Examining the Impact of Website Take-down on Phishing

Tyler Moore and Richard Clayton
Computer Laboratory, University of Cambridge
15 JJ Thomson Avenue, Cambridge CB3 0FD, United Kingdom
{tyler.moore},{richard.clayton}@cl.cam.ac.uk

ABSTRACT

Banks and other organisations deal with fraudulent phishing websites by pressing the hosting service providers to remove the sites from the Internet. Until they are removed, the fraudsters will learn the passwords, personal identification numbers (PINs) and other personal details of the users who are fooled into visiting them. We analyse empirical data on actual phishing website removal times and the number of visitors that the websites attract, and conclude that website removal is part of the answer to phishing, but it is not fast enough to completely mitigate the problem. The removal times have a good fit to a lognormal distribution, but within the general pattern there is ample evidence that some service providers are faster than others at removing sites, and that some brands are able to get fraudulent sites removed more quickly. We particularly examine a major subset of phishing websites (operated by the 'rock-phish' gang) which accounts for around half of all phishing activity and whose architectural innovations have extended their average lifetime. Finally, we provide a ballpark estimate of the total loss being suffered by the banking sector from the phishing websites we observed.

1. INTRODUCTION

Phishing is the process of enticing people into visiting fraudulent websites and persuading them to enter identity information such as usernames, passwords, addresses, social security numbers, personal identification numbers (PINs) and any further information that can be made to seem plausible. This information is then used to impersonate the victim so as to empty their bank account, run fraudulent auctions, launder money, apply for credit cards, take out loans in their name, and so on. Although most current phishing attacks target the banks, phishing websites regularly appear for businesses as diverse as online auctions (eBay), payment sites (PayPal), share dealers (E*Trade), gambling websites (PartyPoker), social-networking sites (MySpace) and online retailers (Amazon).

June 10, 2007

The academic work on phishing has been diverse, with a useful starting point being the book by Jakobsson [5]. Researchers have tried to understand the psychology of the process [3], how to block the spam email containing the initial enticement [8], how server operators might automatically detect fraudulent sites [18], and whether there are patterns to their occurrence [12]. There have been many proposals for browser mechanisms to detect phishing websites [10, 20], and schemes to prevent users from disclosing their secrets to them [13]. Others have looked at disseminating information about the trustworthiness of websites through central repositories (blacklists) or social networks [1], although at present it seems that users generally ignore any cues that tell them that websites are likely to be malicious [14, 19].

In this paper we consider phishing from a completely different angle. The banks (and other organisations being impersonated) are dealing with the fake websites through 'take-down' procedures, so that there is nothing there for a misled visitor to see. Our aim has been to determine how effective this strategy has turned out to be, and whether it is likely to be sufficient on its own to prevent phishing from being profitable.

We monitored the availability of several thousand phishing websites in Spring 2007. Our results show that a typical phishing website can be visited for an average of 58 hours, but this average is skewed by very long-lived sites – we find that the distribution is lognormal – and the median lifetime is just 20 hours. We were able to examine web log summaries at a number of sites, along with some detailed records of visitors that a handful of phishers inadvertently disclosed. This allowed us to estimate the number of visitors who divulged their data on a typical site to be 25 if it remained up for one day, and growing by 10 more per day thereafter.

We also identified a significant subset of websites (about half of all URLs being reported) which were clearly being operated by a single "rock-phish" gang. These sites attacked multiple banks and used pools of IP addresses and domain names. We found that these sites remained available for an average of 94 hours (again with a lognormal distribution, but with a median of 55 hours). A newer architectural innovation dubbed "fast-flux" that used hundreds of different compromised machines per week, extended the website availability to a median of 202 hours. Within the overall figures, we show that some brands are considerably faster than others in getting spoof websites removed, and also that there is a wide disparity in response times from different hosting providers.

We see 'take-down' as a reactive strategy, an increasingly prevalent trend in the way that security issues are being handled. Software vendors wait for vulnerabilities to be discovered and then issue patches. Anti-virus tools update their databases with new signatures as new viruses are identified. In these reactive approaches, the defenders aim to identify the bad guys as quickly as possible to minimise exposure, while the bad guys scramble to open new security holes at a sufficiently fast rate to continue their activities.

