Fighting E-mail Abuse and Phishing in Brazil

Cristine Hoepers
cristine@cert.br

CERT.br – Computer Emergency Response Team Brazil
NIC.br - Network Information Center Brazil
CGI.br - Brazilian Internet Steering Committee
CERT.br Activities

Created in 1997 as the national focal point to handle computer security incident reports and activities related to networks connected to the Internet in Brazil

CERT.br

- Incident Handling
  - Coordination
  - Facilitation
  - Support
  - Statistics

- Training and Awareness
  - Courses
  - Presentations
  - Documents
  - Meetings

- Network Monitoring
  - Distributed Honeypots
  - SpamPots

International Partnerships

Yahoo! Security Week 2009 - June 2009, Sunnyvale, CA
CGI.br is a multi-stakeholder organization that, among the diverse responsibilities, has the main attributions:

• to propose policies and procedures related to the regulation of 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 (IP) and the registration of domain names using <.br>
• **to collect, organize and disseminate information on Internet services, including indicators and statistics**
Agenda

• Overview of the financial fraud scenario
  – New malware rates
  – Antivirus detection rates
• Technical challenges
• Abuse detection and international cooperation
• User awareness initiatives
Profile of Financial Motivated Fraud in Brazil

- Since 2005 fraud enabled by spam is among the top incidents notified to CERT.br
- Most common MO
  - "Generic" spam with links to ID theft malware
    - Could be a direct link to an executable, or
    - A link to a page that redirects to a file download
      - Usually involves an obfuscated scripting code
  - Most spam is sent via abuse of 3rd party networks
    - more on this later in this presentation
Overview of the System that Processes the Malware

**sm2av**
- Select new malware from malware's list
- Send malware copy to each AV vendor that does not detect the malware yet
  - email with the malware copy

**trojanfilter**
- Extract suspicious URLs from emails

**trojancheck**
- Fetch and store malware candidate
- Using AV, confirm if file is really a malware
- Create a list with the confirmed URLs
  - add new URLs
  - list entry
  - IP, date, URL, AV signature

**notify**
- Get IP contacts
- Create email with the list entry data and a email template
- Send notification asking to remove the malware
  - email with the notification

**ISTRonline**
- Try to fetch malware in order to check if it is still online
- Update stats DB including the new date and status of the malware URL
## Phishing Related Malware: 2006–2009/Q1

<table>
<thead>
<tr>
<th>Category</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009/Q1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unique URLs</td>
<td>25087</td>
<td>19981</td>
<td>17376</td>
<td>2695</td>
</tr>
<tr>
<td>Unique trojan samples (unique hashes)</td>
<td>19148</td>
<td>16946</td>
<td>14256</td>
<td>1858</td>
</tr>
<tr>
<td>AntiVirus signatures (unique)</td>
<td>1988</td>
<td>3032</td>
<td>6085</td>
<td>785</td>
</tr>
<tr>
<td>AntiVirus signatures (grouped by “family”)</td>
<td>41</td>
<td>125</td>
<td>447</td>
<td>467</td>
</tr>
<tr>
<td>File Extensions</td>
<td>73</td>
<td>112</td>
<td>112</td>
<td>51</td>
</tr>
<tr>
<td>Domains</td>
<td>5587</td>
<td>7795</td>
<td>5916</td>
<td>1121</td>
</tr>
<tr>
<td>Unique IP Addresses</td>
<td>3859</td>
<td>4415</td>
<td>3921</td>
<td>867</td>
</tr>
<tr>
<td>IP Allocation’s Country Codes</td>
<td>75</td>
<td>83</td>
<td>78</td>
<td>55</td>
</tr>
<tr>
<td>Email notifications sent by CERT.br</td>
<td>18839</td>
<td>17483</td>
<td>15499</td>
<td>2234</td>
</tr>
</tbody>
</table>

Includes:
- Keyloggers
- Screen loggers
- Trojan Downloaders

Does NOT include:
- Bots/Botnets
- Worms
2006 – 2009/Q1 Samples Sent to AVs

Trojan Samples Sent [2006-01-01 -- 2006-12-31]

