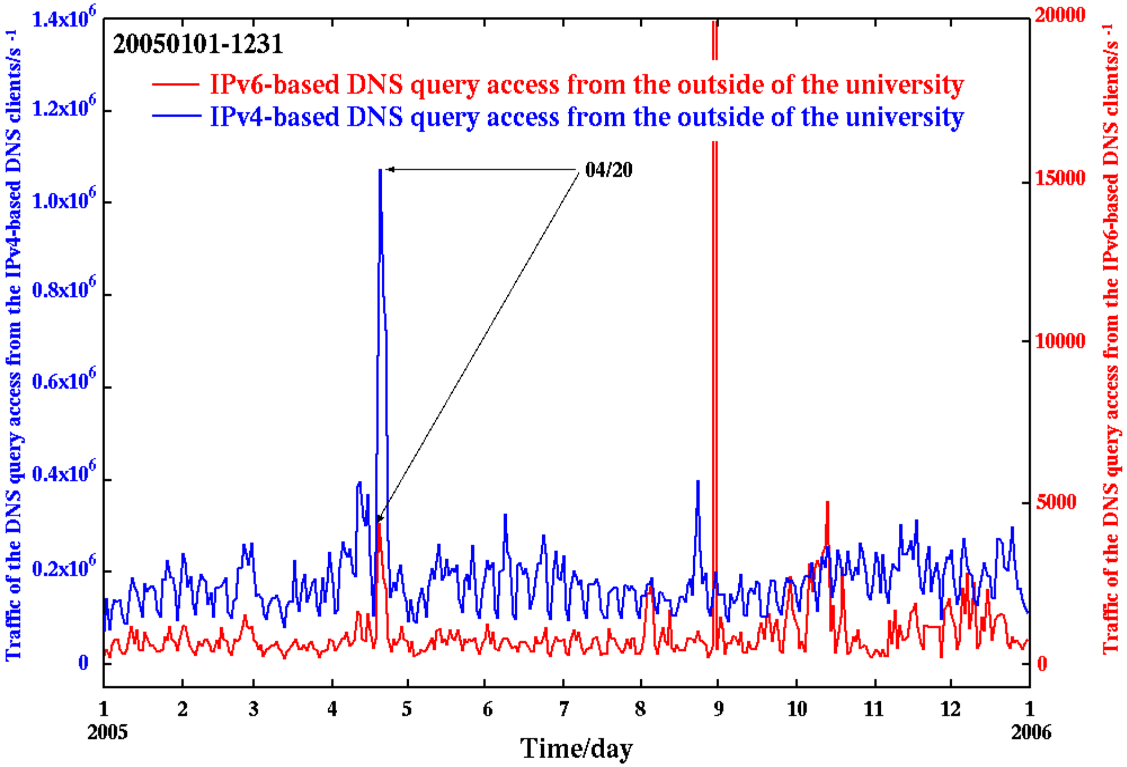
The total DNS query traffic from the IPv6-based DNS clients is mainly driven by that from the outside of the campus network.

# A, PTR, and MX RRs based DNS Query Traffic (IPv6)



**Several interesting peaks can be found: (i) April 20th, (ii) August 5th, (iii) August 30th, (iv) September 28th, (v) October 13th, and (vi) December 10th, 2005.**

# Abnormal A and PTR RRs based DNS Query Traffic



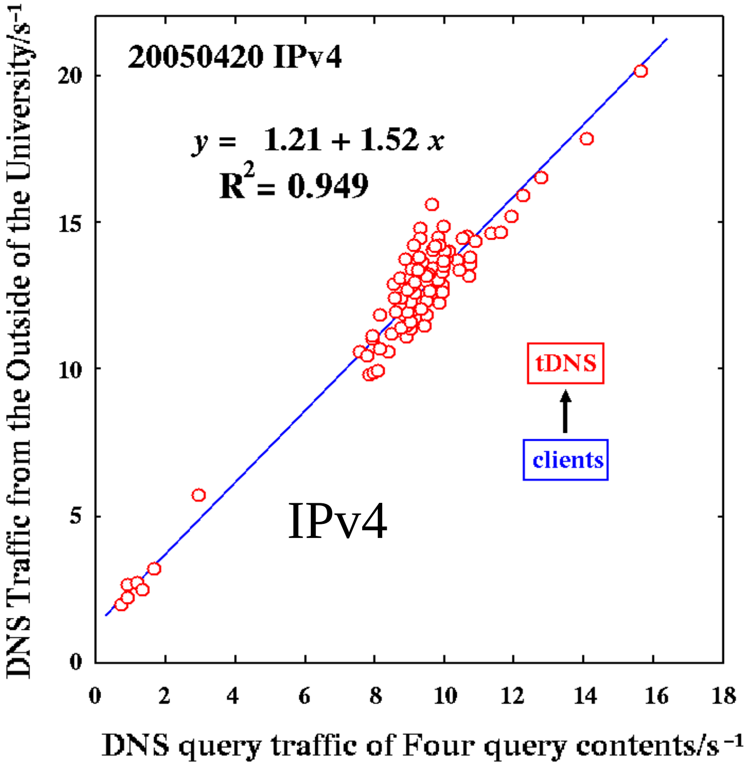
**In April 20th, 2005, both IPv6 and IPv4 based DNS query traffics strike two peaks simultaneously.**

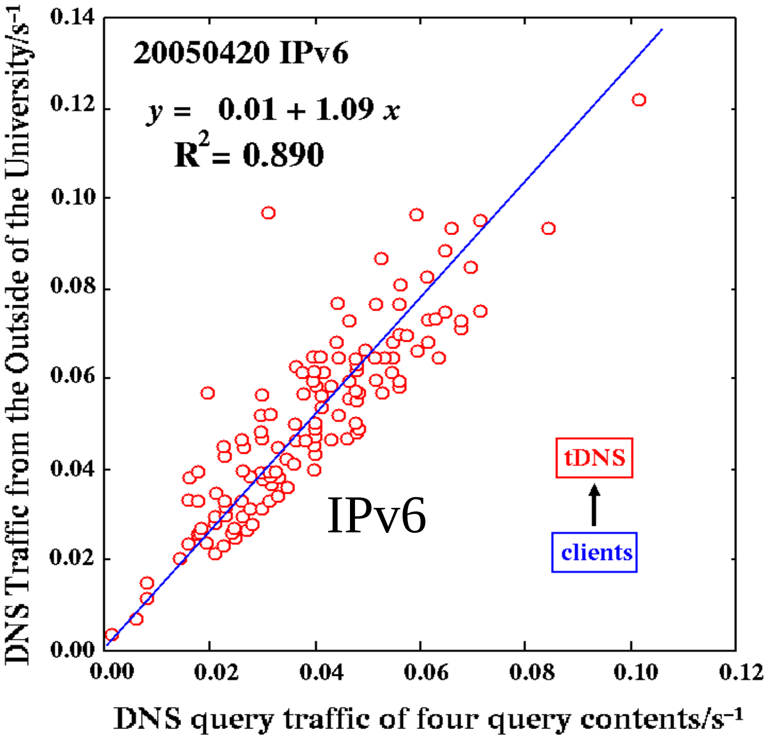
# Abnormal A and PTR RRs based DNS Query Traffic

| DNS query contents | IPv4 | IPv6 |
|---|---|---|
| *********.**.kumamoto-u.ac.jp | 230,729 | 1,345 |
| 133.95.***.** | 216,798 | 265 |
| ***.kumamoto-u.ac.jp | 180,298 | 999 |
| 133.95.***.** | 152,548 | 377 |

