Co-operating to tackle “phishing”? 

Dr Richard Clayton
(joint work with Dr Tyler Moore)
What is “phishing”

• Person receives email from their bank indicating their information must be updated

• URL looks convincing

• Website looks convincing: so they login…
  – usually copied from the real thing!

• Multi-billion dollar losses occurring
  – risk that confidence in online banking will falter
Our research

• Studying phishing since early 2007
• Measuring fake website take-down times
  – removal of sites reduces visitors
• Identified “rock-phish” gang and showed how their methods led to longer lifetimes
  – also tracked rise of hard-to-remove “fast-flux”
• Showed how “mule recruitment” sites ignored by the banking industry [ISEC XXV 2007]
Data sources

• Originally mining PhishTank dataset
  – free and apparently accurate and substantial

• Now getting data from a brand owner and two brand protection companies (plus PhishTank and “Artists Against 419”)

• These phishing “feeds” have common components but turn out to be different…
Feeds are not shared

• Brand-protection companies obtain feeds from many places
• They also run their own detectors
• They sell feeds, but don’t share them
• Hence Company A, who sells services to Bank A1, can be unaware of sites detected by Company B – and doesn’t take them down
Bank A1’s experience as a client of BrandProtection company A
Company A v Company B

• Same pattern continues for top 6 banks for Company A and B, and for all $n$ clients

• However, less pronounced for B: which seems to have a better feed [or maybe just one that is much more aligned with ours!]

• But A’s clients bigger and proportion missed goes up with size; so B’s prowess may be more a structural issue than just extra effectiveness
<table>
<thead>
<tr>
<th>Phishing Lifetimes (hrs)</th>
<th>sites</th>
<th>mean</th>
<th>median</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Free-web hosting</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>395</td>
<td>47.6</td>
<td>0</td>
</tr>
<tr>
<td>brand-owner aware</td>
<td>240</td>
<td>4.3</td>
<td>0</td>
</tr>
<tr>
<td>brand-owner unaware</td>
<td>155</td>
<td>114.7</td>
<td>29</td>
</tr>
<tr>
<td><strong>Compromised machines</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>all</td>
<td>193</td>
<td>49.2</td>
<td>0</td>
</tr>
<tr>
<td>brand-owner aware</td>
<td>105</td>
<td>3.5</td>
<td>0</td>
</tr>
<tr>
<td>brand-owner unaware</td>
<td>155</td>
<td>103.8</td>
<td>10</td>
</tr>
<tr>
<td><strong>Rock-phish domains</strong></td>
<td>821</td>
<td>70.3</td>
<td>33</td>
</tr>
<tr>
<td><strong>Fast-flux domains</strong></td>
<td>315</td>
<td>96.1</td>
<td>25.5</td>
</tr>
</tbody>
</table>
This represents risk

- Longer lifetimes => more visitors
- Hence we can assess impact of longer lifetimes:

<table>
<thead>
<tr>
<th>Exposure figures (6 month totals)</th>
<th>A’s banks</th>
<th>B’s banks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Khour</td>
<td>$m</td>
</tr>
<tr>
<td>Actual values</td>
<td>1005</td>
<td>276</td>
</tr>
<tr>
<td>Expected if sharing</td>
<td>418</td>
<td>113</td>
</tr>
<tr>
<td>Effect of no sharing</td>
<td>587</td>
<td>163</td>
</tr>
</tbody>
</table>
Hence...

• Banks should force brand-protection companies to share feeds
  – cf the anti-virus community for last 15 years
• Brand-protection companies could form a “club” to prevent new entrants from free-riding
  – don’t have to make feeds “free”, just share them
• Expect some excitement as our message begins to sink in during this Autumn…
Co-operating to tackle “phishing”?

http://www.lightbluetouchpaper.org