Trojan Samples Sent [2007-01-01 -- 2007-12-31]

Trojan Samples Sent [2008-01-01 -- 2008-12-31]

Trojan Samples Sent [2009-01-01 -- 2009-03-31]
Technical Challenges (1/3)

- Widespread use of obfuscation in the webpages – impact in automated detection of and response to new malware URLs
  - "Proprietary" obfuscation (e.g. xor, caesar cipher, etc)
  - `JScript.Encode`
    

    "JScript.Encode is a method created by Microsoft used to encode both server and client-side JavaScript or VB Script source code in order to protect the source code from copying."

  - `JavaScript unescape()` function
    
    [http://www.javascripter.net/faq/unescape.htm](http://www.javascripter.net/faq/unescape.htm)

    ```javascript
    unescape("It%27s%20me%21")
    // result: "It's me!"
    ```
Technical Challenges (2/3)

Levels of obfuscation

E-mail with a link to a malicious Web page

Web page

JScript.Encode

xor

Unescape()

VBScript
Technical Challenges (3/3)

- What about links to financial fraud related malware in Social Networks' sites, instant messaging services and alike?
  - It is difficult to report
    - e-mail: just bounce or forward – it is easy to explain to the user
    - when reported, the information is usually incomplete
      - the context is important in cases the malware is encrypted or not yet detected
Understanding and Reducing the Abuse of Brazilian Broadband Networks for sending Spam: SpamPots Project

1st Phase Review
Motivation (1/2)

• Brazil is a big "source" of spam

• Scans for open proxies are always in the top 10 ports in our honeypots' network statistics

  http://www.honeypots-alliance.org.br/stats/

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

• Financial fraud is still using spam
Motivation (2/2)

Spams Reported by SpamCop to CERT.br – Most Common Abuse

Percentage

Open Proxy

Direct Delivery

Spamadvertised Website

Months – January 2006—April 2009
The SpamPots Project

• Main Goals
  – Have metrics about the abuse of our networks
    • Basically measure the problem from a different point of view: abuse of infrastructure X spams received at the destination
  – Help develop the spam characterization research
  – Measure the abuse of end-user machines to send spam

• Structure of the 1st phase
  – Deployment of 10 low-interaction honeypots, emulating open proxy/relay services and capturing spam
    • 5 broadband providers
    • 1 home and 1 business connection each
Location of the Sensors in the 1st Phase

End users broadband computers

Server:
- Collects data daily;
- Monitors the honeypots resources.

spammer

Honeypot emulating an Open Proxy

Computer with Open Proxy

Honeypot emulating an Open Proxy

Computer with Open Proxy

Mail Server 1

Mail Server N

Victim

Victim

Victim

Victim
## Total Data Collected in 466 Days of Operation

Data collected by 10 sensors

<table>
<thead>
<tr>
<th>Category</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-mails captured (injected)</td>
<td>524,585,779</td>
</tr>
<tr>
<td>Potential recipients</td>
<td>4,805,521,964</td>
</tr>
<tr>
<td>Average recipients/e-mail</td>
<td>≈ 9.1</td>
</tr>
<tr>
<td>Average captured e-mails/day</td>
<td>≈ 1.2 Million</td>
</tr>
<tr>
<td>Unique IPs that injected spam</td>
<td>216,888</td>
</tr>
<tr>
<td>Unique Autonomous Systems (AS)</td>
<td>3,006</td>
</tr>
<tr>
<td>Unique Country Codes (CCs)</td>
<td>165</td>
</tr>
</tbody>
</table>
### Distribution by Country Code