**In the query contents of the DNS query packets in the peak at April 20th, 2005, the most largest number of contents mainly consist of an FQDN of a local domain E-mail server, an FQDN of top domain DNS server (tDNS), and two IP addresses that related with PC terminals in the local domain, respectively. Since the E-mail server was pointed out as a spam-sender through the day of 20th April, 2005, the top DNS server are severely accessed by the spam-mail detection system/spam filter world-widely at the day.**

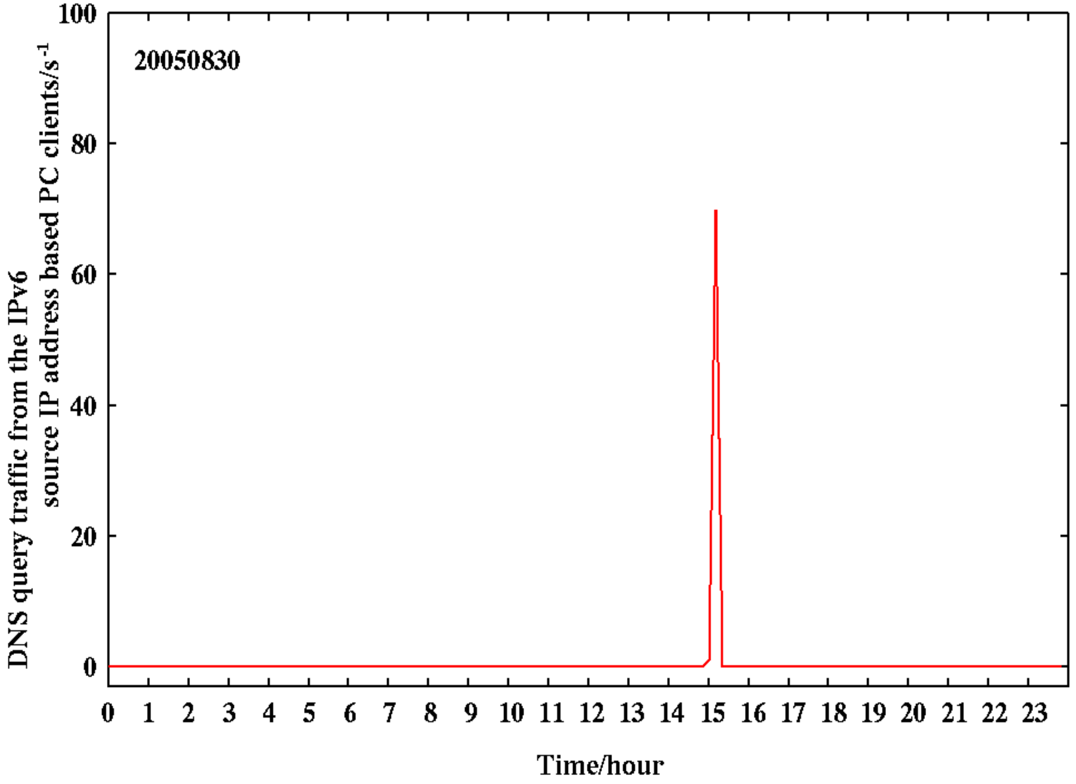
# Abnormal A and PTR RRs based DNS Query Traffic





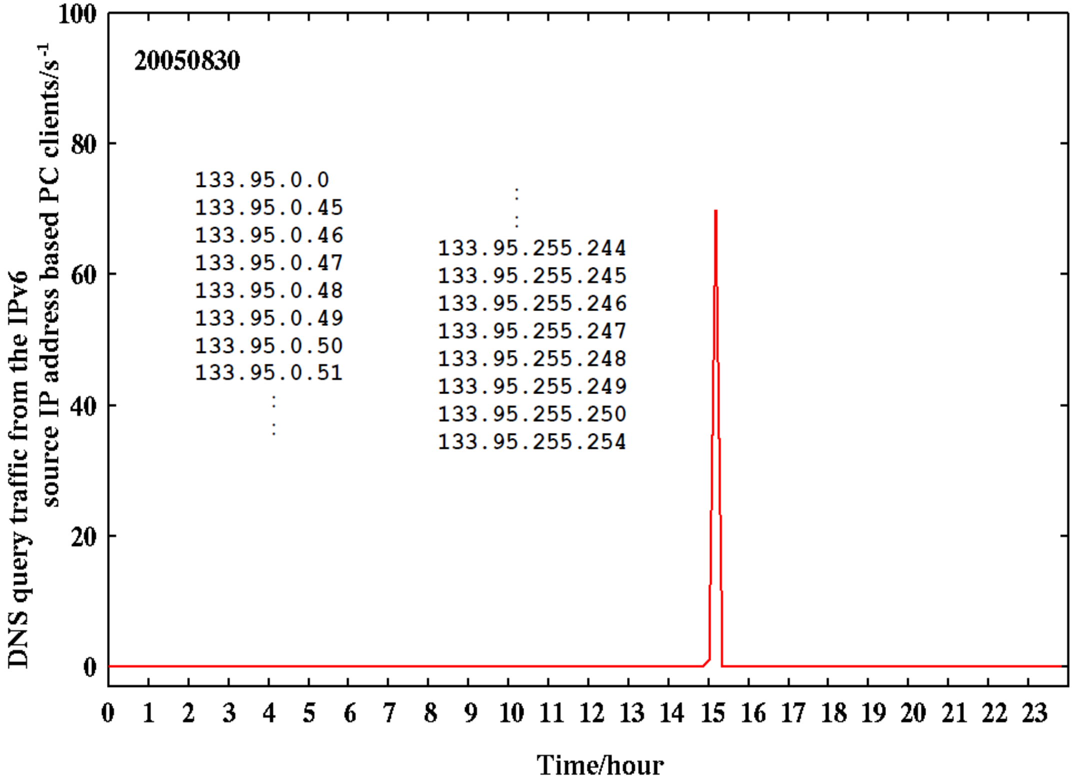
**The DNS query traffic from the outside the campus network correlates well with the IPv4- and IPv6 based DNS query traffics including four keywords.**

