Examining the Impact of Website Take-down on Phishing

Tyler Moore and Richard Clayton

University of Cambridge
Computer Laboratory

APWG eCrime Researchers Summit
Oct. 4, 2007, Pittsburgh, PA, USA

UNIVERSITY OF CAMBRIDGE

Outline

1. Mechanics of phishing
2. Rock-phish attacks
3. Who's winning the phishing arm's race?
4. Comparing performance of defenders

Outline

1. The mechanics of phishing
2. Rock-phish attacks
3. Who's winning the phishing arm's race?
4. Comparing performance of defenders

UNIVERSITY OF CAMBRIDGE

Technical requirements for phishing attacks

- Attackers send out spam impersonating banks with link to fake website
- Hosting options for fake website
  - Free webspace (`http://www.bankname.freespacesitename.com/signin/`)
  - Compromised machine (`http://www.example.com/~user/images/www.bankname.com/`)
  - Registered domain (`bankname-variant.com`) which then points to free webspace or compromised machine
- Personal detail recovery
  - Completed forms forwarded to a webmail address
  - Stored in a text file on the spoof website

Defending against phishing attacks

- Proactive measures
  - Web browser mechanisms to detect fake sites, multi-factor authentication procedures, restricted top-level domains, etc.
  - Not the focus of this paper
- Reactive measures
  - Banks tally phishing URLs
  - Reported phishing URLs are added to a blacklist, which is disseminated via anti-phishing toolbars
  - Banks send take-down requests to the free webspace operator or ISP of compromised machine
  - If a malicious domain has been registered, banks ask the domain name registrar to suspend the offending domain

Data collection methodology

- Phishing website availability
  - Several organizations collate phishing reports; we selected reports from PhishTank
  - PhishTank DB records phishing URLs and relies on volunteers to confirm whether a site is wicked
  - 33 710 PhishTank reports overs 8 weeks early 2007
  - Unfortunately, PhishTank does not indicate exactly when sites are removed and is regularly misled when sites are not disabled, but rather replaced with generic pages
  - We constructed our own testing system to continuously query sites until they stop responding or change
- Caveats to our data collection
  - Sites removed before appearing in PhishTank are ignored
  - We do not follow web-page redirectors

UNIVERSITY OF CAMBRIDGE

Outline

1. The mechanics of phishing
2. Rock-phish attacks
3. Who's winning the phishing arm's race?
4. Comparing performance of defenders

Rock-phish attacks are different!

- 'Rock-phish' gang operate different to 'ordinary' phishing sites

1. Purchase several innocuous-sounding domains (e.g., `lof80.info`)
2. Send out phishing email with URL `http://www.volksbank.de.netw.oid3614061.lof80.info/vr`
3. Gang-hosted DNS server resolves domain to IP address of one of several compromised machines
4. Compromised machines run a proxy to a back-end server
5. Server loaded with many fake websites (around 20), all of which can be accessed from any domain or compromised machine

Rock-phish attacks (cont'd.)

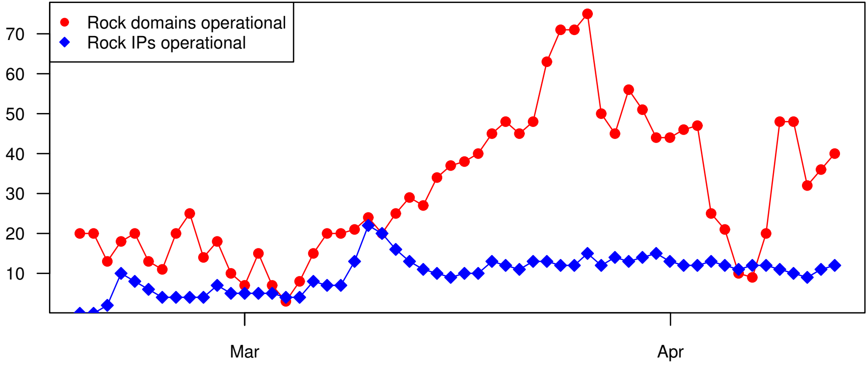
- Rock-phish strategy is more resilient to failure
  - Dynamic pool of domains maps to another pool of IP addresses
- Also increase confusion by splitting the attack components over disjoint authorities
  - Registrars see non-bank domains
  - Compromised machine owners don't see bank webpages
- Wildcard DNS confuses phishing-report collators
  - 18 680 PhishTank reports during 8 week sample (52.6% of all reports)
  - Reduces to 421 unique domains