<table>
<thead>
<tr>
<th>#</th>
<th>CC</th>
<th>E-mails</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>TW</td>
<td>385,189,756</td>
<td>73.43</td>
</tr>
<tr>
<td>02</td>
<td>CN</td>
<td>82,884,642</td>
<td>15.80</td>
</tr>
<tr>
<td>03</td>
<td>US</td>
<td>29,764,293</td>
<td>5.67</td>
</tr>
<tr>
<td>04</td>
<td>CA</td>
<td>6,684,667</td>
<td>1.27</td>
</tr>
<tr>
<td>05</td>
<td>JP</td>
<td>5,381,192</td>
<td>1.03</td>
</tr>
<tr>
<td>06</td>
<td>HK</td>
<td>4,383,999</td>
<td>0.84</td>
</tr>
<tr>
<td>07</td>
<td>KR</td>
<td>4,093,365</td>
<td>0.78</td>
</tr>
<tr>
<td>08</td>
<td>UA</td>
<td>1,806,210</td>
<td>0.34</td>
</tr>
<tr>
<td>09</td>
<td>DE</td>
<td>934,417</td>
<td>0.18</td>
</tr>
<tr>
<td>10</td>
<td>BR</td>
<td>863,657</td>
<td>0.16</td>
</tr>
</tbody>
</table>

**Subtotal:** 99.50

### Percentage of Emails Received – Over the Period

![Graph showing the percentage of emails received over time by different countries.](image-url)
### Distribution by Autonomous System

<table>
<thead>
<tr>
<th>#</th>
<th>AS</th>
<th>CC</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>TFN-TW</td>
<td>TW</td>
<td>32.60</td>
</tr>
<tr>
<td>02</td>
<td>HINET</td>
<td>TW</td>
<td>25.04</td>
</tr>
<tr>
<td>03</td>
<td>CNCGROUP</td>
<td>CN</td>
<td>12.43</td>
</tr>
<tr>
<td>04</td>
<td>SEEDNET</td>
<td>TW</td>
<td>10.38</td>
</tr>
<tr>
<td>05</td>
<td>NCIC-TW</td>
<td>TW</td>
<td>1.75</td>
</tr>
<tr>
<td>06</td>
<td>CHINA169</td>
<td>CN</td>
<td>1.72</td>
</tr>
<tr>
<td>07</td>
<td>NDCHOST</td>
<td>US</td>
<td>1.59</td>
</tr>
<tr>
<td>08</td>
<td>CHINANET</td>
<td>CN</td>
<td>1.39</td>
</tr>
<tr>
<td>09</td>
<td>EXTRALAN</td>
<td>TW</td>
<td>1.29</td>
</tr>
<tr>
<td>10</td>
<td>LOOKAS</td>
<td>CA</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>89.26</td>
</tr>
</tbody>
</table>

### Percentage of Emails Received – Over the Period

The chart shows the percentage of emails received over a given period, categorized by different Autonomous Systems (ASes). Each line represents a specific AS, color-coded for easy identification.

- **ASN 9924** (TFN-TW/TW)
- **ASN 3462** (HINET/TW)
- **ASN 17623** (CNCGROUP/CN)
- **ASN 4780** (SEEDNET/TW)
- **ASN 9919** (NCIC-TW/TW)
- **ASN 4837** (CHINA169-BACKBONE/CN)
- **ASN 33322** (NDCHOST/US)
- **Others**

The x-axis represents the months from 2006 to 2007, while the y-axis measures the percentage of emails received.
<table>
<thead>
<tr>
<th>#</th>
<th>TCP Port</th>
<th>Protocol</th>
<th>Usual Service</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>1080</td>
<td>SOCKS</td>
<td>socks</td>
<td>37.31</td>
</tr>
<tr>
<td>02</td>
<td>8080</td>
<td>HTTP</td>
<td>alternate http</td>
<td>34.79</td>
</tr>
<tr>
<td>03</td>
<td>80</td>
<td>HTTP</td>
<td>http</td>
<td>10.92</td>
</tr>
<tr>
<td>04</td>
<td>3128</td>
<td>HTTP</td>
<td>Squid</td>
<td>6.17</td>
</tr>
<tr>
<td>05</td>
<td>8000</td>
<td>HTTP</td>
<td>alternate http</td>
<td>2.76</td>
</tr>
<tr>
<td>06</td>
<td>6588</td>
<td>HTTP</td>
<td>AnalogX</td>
<td>2.29</td>
</tr>
<tr>
<td>07</td>
<td>25</td>
<td>SMTP</td>
<td>smtp</td>
<td>1.46</td>
</tr>
<tr>
<td>08</td>
<td>4480</td>
<td>HTTP</td>
<td>Proxy+</td>
<td>1.38</td>
</tr>
<tr>
<td>09</td>
<td>3127</td>
<td>SOCKS</td>
<td>MyDoom Backdoor</td>
<td>1.00</td>
</tr>
<tr>
<td>10</td>
<td>3382</td>
<td>HTTP</td>
<td>Sobig.f Backdoor</td>
<td>0.96</td>
</tr>
<tr>
<td>11</td>
<td>81</td>
<td>HTTP</td>
<td>alternate http</td>
<td>0.96</td>
</tr>
</tbody>
</table>
Requests to the HTTP and SOCKS Modules