# Abnormal PTR RR based DNS Query Traffic



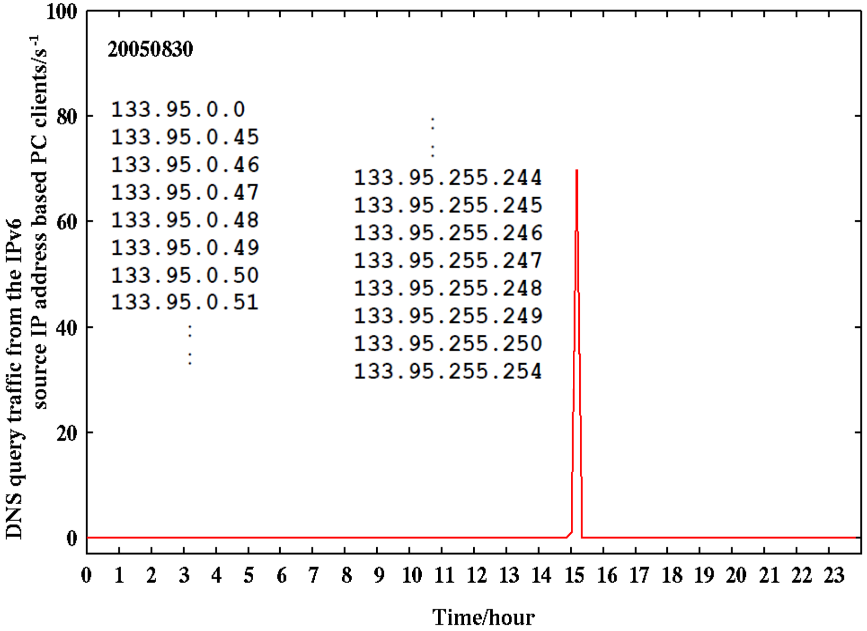
**The abnormal DNS query traffic is observed in the short period of time through 15:09-15:19 at August 30th, 2005. In the traffic, the two top DNS clients are found and they belong to the same site. The traffic mainly consists of the PTR record based DNS query packets including internal IP addresses of the university. Probably, the DNS clients tried to scan the hosts in the university.**

# Abnormal PTR RR based DNS Query Traffic



**The abnormal DNS query traffic is observed in the short period of time through 15:09-15:19 at August 30th, 2005. In the traffic, the two top DNS clients are found and they belong to the same site. The traffic mainly consists of the PTR record based DNS query packets including internal IP addresses of the university. Probably, the DNS clients tried to scan the hosts in the university.**

# Abnormal PTR RR based DNS Query Traffic



| DNS client IP address | Top access clients |
|---|---|
| 2001:1***:10**::2 | 22,001 |
| 2001:1***:10**::4 | 20,538 |
| 2001:2f8:14:**::64 | 229 |
| 3ffe:8200:0:10:250:****:fe00:**** | 135 |
| 3ffe:8200:0:10:250:****:fe00:**** | 67 |

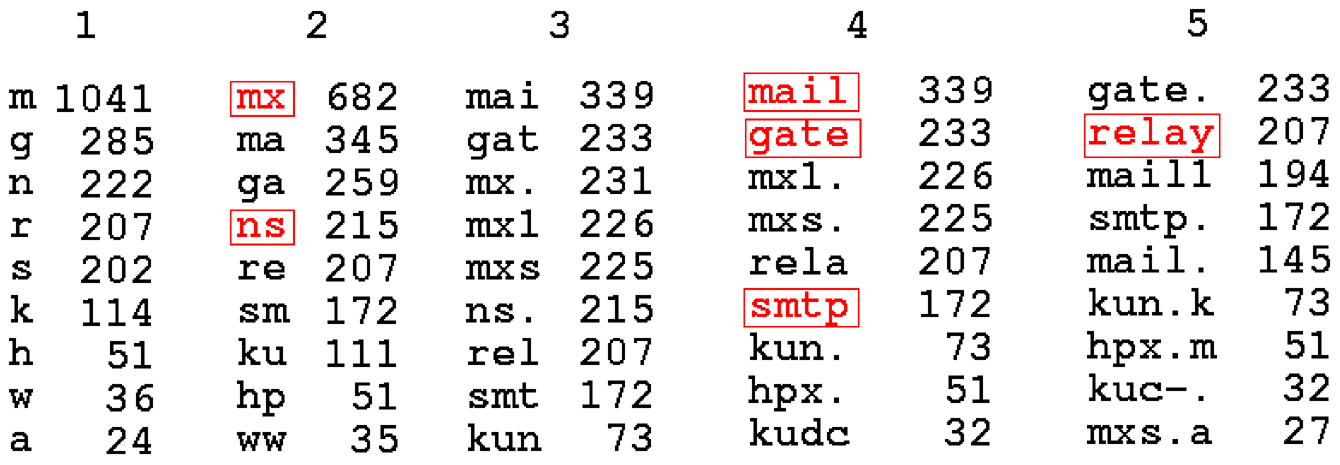
**The abnormal DNS query traffic is observed in the short period of time through 15:09-15:19 at August 30th, 2005. In the traffic, the two top DNS clients are found and they belong to the same site. The traffic mainly consists of the PTR record based DNS query packets including internal IP addresses of the university. Probably, the DNS clients tried to scan the hosts in the university.**

# Statistics of the source IP address based Abnormal PTR RR type DNS Query Traffic

| DNS client IP address | Top access clients |
|---:|---:|
| 2001:1***:10**::2 | 22,001 |
| 2001:1***:10**::4 | 20,538 |
| 2001:2f8:14:**::64 | 229 |
| 3ffe:8200:0:10:250:****:fe00:**** | 135 |
| 3ffe:8200:0:10:250:****:fe00:**** | 67 |

**The abnormal DNS query traffic is observed in the short period of time through 15:09-15:19 at August 30th, 2005. In the traffic, the two top DNS clients are found and they belong to the same site. The traffic mainly consists of the PTR record based DNS query packets including internal IP addresses of the university. Probably, the DNS clients tried to scan the hosts in the university.**