UNIVERSITY OF CAMBRIDGE

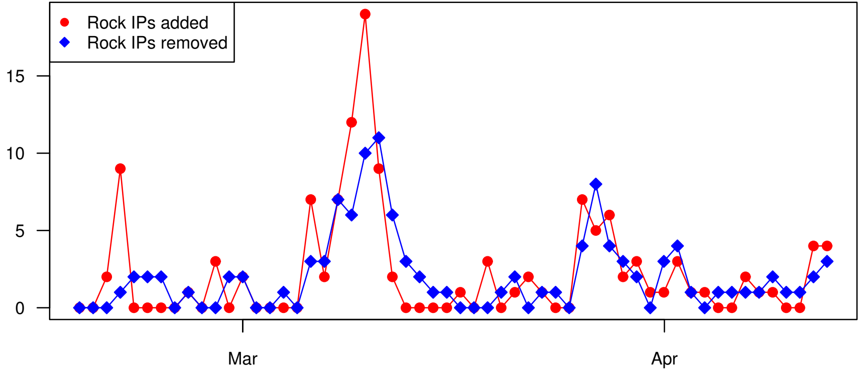
'Fast-flux' phishing domains

- Rock-phish gang's strategy is evolving fast
- In a fast-flux variant, domains resolve to a set of 5 IP addresses for a short time, then abandon them for another 5
- Burn through 400 IP addresses per week, but the upside (for the attacker) is that machine take-down becomes impractical
- Fast-flux strategy demonstrates just how cheap compromised machines are