Number of requests received by the modules, divided according to outbound requested connection type:

<table>
<thead>
<tr>
<th>Type</th>
<th>Requests</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>connect to 25/TCP</td>
<td>89,496,969</td>
<td>97.62</td>
</tr>
<tr>
<td>connect to others</td>
<td>106,615</td>
<td>0.12</td>
</tr>
<tr>
<td>get</td>
<td>225,802</td>
<td>0.25</td>
</tr>
<tr>
<td>errors</td>
<td>1,847,869</td>
<td>2.01</td>
</tr>
<tr>
<td>total</td>
<td>91,677,255</td>
<td>100.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Requests</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>connect to 25/TCP</td>
<td>46,776,884</td>
<td>87.31</td>
</tr>
<tr>
<td>connect to others</td>
<td>1,055,081</td>
<td>1.97</td>
</tr>
<tr>
<td>get</td>
<td>1,055,081</td>
<td>1.97</td>
</tr>
<tr>
<td>errors</td>
<td>5,741,908</td>
<td>10.72</td>
</tr>
<tr>
<td>total</td>
<td>53,573,873</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Among Other Misc Activities Observed...

• Among the outgoing activity that was not aimed at port 25/TCP:
  – attempts to connect to Yahoo! servers using the Yahoo! Messenger Protocol, via the abuse of SOCKS proxies
Current Anti-spam Activities
Data Mining: Characterization of Spam Campaigns

- Frequent Pattern Tree showing different spam campaigns
  - node’s color represents a different feature that varied among the messages at that level
  - diameter of the node is proportional to the log of the frequency of the characteristic in the campaign

- Some characteristics taken into account:
  - Common keywords
  - Message layout
  - Language
  - Encoding type
  - Similar URLs
  - Services abused

Work developed by researchers from the Federal University of Minas Gerais (DCC/UFMG)
Collaboration with TW Authorities

- MoU with TW NCC (National Communications Commission), TWCERT/CC and TWIA (Taiwan Internet Association)
  - Send data weekly about spam coming from and returning to Taiwan
  - They are identifying and shutting down spammers operations
  - We are discussing the implementation of a sensor in Taiwan
Collaboration with JP Authorities

- In the past few months the activities seen changed
  - IPs assigned to Philippines are attempting to send spam to mobile phones in Japan

- JPCERT/CC and the Japanese Embassy in Brazil contacted us regarding "spam coming from Brazil"
  - the data being collected at the active sensors is being sent to them so they can pursue their investigations
  - They are sharing a case study on the success of Port 25 Management adoption in Japan, regarding the abuse of Japanese networks for sending spam
Port 25 Management Adoption Task Force

- The scenario in Brazil is very unique: regulation split the services between:
  - broadband provider – provides connectivity and IP address (responsible for network services, filters, etc)
  - ISP – authenticate the user and provide services like e-mail, web, etc
- The adoption of port 25 management need to be articulated among competing sectors
Deployment of spampots' sensors worldwide

- Global view of the data
- Help other networks to understand and prevent being abused by spammers
- Better understand the abuse of the Internet infrastructure by spammers
- Use the spam collected to improve antispam filters
- Develop better ways to
  - identify phishing and malware
  - identify botnets via the abuse of open proxies and relays
- Provide data to trusted parties
  - help the constituency to identify infected machines
  - identify malware and scams targeting their constituency
We are Looking for Partners Interested in...