# Statistics for the DNS query contents of the Abnormal A RR based DNS Query Traffic

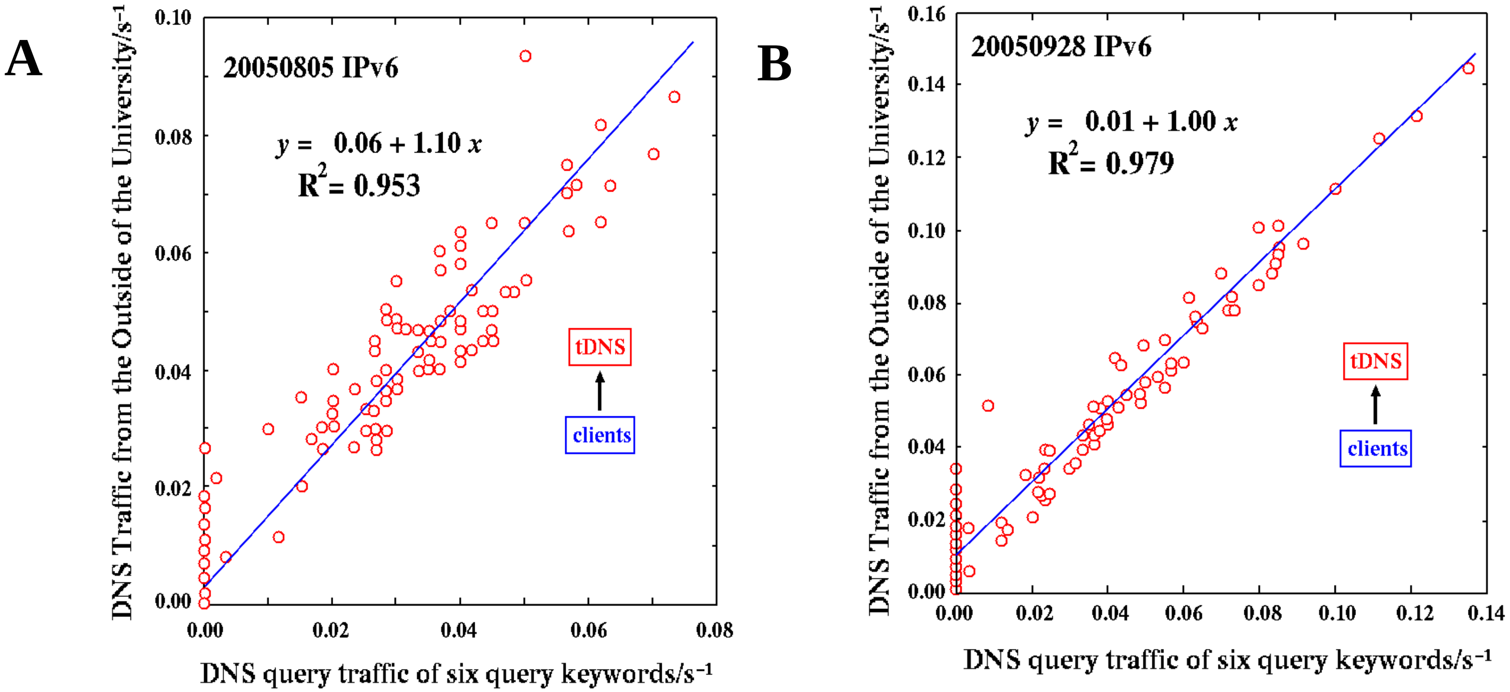


| 1 | | 2 | | 3 | | 4 | | 5 | |
|---|---|---|---|---|---|---|---|---|---|
| m | 1041 | mx | 682 | mai | 339 | mail | 339 | gate. | 233 |
| g | 285 | ma | 345 | gat | 233 | gate | 233 | relay | 207 |
| n | 222 | ga | 259 | mx. | 231 | mx1. | 226 | mail1 | 194 |
| r | 207 | ns | 215 | mx1 | 226 | mxs. | 225 | smtp. | 172 |
| s | 202 | re | 207 | mxs | 225 | rela | 207 | mail. | 145 |
| k | 114 | sm | 172 | ns. | 215 | smtp | 172 | kun.k | 73 |
| h | 51 | ku | 111 | rel | 207 | kun. | 73 | hpx.m | 51 |
| w | 36 | hp | 51 | smt | 172 | hpx. | 51 | kuc-. | 32 |
| a | 24 | ww | 35 | kun | 73 | kudc | 32 | mxs.a | 27 |

**In August 5th, 2005, we can observe that the A RR based DNS query traffic includes several typical keywords as in their query contents *i.e.* "mx", "ns", "mail", "gate", "smtp", and "smtp" that were included in the A RR based DNS query traffic from the bot worm (BW) like a W32/Mytob.A BW.**

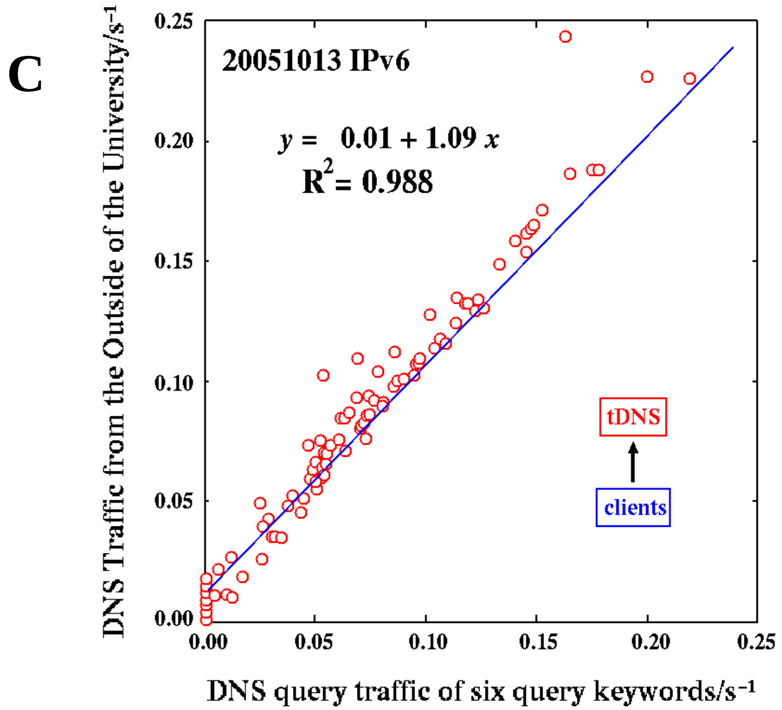
**Musashi, Y., etal., *IPSJ SIG Technical Reports, DSM38* , Vol. 2005, No. 83, pp.23-28 (2005).**

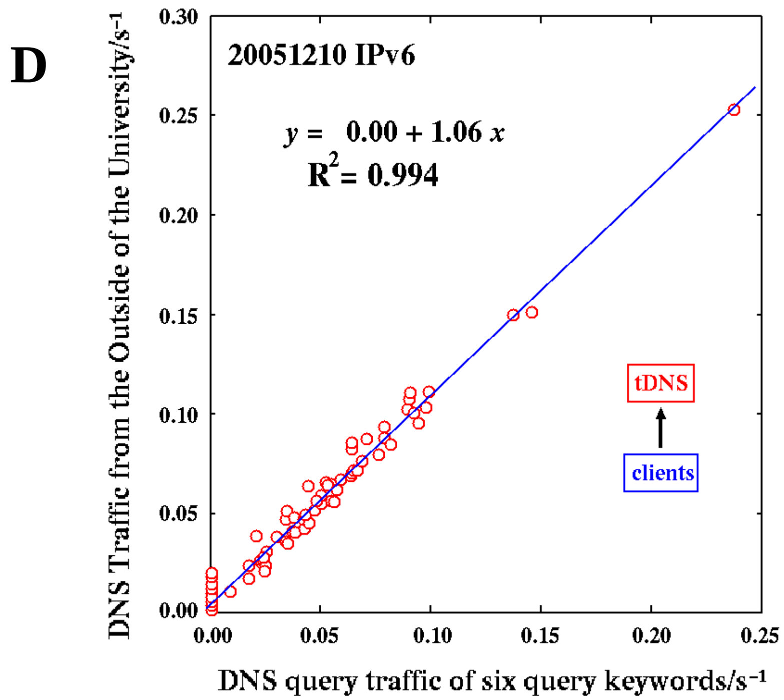
# Several Keywords for Spam Bots in IPv6 based DNS Query Traffic



**In August 5th, September 28th, October 13th, and December 10th, 2005, we can observe that the A RR based DNS query traffic includes several typical keywords as in their query contents *i.e.* "mx", "ns", "mail", "gate", "smtp", and "smtp" that transmitted by W32/Zotob variants-infected PCs.**

