# Ethical Dilemmas in Take-down Research

Richard Clayton
**richard.clayton@cl.cam.ac.uk**

joint work with

Tyler Moore
**tmoore@seas.harvard.edu**

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

UAB
24th June 2011

NPL
National Physical Laboratory

# Phishing research

- Phishing is the theft of credentials by the use of fake websites
  - though “phish by email reply” is also relevant these days
- We have a series of papers measuring this from 2007 onwards
  - all are branded “T. Moore & R. Clayton”
- So we are studying the actions of criminals – but we mainly studied (and carefully measured) the main countermeasure which is the “take-down” (the removal of) the fake websites
- This “take-down” was first done by the banks themselves, but various “brand-protection” companies now do most of the work
- Tyler and I had a paper at WECSR in March which sets out nine ethical issues that we have run up against during the course of our research. It’s “war stories” rather than philosophy!

# Dilemma 1: Should researchers notify affected parties in order to expedite take-down?

- We were measuring take-down and didn't want to interfere
  - found log-normal distributions (long tails) and that lack of information sharing was damaging effectiveness
- Who could we tell anyway?
  - no organised way to report data to banks
- Our NDAs forbade this!
  - take-down companies make money by selling data feeds
- c.f. clinical trials:: These trials can and should be stopped prematurely once the results become statistically significant and the divergence in treatment outcome is substantial
- We recommend that researchers avoid direct interference during data collection, but once the conclusions have been drawn, assistance to relevant stakeholders should be encouraged.

# Dilemma 2: Should researchers intervene to assist victims?

- We reported on 414 compromised users whose details we recovered from phishing sites (and have found more since)
- We repatriated these to banks where we could (and there is since 2010 a formal scheme for this run by the NCFTA)
- Common issue:: Torpig takeover, 180000 infections, 70G data, 1 million Windows passwords, 100000 SMTP logins, 12000 FTP credentials. Disrupted research for 6+ months
- But may be hard to locate victims, BBC's "Click" changed "wallpaper" on botted machines; apparently without having considered that this is clearly a s1 offence under the Computer Misuse Act 1990.
  - FBI was proactive on Coreflood, but only in the jurisdiction

# Rod Rasmussen (Internet Identity)

*The normal admin for the machine had been deployed to Iraq as part of his National Guard unit, and his backup was busy and hundreds of miles away that weekend because of his father's funeral. There were plenty of people looking at the machine (as in had their physical eyeballs on it) including the local sheriff, but no one was touching it since it ran the 911 Dispatch system and no one had the knowledge (as in passwords and expertise) to fix it.*

*We've also had take-downs on machines that were in hospitals, railroad stations, airports, and government facilities. While those could be just public access terminals, there's no way we can tell from the outside if that is the case or they are running life-saving equipment, switching operations, air-traffic control systems, or have sensitive data on them respectively. That's why we have a very bright line barring any sort of "write access", resetting or otherwise monkeying with content on compromised servers. Not only is it usually illegal in the US, someone's life can literally be on the line!*

# Dilemma 3: Should researchers fabricate plausible content to conduct “pure” experiments of take-down?

- Most empirical research in computer security is “observational”
- Some attempts at experiments, eg copyright issues, but considerable flaws since didn’t investigate whole process (especially US DMCA “put-back”)
- Risk of wasting the time and energy of frontline responders on fabricated requests suggests real harm is caused by the experiments. In particular, the responders typically have substantial resource constraints and already find it difficult to keep up with the number of legitimate take-down requests
- We believe the fabrication of reports to study take-down is usually unethical

# Dilemma 4: Should researchers collect world-readable data from “private” locations?

- We used “Webalizer” data from compromised websites to study the number of victims and how sites were located by criminals
- Owners of these sites may not have intended to publish their visitor data – but they did; so there’s an ethics question
- Our view was that the data enabled us to answer questions that we could not have done otherwise; and it was not personally identifiable data, just summary data for website visits
- On balance, we feel the opportunities for scientific advancement outweigh the risks to an individual website operator in collecting the data. However, it is a judgment call, and one that should be weighed on a case-by-case basis.

# Dilemma 5: What if our analysis will assist criminals?

- This is always an issue; the issue is sometimes described as the suitability of “full disclosure”
- We observed that fast-flux systems used multiple servers but this is unnecessarily cautious.
- We also made some observations about DNS time-to-live values
- We tried not to emphasis these issues and the criminals have not changed how they operate
- Long history of ethical analysis of this issue, for example Hobbs (1853) re locksmiths, and Wilkins (1641) re crypto
- General view is that the benefits of explaining how the criminals operate benefits the good guys far more than the bad guys, who already know how to do their crimes and already understand what works and what doesn’t

# Dilemma 6: Should investigatory techniques be revealed?

- It’s not science if an academic paper does not explain the methodology of an investigation
- The main effect of this has been for us to suppress more minor bits of research
- Others take a different line, such as Netcraft explaining about phishing kit “back doors” and Billy Rios discussing a file injection vulnerability in Zeus
- We don’t agree that “full disclosure” is always the overriding principle and so we choose not to publish when the details of our paper would disrupt investigations by the authorities

# Dilemma 7: When should datasets be made public or kept secret?

- Phishtank lists phishing websites, but others (Google, Microsoft, Netcraft and the take-down companies) keep lists private
- We found that sites on Phishtank were less likely to be recompromised & we found that sharing data between take-down companies would reduce phishing website lifetimes
- However, there are downsides to public sharing, it shows what the defenders know, it “names & shames” – and criminals can steal caches of credentials (as we did)
- Many systems hash the URLs so they can be compared, but the actual values are not revealed – this prevents research and prevents defenders being proactive.
- Decisions on publishing are ethical decisions

# Dilemma 8: Is the fix realistic, and does it consider the incentives of all the participants?

- We believe in security economics, and so one of the ways we look to solve security problems is to align incentives
- However, when we proposed list sharing we did not take full account of the incentives of the take-down companies; though we fixed this after they complained of our naivety
- We have proposed improvements in the take-down of child sexual abuse, unfortunately they are not realistic whilst INHOPE prevents cross-border notifications. We still think this would be the right thing to do.
- It is unethical to propose fixes to security problems that cannot be made to work in the real world

# Dilemma 9: What if the fix is worse than the problem?

- Restrictions on registering domain names would help the problem of misleading names but would not be proportionate, which is why we never suggest that type of solution
- Paul Vixie has proposed an efficient way (RPZ) of publishing lists of domains that are to be suppressed by Domain Name Servers (similar to the way that spam senders are currently handled). However, it seems unwise to institutionalise this type of suppression in a world where many groups see removal of domain names as a way to impose their world view
- We think that it is unethical to propose fixes without considering their impact, and seeking to minimise the side-effects

# Conclusion

**http://www.cl.cam.ac.uk/~rnc1/ntdethics.pdf**

- We're not proposing ethical principles but telling "war stories"
- But philosophers might usefully take our stories into account
- Also, if you're going to work in this area it would be worthwhile learning from our mistakes!

# Ethical Dilemmas in Take-down Research

**http://www.lightbluetouchpaper.org**

**http://www.cl.cam.ac.uk/~rnc1/publications.html**

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

NPL
National Physical Laboratory