SpamPots Project: Using Honeypots to Measure the Abuse of End-User Machines to Send Spam

Klaus Steding-Jessen

jessen@cert.br

CERT.br – Computer Emergency Response Team Brazil
NIC.br – Network Information Center Brazil
CGI.br – Brazilian Internet Steering Committee
Our Parent Organization: CGI.br

Among the diverse responsibilities of The Brazilian Internet Steering Committee – CGI.br, the main attributions are:

- to propose policies and procedures related to the regulation of the Internet activities
- to recommend standards for technical and operational procedures
- to establish strategic directives related to the use and development of Internet in Brazil
- to promote studies and technical standards for the network and services’ security in the country
- to coordinate the allocation of Internet addresses (IPs) and the registration of domain names using <.br>
- to collect, organize and disseminate information on Internet services, including indicators and statistics
CGI.br Structure

GOVERNMENT
01- Ministry of Science and Technology
02- Ministry of Communications
03- Presidential Cabinet
04- Ministry of Defense
05- Ministry of Development, Industry and Foreign Trade
06- Ministry of Planning, Budget and Management
07- National Telecommunications Agency
08- National Council of Scientific and Technological Development
09- National Forum of Estate Science and Technology Secretaries
10- Internet Expert

N.S.
11- Internet Service Providers
12- Telecom Infrastructure Providers
13- Hardware and Software Industries
14- General Business Sector Users
15- Non-governamental Entity
16- Non-governamental Entity
17- Non-governamental Entity
18- Non-governamental Entity
19- Academia
20- Academia
21- Academia

CIVIL SOCIETY
About CERT.br

*Created in 1997 to receive, review and respond to computer security incident reports and activities related to networks connected to the Internet in Brazil.*

- National focal point for reporting security incidents
- Establishes collaborative relationships with other entities
- Helps new CSIRTs to establish their activities
- Provides training in incident handling
- Provides statistics and best practices’ documents
- Helps raise the security awareness in the country

http://www.cert.br/mission.html
Agenda

Motivation

The SpamPots Project
Open Proxy Abuse Scenario
Architecture
Honeypots
Server

Statistics

Future Work

References
Motivation

- Spam is a source of
  - malware/phishing
  - decrease in productivity
  - increase in infrastructure costs

- Spam complaints related to open proxy abuse have increased in the past few years

- Scans for open proxies are always in the top 10 ports in our honeypots’ network stats
Motivation (2)

- Brazil is usually listed as a big source of spam
  - is it really the source or is it just being abused by others?

- Need to better understand the problem and have more data about it
  - generate metrics that can help the formulation of policies
The SpamPots Project

- Supported by the CGI.br/NIC.br
  - as part of the Anti-spam Commission work

- Deployment of 10 low-interaction honeypots, emulating open proxy/relay services and capturing spam

- Installed on Brazilian ADSL/cable networks, for one year
  - 5 broadband providers, 1 residential and 1 business connection each

- Measure the abuse of end-user machines to send spam
Open Proxy Abuse Scenario

End users broadband computers

- Computer with Open Proxy
  - Computer with Open Proxy
  - Computer with Open Proxy
  - Computer with Open Proxy

Mail Server 1

Mail Server N

Victim

Victim

Victim

Victim

Victim

Victim

Victim

Victim

Victim

Victim

Victim

Victim

Victim

Victim

End users broadband computers

spammer
Architecture

End users broadband computers

spammer

Computer with Open Proxy

Honeypot emulating an Open Proxy

Server:
Collects data daily;
Monitors the honeypots resources.

Mail Server 1

Mail Server N

Victim

Victim

Victim

Computer with Open Proxy

Honeypot emulating an Open Proxy

Honeypots

- OpenBSD as the base OS
  - good proactive security features
  - pf packet filter: stateful, integrated queueing (ALTQ), port redirect
  - logs in libpcap format: allows passive fingerprinting

- Honeyd emulating services
  - Niels Provos’ SMTP and HTTP Proxy emulator (with minor modifications)
  - SOCKS 4/5 emulator written by ourselves
  - pretends to connect to the final SMTP server destination and starts receiving the emails
  - doesn’t deliver the emails