# Several Keywords for Spam Bots in IPv6 based DNS Query Traffic





**In August 5th, September 28th, October 13th, and December 10th, 2005, we can observe that the A RR based DNS query traffic includes several typical keywords as in their query contents *i.e.* "mx", "ns", "mail", "gate", "smtp", and "smtp" that transmitted by W32/Zotob variants-infected PCs.**

# Detection Strategies

**Statistical Analysis on:**

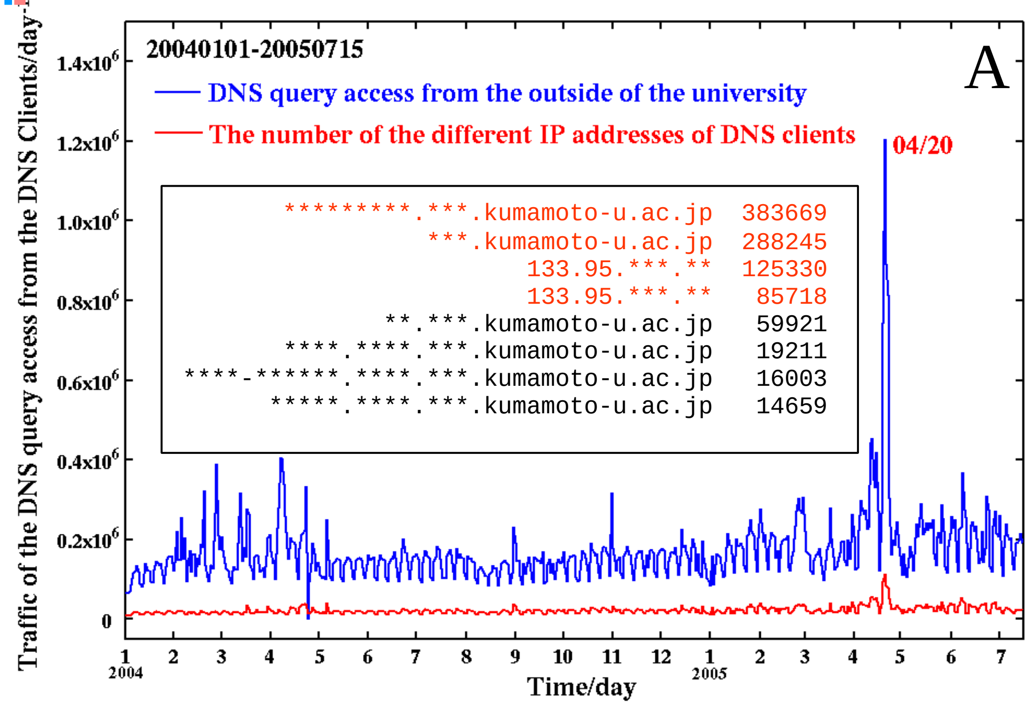
**(1) the source IP address based DNS query traffic from the bot worm (BW)-infected PC terminals in the campus network,**

**(2) the IPv6-source IP based DNS query traffic from the bot worm (BW)-infected PC terminals in the campus network, and**

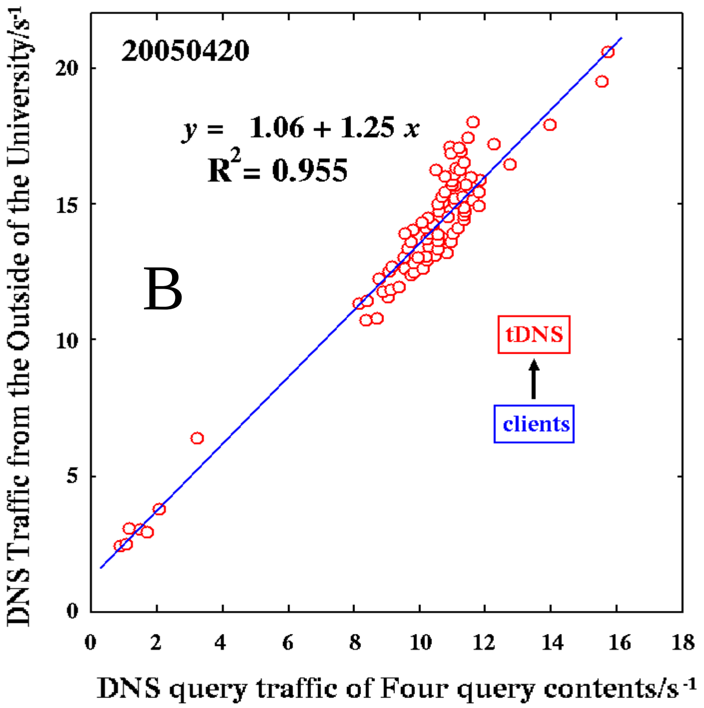
**(3) the query contents based DNS query traffic from detection systems on the internet (the other sites) like IDS/IPS, spam filter, etc.**

**IDS/IPS=Intrusion Detection/Prevention System**

# DNS Resolution Reflection/Degree of Attention?

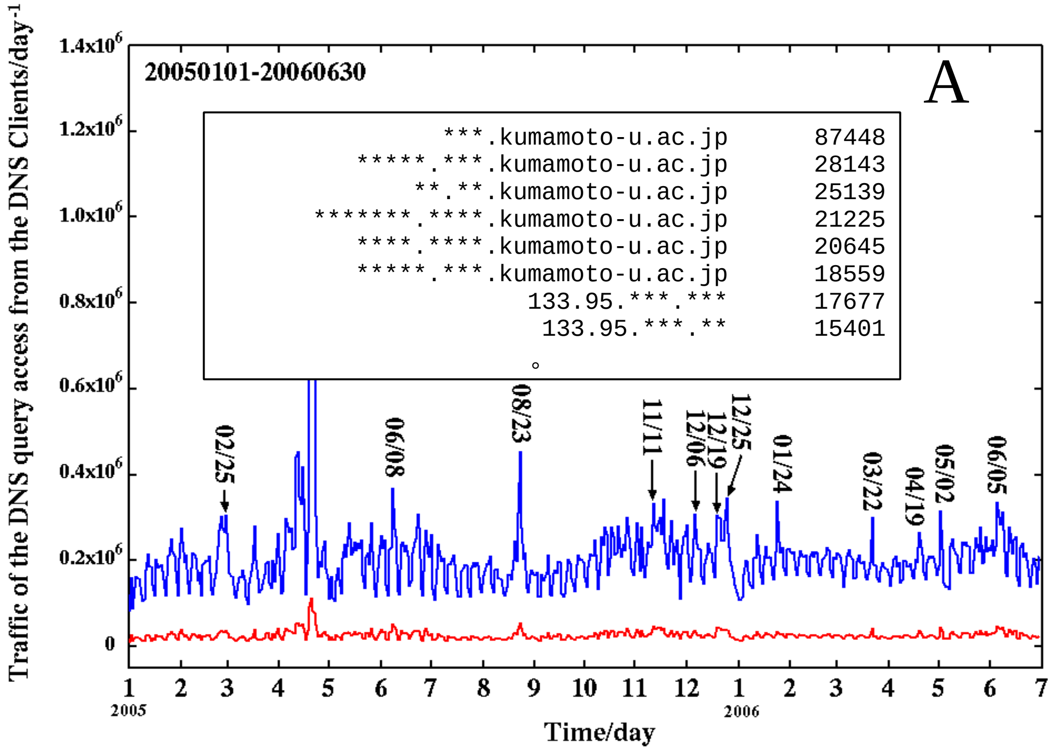


| | |
|---|---|
| *********.***.kumamoto-u.ac.jp | 383669 |
| ***.kumamoto-u.ac.jp | 288245 |
| 133.95.***.** | 125330 |
| 133.95.***.** | 85718 |
| **.***.kumamoto-u.ac.jp | 59921 |
| ****.****.***.kumamoto-u.ac.jp | 19211 |
| ****-******.****.***.kumamoto-u.ac.jp | 16003 |
| *****.****.***.kumamoto-u.ac.jp | 14659 |