• Receiving data
  – spams, URLs, IPs abusing the sensors, etc
• Hosting a sensor
• Helping to improve the technology
  – Analysis, capture, collection, correlation with other data sources, etc
• All partners will have access to all data if they want
• We are currently working with networks in the following countries/economies: AU, UY, PL, TW, HK and JP.
User Awareness
Antispam.br Website - Malicious Code Through E-mail

Tipos de spam

Códigos maliciosos

São programas que executam ações maliciosas em um computador. Diversos tipos de códigos maliciosos são inseridos em e-mail, contendo texto que se valem de métodos de engenharia social para convencer o usuário a executar o código malicioso em anexo. Em geral, esses códigos também são utilizados em spams enviados por fraudadores.

Dentre os códigos mais comuns enviados via spam, pode-se citar as seguintes categorias:

- **Backdoor**: Programa que permite a um invasor retornar a um computador comprometido. Normalmente o programa é colocado de forma a não ser notado.

- **Spyware**: Terno utilizado para se referir a uma grande categoria de software que têm o objetivo de monitorar atividades de um sistema e enviar as informações coletadas para terceiros. Podem ser utilizados de forma legítima, mas, na maioria das vezes, são utilizados de forma disfarçada, não autorizada e maliciosa.

- **Keylogger**: Programa capaz de capturar e armazenar as teclas digitadas pelo usuário no teclado de um computador. Normalmente, a ativação do keylogger é condicionada a uma ação prévia do usuário, como por exemplo, após o acesso a um site de comércio eletrônico ou Internet Banking, para a captura de senhas bancárias ou número de cartões de crédito.

- **Screentagger**: Forma avançada do keylogger, capaz de armazenar a posição do cursor e a tela apresentada no monitor, nos momentos em que o mouse é clicado, ou armazenar a região que circunda a posição onde o mouse é clicado.

- **Cavalo de tróia**: Programa, normalmente recebido como um "presente" (por exemplo, cartão virtual, álbum de fotos, protetor de tela, jogo, etc.), que além de executar funções para as quais foi aparentemente projetado, também executa outras funções maliciosas e sem o conhecimento do usuário.
Antispam.br Website - Fraud, Phishing, Scam, etc

Tipos de spam

Fraudes

Normalmente, não é uma tarefa simples atacar e fraudar dados em um servidor de uma instituição bancária ou comercial. Entretanto, atacantes têm concentrado seus esforços na exploração de fragilidades dos usuários, para realizar fraudes comerciais e bancárias através da Internet.

Para obter vantagens, os fraudadores têm utilizado amplamente e-mails com discursos que, na maioria dos casos, envolvem engenharia social e que tentem persuadir o usuário a fornecer seus dados pessoais e financeiros. Em muitos casos, o usuário é induzido a instalar algum código malicioso ou acessar uma página fraudulenta, para que dados pessoais e sensores, como senhas bancárias e números de cartões de crédito, possam ser roubados.

Golpes (Scams)

Phishing: situações em que o usuário pode receber um e-mail com links que contêm links para programas maliciosos
Com o fraudador conseguem acesso ao seu computador
Com identificador
Recomendações

Golpes (Scams)
Cartoons

• 4 videos – ≈ 4 minutes each
  • The Internet
  • The Intruders
  • Spam
  • The Defense
    – Freely available on the Internet
    – In several formats and resolutions

• English version (subtitles) already available: http://www.antispam.br/videos/english/

• English (voice-over and written texts) to be released very soon

• Q-CERT interested in making an Arabic voice-over
Video 1: The Internet
Video 2: The Intruders
Video 3: Spam
Video 4: The Defense
Additional References

• This presentation (next week)
  http://www.cert.br/docs/presentations/

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

• Contact information:
  Cristine Hoepers <cristine@cert.br>