- Fools spammers’ confirmation attempts
Server

- Collects and stores data from honeypots
  - initiates transfers through ssh connections
  - uses rsync over ssh to copy spam from the honeypots

- Performs status checks in all honeypots
  - daemons, ntp, disk space, load, rsync status

- Web page interface
  - honeypot status
  - emails stats: daily, last 15min
  - MRTG: bandwidth, ports used, emails/min, etc
Statistics
## Statistics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>period</strong></td>
<td>2006-06-10 to 2007-04-30</td>
</tr>
<tr>
<td><strong>days</strong></td>
<td>325</td>
</tr>
<tr>
<td><strong>emails</strong></td>
<td>≈ 370M</td>
</tr>
<tr>
<td><strong>recipients</strong></td>
<td>≈ 3.2G</td>
</tr>
<tr>
<td><strong>avg. recpts/email</strong></td>
<td>≈ 8.9</td>
</tr>
<tr>
<td><strong>unique IPs</strong></td>
<td>≈ 160K</td>
</tr>
<tr>
<td><strong>unique ASNs</strong></td>
<td>2813</td>
</tr>
<tr>
<td><strong>unique CCs</strong></td>
<td>157</td>
</tr>
</tbody>
</table>
Spams captured / day

Emails Received [2006-06-10 -- 2007-04-30]

emails / day

0 0.5M 1.0M 1.5M 2.0M

01/07 01/08 01/09 01/10 01/11 01/12 01/01 01/02 01/03 01/04

(2006 - 2007)
Top ASNs sending spam

- Top 10 emails/ASN:

<table>
<thead>
<tr>
<th>#</th>
<th>ASN</th>
<th>ASN Name</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>9924</td>
<td>TFN-TW Taiwan Fixed Network</td>
<td>32.08</td>
</tr>
<tr>
<td>02</td>
<td>3462</td>
<td>HINET Data Communication</td>
<td>25.41</td>
</tr>
<tr>
<td>03</td>
<td>17623</td>
<td>CNCGROUP-SZ CNCGROUP</td>
<td>13.37</td>
</tr>
<tr>
<td>04</td>
<td>4780</td>
<td>SEEDNET Digital United</td>
<td>12.21</td>
</tr>
<tr>
<td>05</td>
<td>9919</td>
<td>NCIC-TW</td>
<td>02.25</td>
</tr>
<tr>
<td>06</td>
<td>4837</td>
<td>CHINA169-BACKBONE CNCGROUP</td>
<td>01.69</td>
</tr>
<tr>
<td>07</td>
<td>7271</td>
<td>LOOKAS - Look Communications</td>
<td>01.51</td>
</tr>
<tr>
<td>08</td>
<td>7482</td>
<td>APOL-AS Asia Pacific On-line</td>
<td>00.98</td>
</tr>
<tr>
<td>09</td>
<td>18182</td>
<td>SONET-TW Sony Network Taiwan</td>
<td>00.96</td>
</tr>
<tr>
<td>10</td>
<td>18429</td>
<td>EXTRALAN-TW</td>
<td>00.89</td>
</tr>
</tbody>
</table>
Top ASNs sending spam (2)

Emails Received / ASN [2006-06-10 -- 2007-04-30]

- ASN 9924 (TFN-TW)
- ASN 3462 (HINET)
- ASN 17623 (CNCGROUP)
- ASN 4780 (SEEDNET)
- Others

Months (2006 - 2007)
Top CCs sending spam

- Top 10 emails/CC:

<table>
<thead>
<tr>
<th>#</th>
<th>emails</th>
<th>CC</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>281601310</td>
<td>TW</td>
<td>76.05</td>
</tr>
<tr>
<td>02</td>
<td>58912303</td>
<td>CN</td>
<td>15.91</td>
</tr>
<tr>
<td>03</td>
<td>14939973</td>
<td>US</td>
<td>04.03</td>
</tr>
<tr>
<td>04</td>
<td>6677527</td>
<td>CA</td>
<td>01.80</td>
</tr>
<tr>
<td>05</td>
<td>1935648</td>
<td>KR</td>
<td>00.52</td>
</tr>
<tr>
<td>06</td>
<td>1924341</td>
<td>JP</td>
<td>00.52</td>
</tr>
<tr>
<td>07</td>
<td>816072</td>
<td>HK</td>
<td>00.22</td>
</tr>
<tr>
<td>08</td>
<td>776245</td>
<td>DE</td>
<td>00.21</td>
</tr>
<tr>
<td>09</td>
<td>642446</td>
<td>BR</td>
<td>00.17</td>
</tr>
<tr>
<td>10</td>
<td>355622</td>
<td>PA</td>
<td>00.10</td>
</tr>
</tbody>
</table>
Top CCs sending spam (2)