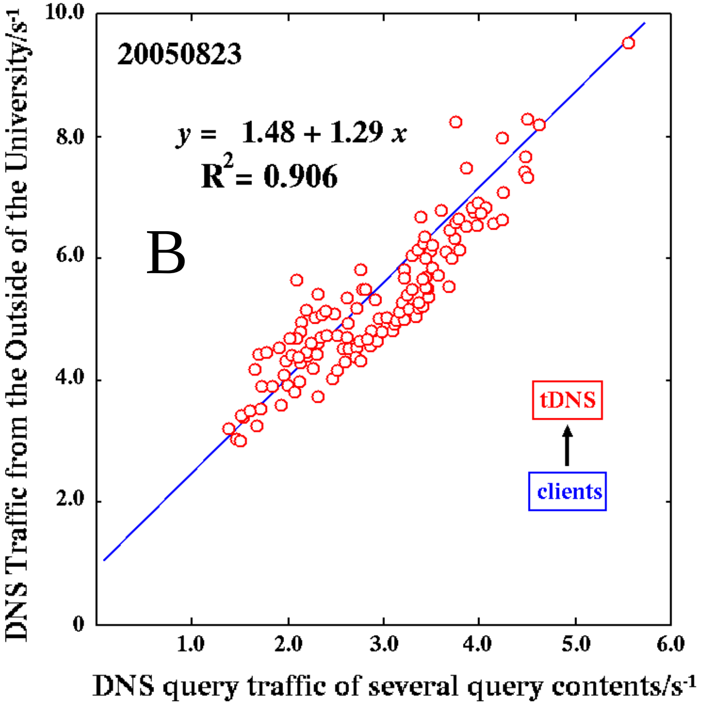


**In the query contents of the DNS query packets in the latter peak, the most largest number of contents mainly consist of an FQDN of a subdomain E-mail server, an FQDN of top domain DNS server (tDNS), and two IP addresses that related with the subdomain, respectively. Since the E-mail server is claimed as a spam-sender through the the day of 20th April, 2005, the top DNS server are severely accessed by the spam-mail detection system/spam filter world-widely at the day.**

# BW detection by watching the DNS traffic from the outside?

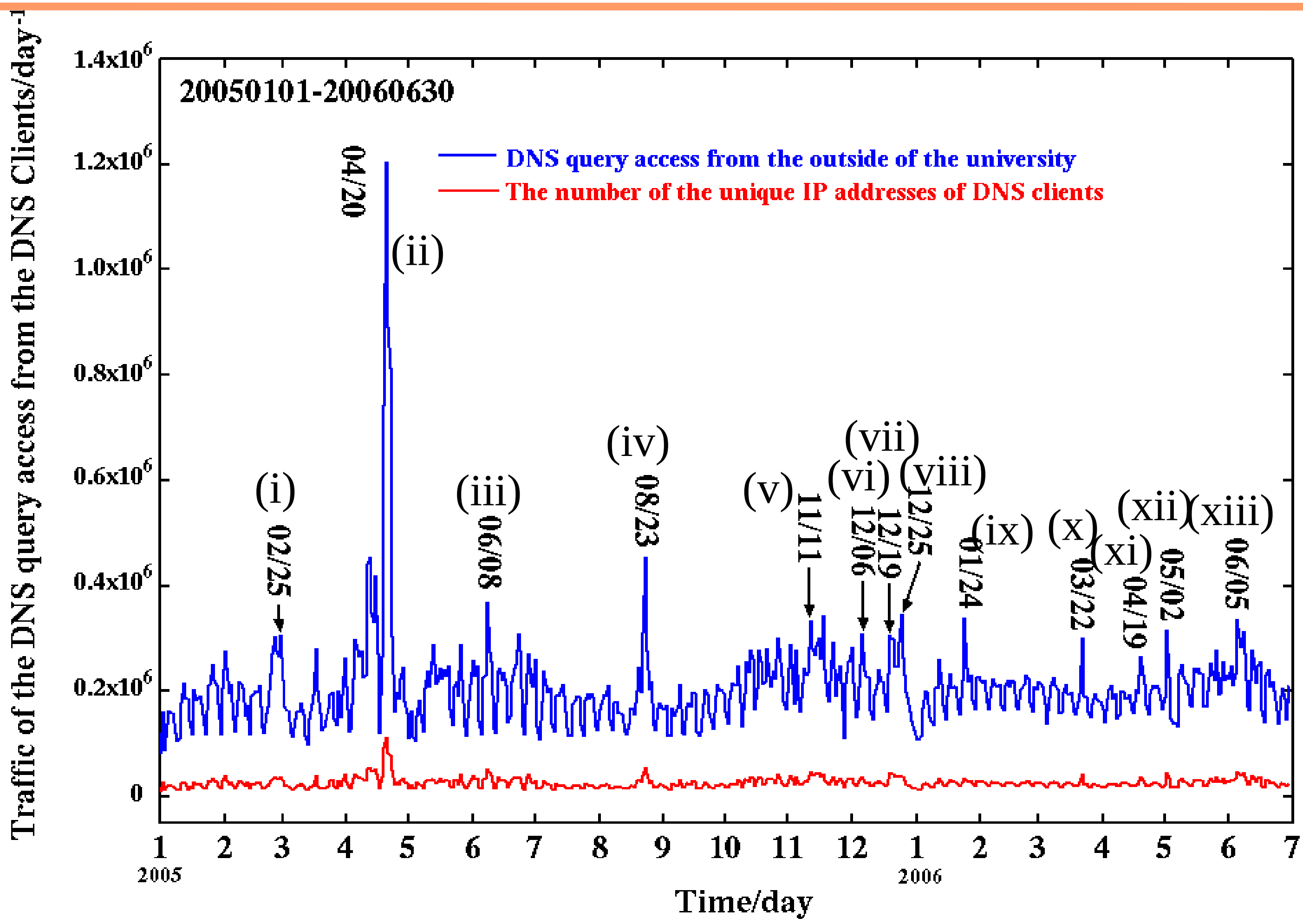


| | |
|---|---|
| ***.kumamoto-u.ac.jp | 87448 |
| *****.***.kumamoto-u.ac.jp | 28143 |
| **.**.kumamoto-u.ac.jp | 25139 |
| *******.****.kumamoto-u.ac.jp | 21225 |
| ****.****.kumamoto-u.ac.jp | 20645 |
| *****.***.kumamoto-u.ac.jp | 18559 |
| 133.95.***.*** | 17677 |
| 133.95.***.** | 15401 |