Rock-phish site activity per day



New and removed rock-phish IPs per day

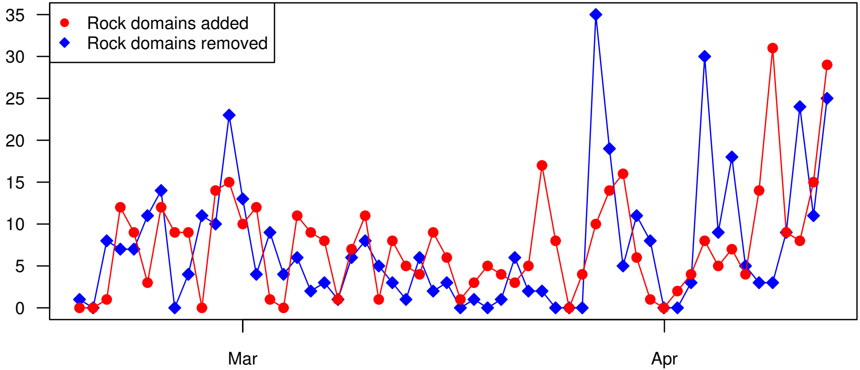


Correlation coefficient r: 0.740

Synchronized $\implies$ automated replenishment

UNIVERSITY OF CAMBRIDGE

New and removed rock-phish domains per day

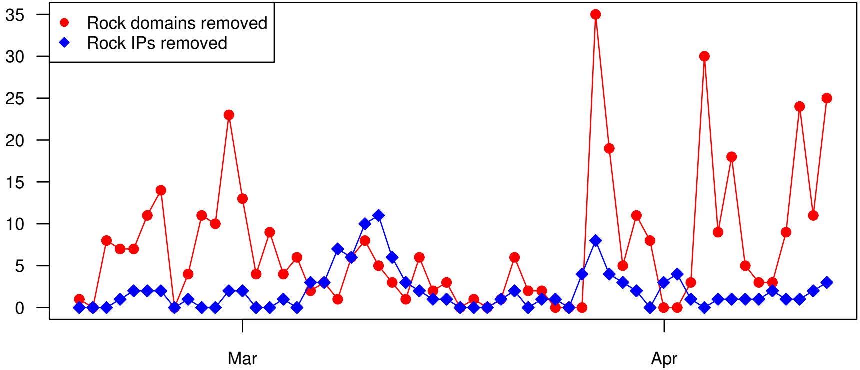


Correlation coefficient r: 0.340

Unsynchronized $\implies$ manual replenishment

UNIVERSITY OF CAMBRIDGE

Rock-phish domain and IP removal per day



Correlation coefficient r: 0.142

Unsynchronized $\Longrightarrow$ uncoordinated response

UNIVERSITY OF CAMBRIDGE

Outline

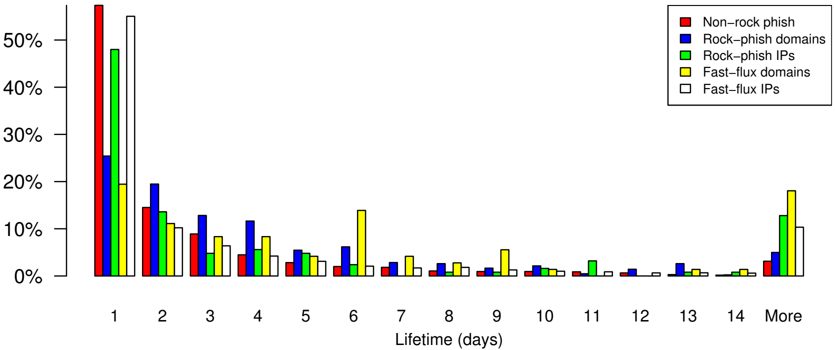
1. The mechanics of phishing
2. Rock-phish attacks
3. Who's winning the phishing arm's race?
4. Comparing performance of defenders

Phishing-site lifetimes

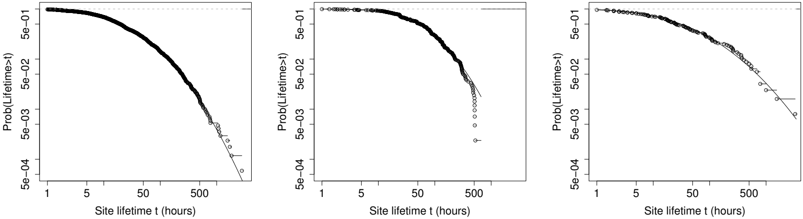
		Lifetime (hours)	
	Sites	Mean	Median
Non-rock	1 695	61.69	19.52
Rock domains	421	94.68	55.14
Rock IPs	125	171.8	25.53
Fast-flux domains	57	196.2	111.0
Fast-flux IPs	4 287	138.6	18.01

UNIVERSITY OF CAMBRIDGE

Histogram of phishing-site lifetimes



And now for some curve fitting

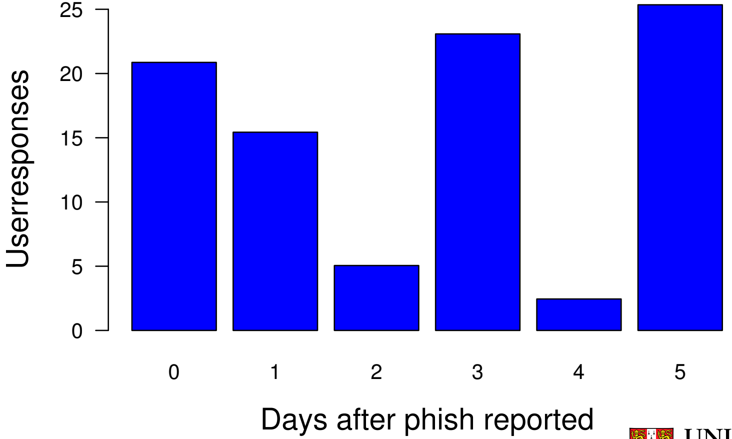


	Lognormal				Kolmogorov-Smirnov	
	μ	Std err.	σ	Std err.	D	p-value
Non-rock	3.011	0.03562	1.467	0.02518	0.03348	0.3781
Rock domains	3.922	0.05966	1.224	0.04219	0.06289	0.4374
Rock IPs	3.434	0.1689	1.888	0.1194	0.09078	0.6750