Emails Received / Country Code [2006-06-10 -- 2007-04-30]

- TW
- CN
- US
- CA
- Others

Months (2006 - 2007)
Top TCP ports used

- TCP ports used:

<table>
<thead>
<tr>
<th>#</th>
<th>TCP Port</th>
<th>protocol</th>
<th>used by</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>8080</td>
<td>HTTP</td>
<td>alt http</td>
<td>42.68</td>
</tr>
<tr>
<td>02</td>
<td>1080</td>
<td>SOCKS</td>
<td>socks</td>
<td>34.66</td>
</tr>
<tr>
<td>03</td>
<td>80</td>
<td>HTTP</td>
<td>http</td>
<td>11.22</td>
</tr>
<tr>
<td>04</td>
<td>3128</td>
<td>HTTP</td>
<td>Squid</td>
<td>06.61</td>
</tr>
<tr>
<td>05</td>
<td>3127</td>
<td>SOCKS</td>
<td>MyDoom</td>
<td>01.28</td>
</tr>
<tr>
<td>06</td>
<td>25</td>
<td>SMTP</td>
<td>smtp</td>
<td>01.18</td>
</tr>
<tr>
<td>07</td>
<td>3382</td>
<td>HTTP</td>
<td>Sobig.f</td>
<td>01.07</td>
</tr>
<tr>
<td>08</td>
<td>81</td>
<td>HTTP</td>
<td>alt http</td>
<td>00.51</td>
</tr>
<tr>
<td>09</td>
<td>8000</td>
<td>HTTP</td>
<td>alt http</td>
<td>00.37</td>
</tr>
<tr>
<td>10</td>
<td>6588</td>
<td>HTTP</td>
<td>AnalogX</td>
<td>00.27</td>
</tr>
<tr>
<td>11</td>
<td>4480</td>
<td>HTTP</td>
<td>Proxy+</td>
<td>00.15</td>
</tr>
</tbody>
</table>
Top TCP ports used (2)

Emails Received / TCP Ports [2006-06-10 -- 2007-04-30]
Top Source OS used

- tcpdump/pf.os used to fingerprint the OS of hosts originating IPv4 TCP connections

<table>
<thead>
<tr>
<th>#</th>
<th>emails</th>
<th>Src OS</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>235990984</td>
<td>Windows</td>
<td>63.74</td>
</tr>
<tr>
<td>02</td>
<td>133276691</td>
<td>Unknown</td>
<td>36.00</td>
</tr>
<tr>
<td>03</td>
<td>945642</td>
<td>Unix</td>
<td>00.26</td>
</tr>
<tr>
<td>04</td>
<td>50096</td>
<td>Other</td>
<td>00.01</td>
</tr>
</tbody>
</table>

[http://www.openbsd.org/cgi-bin/man.cgi?query=pf.os](http://www.openbsd.org/cgi-bin/man.cgi?query=pf.os)
Top Source OS used (2)

Emails Received / Source OS [2006-06-10 -- 2007-04-30]

Months (2006 - 2007)
Future Work
Future Work

- Comprehensive spam analysis
  - using Data Mining techniques
  - determine patterns in language, embedded URLs, etc
  - phishing and other online crime activities

- Propose best practices to ISPs
  - port 25 management
  - proxy abuse monitoring

- International cooperation
References

• This presentation can be found at:
  http://www.cert.br/docs/presentations/

• Computer Emergency Response Team Brazil – CERT.br
  http://www.cert.br/

• NIC.br
  http://www.nic.br/

• Brazilian Internet Steering Comittee – CGI.br
  http://www.cgi.br/

• OpenBSD
  http://www.openbsd.org/

• Honeyd
  http://www.honeyd.org/

• Brazilian Honeypots Alliance
  http://www.honeypots-alliance.org.br/