**In the query contents of the DNS query packets in the peak at 23rd August, 2005, the most largest number of contents mainly consist of several FQDNs and IP addresses that related with the local networks. This situation can be already observed in 20th April, 2005, and this feature shows that the query contents-based detection is useful for detection of the BW-infected PCs in the campus network, since infection of new W32/Zotob variants started aftert the middle days of August, 2005.**

# DNS traffic from the outside of the Campus Network

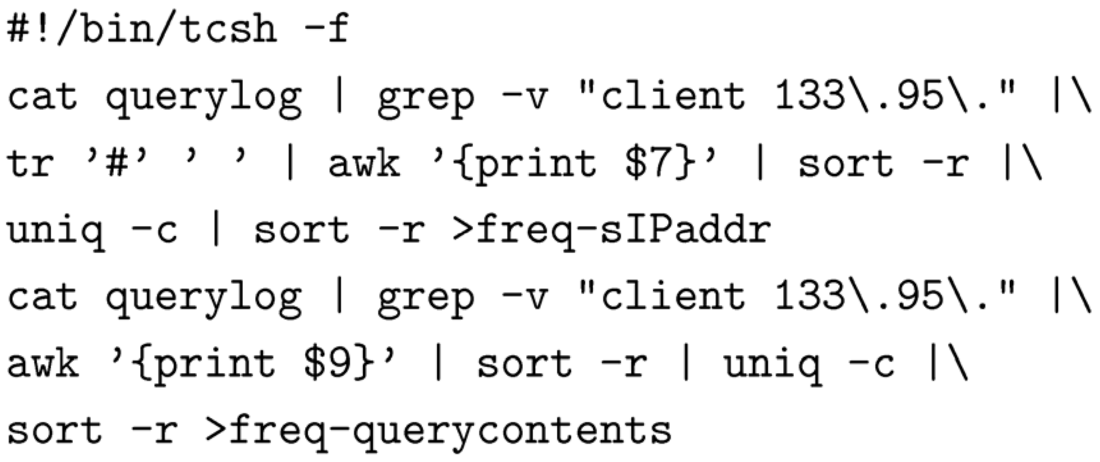


**It is of considerable importance to study more on the DNS traffic from the outside of the university.**

# Entropy

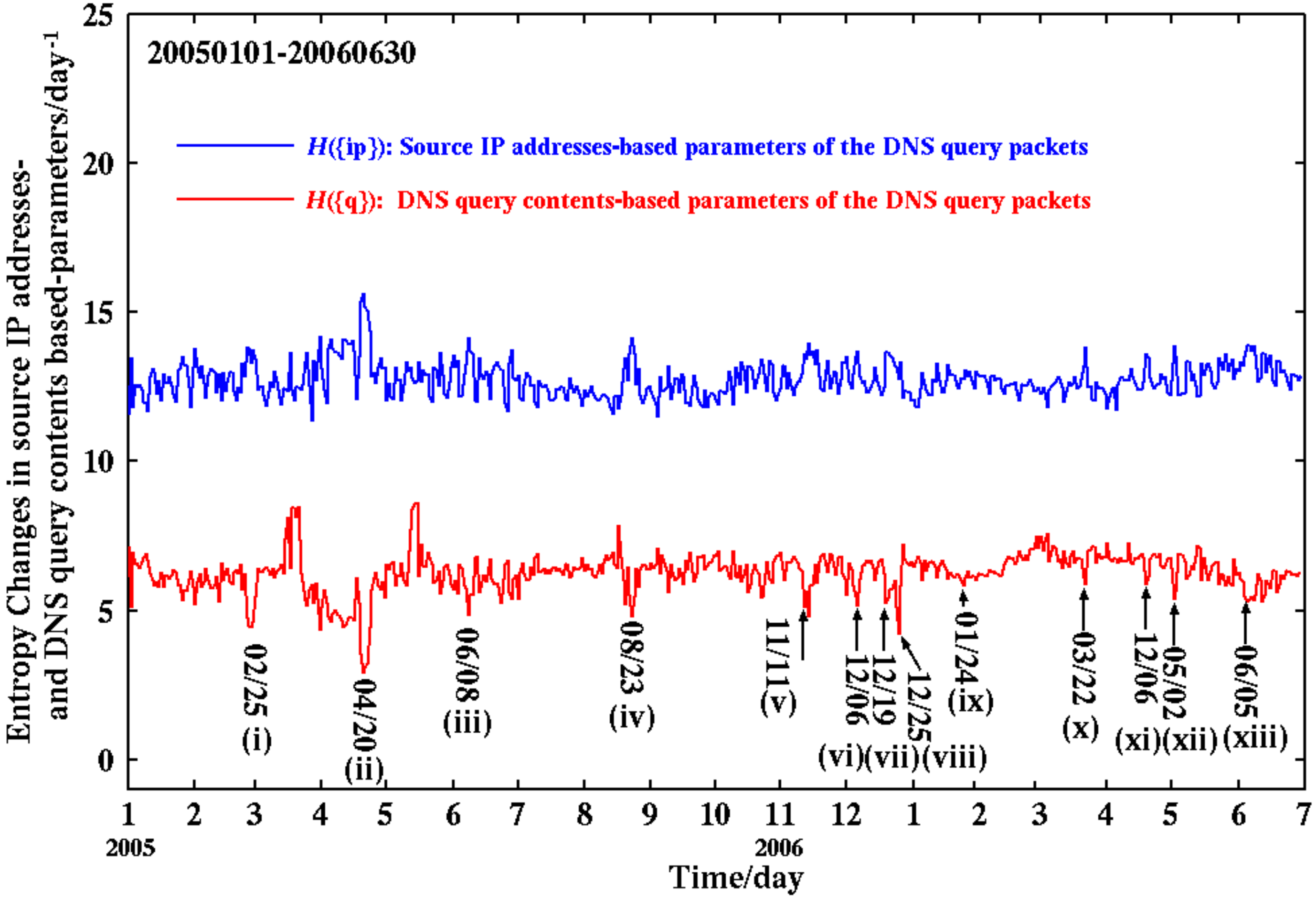
$$H(X) = -\sum_{i \in X} P(i) \log_2 P(i) \qquad \textbf{(1)}$$