User response to phishing

- Webalizer data
  - Web page usage statistics are sometimes set up by default in a world-readable state
  - Gives daily updates of which URLs are visited
  - We can view how many times a 'thank you' page is visited
  - We automatically checked all sites reported to PhishTank for the Webalizer package, revealing over 700 sites
- On-site text files
  - We retrieved around two dozen text files with completed user details from phishing sites
  - 200 of the 414 responses appeared legitimate

User responses to phishing sites over time



Estimating the cost of phishing attacks

- Having measured how many phishing sites exist, how long they stick around, and how many people give away their details, we can estimate the losses due to phishing
- DISCLAIMER: Cost is the product of several fuzzy estimates
  1. 1 438 banking phishing sites implies 9 347 p.a.
  2. 61 hours on average implies 30 victims per site
  3. Gartner estimate cost of identity theft to be $572 per victim
  4. $9\,347 * 30 = 280\,410$ victims $* \$572 = \$160.4$m

Estimating the cost of phishing attacks (cont'd.)

- Estimate ignores rock-phish and fast-flux
  - Since rock-phish account for a large proportion of spam, we assume that they are at least as successful as ordinary phishing sites
  - Our final minimum cost estimate: $320m p.a.
- Gartner estimates 3.5m people fall victim to identity theft at a cost of $2Bn p.a.
  - Part of the disparity can be accounted for our conservative counting of sites
  - The difference can also be accounted for by other types of identity theft (theft of merchant databases, Trojan programs operating keyloggers, etc.)

UNIVERSITY OF CAMBRIDGE

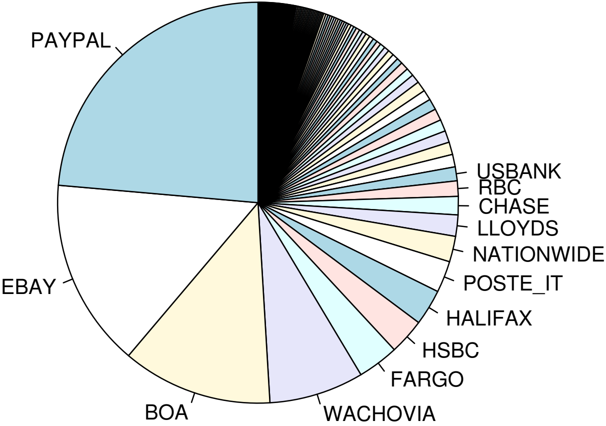
Outline

1. The mechanics of phishing
2. Rock-phish attacks
3. Who's winning the phishing arm's race?
4. Comparing performance of defenders

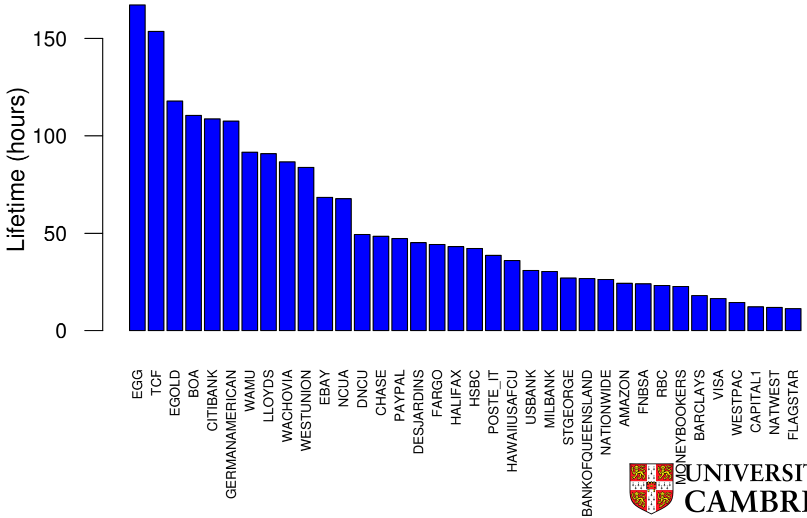
Breaking down site lifetimes

- Phishing site lifetimes vary greatly, but can we make sense of the differences?
  - We have already established that the rock-phish gang are more effective than other attackers
  - Do some banks perform better than others?
  - Do some ISPs respond better than others?
  - Does the timing of attacks make any difference?
- Identifying exceptional performers (both good and bad) could help encourage improved response times

