Measuring the Security of Internet Infrastructure

Richard Clayton

8th October 2012
Richard Clayton

- Background as a software developer for the mass market
  - Amstrad CPC464, Amstrad PCW (LocoScript) and Turnpike
- Company bought by Demon Internet in 1995
- Returned to Cambridge for PhD in 2000
- Have stayed on as an academic working on security issues
- One of a number of “post-docs” funded by EPSRC/NPL
- Three year stint, July 2010 to June 2013
- Spend half my time at NPL
- My field is security economics... looking at the economics is usually a more valuable way of understanding security failures than by considering the ‘computer science’
Recent work

• R. Clayton: *Online traceability: who did that?* Consumer Focus, 26 July 2012.
  ▪ advice to OFCOM on how to write DEA initial requirements code

  ▪ debunking 27billion estimates from Detica …

  ▪ academic summary of major report for ENISA, that examined inter-ISP routeing issues, concluding that Internet has been extremely resilient thus far, but unfortunately the economics work against this continuing to be the case
ICANN “whois” project

• Shortly after I started at NPL the opportunity arose to pitch for an ICANN project to measure the usage of “privacy” and “proxy” services by Internet criminals
  ▪ our submission was to supplement the raw data collection with phone calls to putative domain owners

• Original proposal was submitted 20 July 2010 .. we were finally awarded the contract in March 2012

• Intending to be completed this year and presented at ICANN meeting (possibly Peking, April 2013)
### Whois project results (WP1: phishing)

<table>
<thead>
<tr>
<th></th>
<th>Compromised</th>
<th>Infrastructure</th>
<th>Malicious</th>
</tr>
</thead>
<tbody>
<tr>
<td>uses privacy/proxy service</td>
<td>525 24.8%</td>
<td>34 12.9%</td>
<td>139 31.0%</td>
</tr>
<tr>
<td>no phone number in whois</td>
<td>109 5.1%</td>
<td>11 4.2%</td>
<td>24 5.3%</td>
</tr>
<tr>
<td>invalid phone number or does not connect</td>
<td>623 29.4%</td>
<td>83 31.7%</td>
<td>223 49.6%</td>
</tr>
<tr>
<td>number is not answered</td>
<td>261 12.3%</td>
<td>25 9.5%</td>
<td>20 4.5%</td>
</tr>
<tr>
<td>inconclusive call or answering machine</td>
<td>80 3.8%</td>
<td>23 8.9%</td>
<td>6 1.4%</td>
</tr>
<tr>
<td>number does not work to reach registrant</td>
<td>20 0.9%</td>
<td>0 0.0%</td>
<td>29 6.4%</td>
</tr>
<tr>
<td>number works to reach registrant</td>
<td><strong>502 23.7%</strong></td>
<td>86 32.8%</td>
<td>8 1.8%</td>
</tr>
</tbody>
</table>

NPL Management Ltd - Internal
deft-whois

- I’ve spent much of my time since August 2011 working on a new way of processing “whois” information. This information is basically formatted output from a database. Traditionally one processes this with hundreds (literally) of regular expressions that extract the raw material – this gets very messy

- New approach is to provide a template for each registrar with placeholders for variable information

- Whois results are then parsed against these templates and the information extracted

- Once ICANN project is complete (and there is confidence in the coverage of the templates) this will be turned into an open source project, building a community to maintain the templates
Simple template

Domain Name : <DOMAIN>
::Registrant::
Name : <OWNER: name>
Email : <OWNER: email>
Address : <OWNER: addr>
Zipcode : <OWNER: zip>
Nation : <OWNER: cc>
Tel : <OWNER: phone>
Fax : <OWNER: fax>

::Administrative Contact::
Name : <ADMIN: name>
Email : <ADMIN: email>
Address : <ADMIN: addr>
Zipcode : <ADMIN: zip>
Nation : <ADMIN: cc>
Tel : <ADMIN: phone>
Fax : <ADMIN: fax>
Some complexities

- `<*REPEATLINE>`
  - next line is repeated whilst matches

- `<*OPTIONAL>`
  - Next line may or may not be present

- `<*OPTBLOCK> ... <*ENDBLOCK>`
  - Next block may or may not be present

- `<*ALTBLOCK> ... <*ENDBLOCK><*ALTBLOCK>...`
  - One of these alternatives will be present

- and handling for blocks of info referenced by identifier:
  - `<*COUNTER>
    `<*REPEAT>`
    `<*REPEAT>`
    `nic-hdl-br: <%INDIRECT>`
    `person: <%INDIRECT: name>丰满ıkları`