$$P(i) = \frac{freq(i)}{\sum_{j} freq(j)} \qquad \textbf{(2)}$$



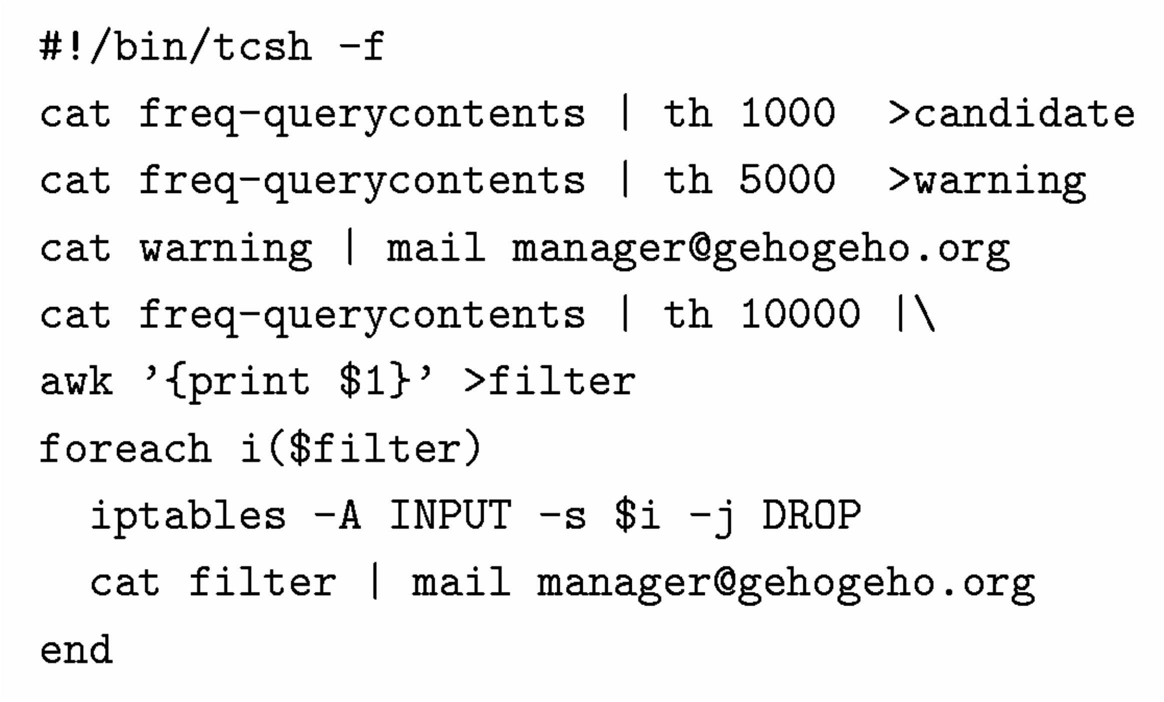
```
#!/bin/tcsh -f
cat querylog | grep -v "client 133\.95\." |\
tr ’#’ ’ ’ | awk ’{print $7}’ | sort -r |\
uniq -c | sort -r >freq-sIPaddr
cat querylog | grep -v "client 133\.95\." |\
awk ’{print $9}’ | sort -r | uniq -c |\
sort -r >freq-querycontents
```

# Entropy Analysis on the unique Source IP address and the DNS query contents in the DNS traffic



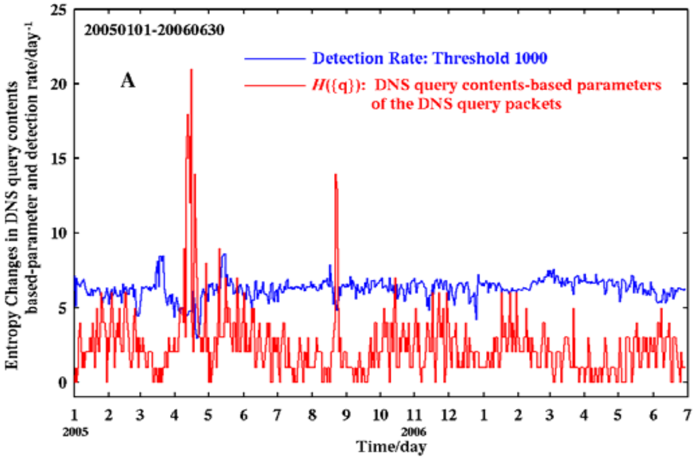
**Especially, the peaks (i)-(xiii) in the DNS query contents-based entropy curve synchronize in the previous traffic curve of DNS query packets from the outside of the campus network.**

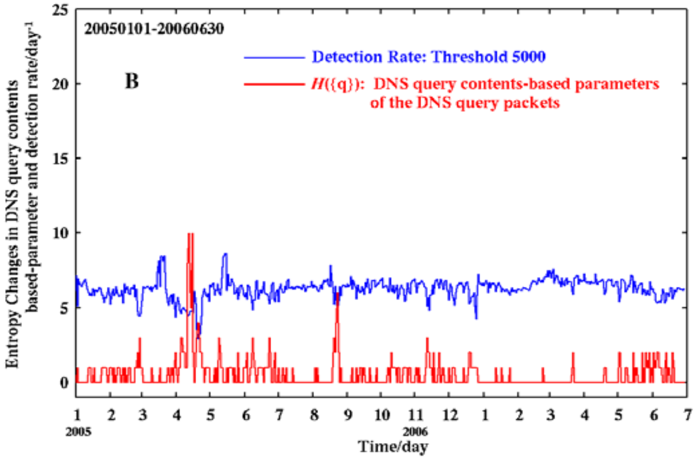
# Prototype of Detection System

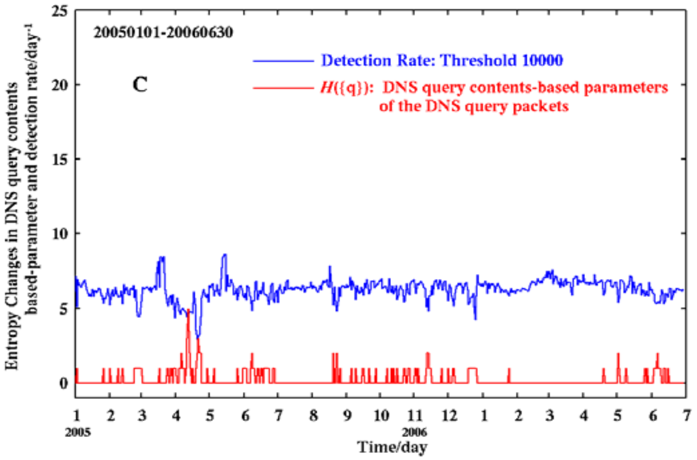


```
#!/bin/tcsh -f
cat freq-querycontents | th 1000  >candidate
cat freq-querycontents | th 5000  >warning
cat warning | mail manager@gehogeho.org
cat freq-querycontents | th 10000 |\
awk '{print $1}' >filter
foreach i($filter)
  iptables -A INPUT -s $i -j DROP
  cat filter | mail manager@gehogeho.org
end
```

# Estimation of Entropy and Detection Rate







**Threshold=1000 (Candidate as Listed):**

- **False Positive = High**
- **False Negative = Low**

**Threshold=5000 (Warning):**

- **False Positive = Medium**
- **False Negative = Medium**

**Threshold=10000 (Emergency or Critical):**

- **False Positive = Low**
- **False Negative = High**

# Conclusion and Future Work

We performed detailed statistical analysis on the traffic of the DNS query packets to the top domain DNS (tDNS) server in order to find out a detection method of the bot worm (BW)-infected PC terminals.

(1) We can observe the source IP address based DNS query traffic from the BW-infected PC terminals, especially the A RR based DNS query traffic including several keywords.

(2) We should pay much attention on the IPv6 address based DNS query packets that can be used to evade a detection system.

(3) We can also observe the useful DNS query traffic from the outside of the campus network including information on the BW-infected PC terminals in the campus network.

We are just testing the hybridized detection method and developing the zero-day incident detection system.

# Any Questions?