Number of phishing sites per bank



Phishing-site lifetimes per bank (only banks >= 5 sites)



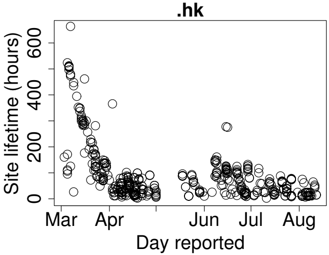
UNIVERSITY OF CAMBRIDGE

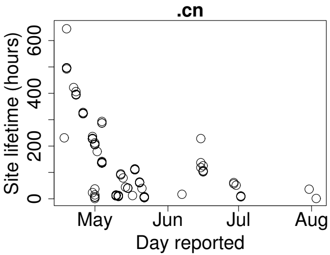
Take-down performance of free-website hosts

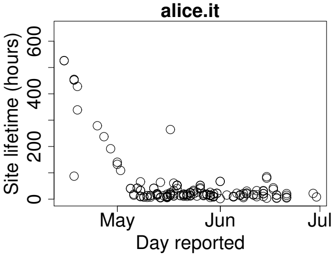
- Some phishing attacks are hosted on free webspace
- Overall, these sites are removed more quickly than sites hosted on compromised web servers
- But significant variation remains

	Sites	Mean lifetime	Median lifetime
yahoo.com	174	23.79 hours	6.88 hours
doramail	155	32.78 hours	18.06 hours
pochta.ru	1 253	33.79 hours	16.83 hours
alice.it	159	52.43 hours	18.83 hours
by.ru	254	53.11 hours	38.16 hours

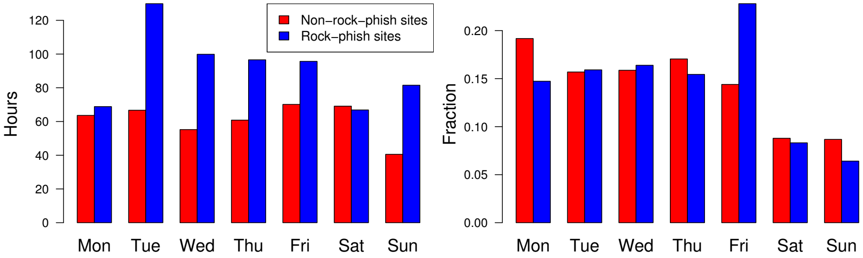
'Clued-up' effect on free host & registrar take-down times







Do weekends adversely impact phishing site removal?



Phishing site lifetime by weekday (left) and number of reported phishing sites by weekday (right)

Discussion

- Collusion dividend for rock-phish gang
  - Cooperation has strengthened the gang: pooling resources to swap between machines while impersonating many banks per domain
  - Should have attracted more attention from the banks, but perhaps sum-of-efforts nature of the cooperation enables banks to free-ride off each other's vigilance
- Countermeasures
  - Direct tactics like reducing the # of compromised machines available or rate-limiting domain registration appears futile
  - Transparency could help: publishing take-down performance by bank, ISP and country may pressure improvements
  - Increasing awareness to targeted banks of rock-phish tactics may trigger cooperation

Conclusions

- We have established that there is wide disparity in phishing site lifetimes
    - Long-tailed distribution of lifetimes implies that a few long-lived sites are undermining the effectiveness of take-down countermeasures
    - Some banks and ISPs are doing better than others
    - Disparity also suggests there is room for improvement through better monitoring
- We have also seen that attackers innovate: rock-phish sites outlive ordinary phishing sites through clever adaptations in strategy
- For more, see `http://www.cl.cam.ac.uk/~twm29/` and `http://www.lightbluetouchpaper.org/`

UNIVERSITY OF CAMBRIDGE