In this case our figures demonstrate that a reactive strategy does reduce the damage done by phishing websites. However, it is clearly not occurring fast enough to prevent losses from occurring, and so it cannot be the only response. In particular, we use the lifetime and visitor numbers above to show that, on fairly conservative extrapolations, the banks' losses that can be directly attributed to ordinary phishing websites are some $178m per annum, with a similar amount being raked in by the rock-phish gang.

The rest of the paper is arranged as follows. We first set out a model of the mechanics of a phishing attack in Section 2, presenting the arms race resulting from the tactics available to both attacker and defender. In Section 3.1 we set out our methodology for gathering data about phishing websites to compute take-down times, and in Section 3.2 explain how we estimate the time distribution of phishing responses. In Section 4 we describe a particularly pernicious category of phishing site called '*rock-phish*', which simultaneously impersonates many banks and regularly cycles through domain names and IP addresses. In Section 5 we analyse our results and find that by the time phishing sites are removed, damage has already been done: many responses have been received and the attackers are moving on to new sites. Finally, in Section 6, we discuss what our results mean in terms of practical strategies for the banks (and the phishing attackers).

2. THE MECHANICS OF PHISHING

To carry out phishing scams, *attackers* transmit large numbers of *spam emails* which include *links* (URLs) to websites under their control. The spam emails must resemble legitimate email, so that unsuspecting users will consider them genuine. The spam must also contain an appropriate message so that users will act upon it, be it an impending account suspension, a payment for a marketing survey, or a report of a transaction that the user will know to be fake and must therefore be cancelled [3]. The email must also be able to evade the user's spam *filters*. Looking like genuine email clearly helps, but the filters may also have access to a *blacklist* of URLs that are currently being promoted, so that there is value in varying the URL to prevent matches occurring.

The user connects to a *spoof website* by clicking on a link in the email. Their *web browser* may access the website directly or be redirected from an initial site (perhaps exploiting apparently legitimate redirector systems at Google[1] or eBay) to the actual phishing web pages. At this stage browsers may apply their own heuristics and consult their own blacklists to determine if the site should be blocked as clearly illegitimate. Provided the browser does not interfere, the user will then be presented with an accurate imitation of the legitimate company's pages (often including all the links to warnings about fraud), and thus reassured will then fill in their *personal details*. Although a handful of sites validate these details immediately, it is more common for any response at all to be accepted.

The compromised details are usually emailed to a *webmail address*, but are sometimes stored in *plain text files* at the spoof website, awaiting direct collection by the fraudster. Once they have received the compromised details they will discard the obviously fake and then sell on the details to *cashiers* who will empty the bank accounts [15], perhaps transferring the money via a *mule* who has been recruited via further spam email seeking 'financial consultants' to accept and relay payments for a commission.

The spoof website is sometimes hosted on 'free' webspace, where just anyone can register and upload pages, but it is more usually placed on a compromised machine; perhaps a residential machine, but often a server in a data centre. The hijacked machine will have come under the attacker's control either through a security vulnerability (typically unpatched applications within a semi-abandoned 'blog' or message-board), or because the user is running some malware, delivered by email or downloaded during a visit to a malicious website.

If the website is on 'free' webspace a typical URL will be of the form `http://www.bankname.freespacesitename.com/signin/` where the `bankname` is chosen to match or closely resemble the domain name of the financial institution being attacked. Changing the hostname is not always possible for compromised machines, and attackers may have restricted permissions, so they will add their own web pages within an existing structure, leading to URLs of the typical form `http://www.example.com/~user/www.bankname.com/` where the `bankname` is present to lend specious legitimacy should the user check which site they are visiting, yet fail to appreciate the way in which URLs are really structured.

To avoid the use of `example.com`, the URL may use just the IP address of the compromised machine, perhaps encoded into hexadecimal to obscure its nature. However, to further allay suspicion, the fraudsters will sometimes go to the effort of registering their own domain name, which they will then point at either free webspace, which can often be configured to allow this to work, or to a compromised machine where they have sufficient control of the web server configuration. The domain names are usually chosen to be a variation on `bankname.com` such as `bankname-usa.com`, or they will use the bank's name as a subdomain of some plausible, but superficially innocuous domain, such as `bankname.xtrasecuresite.com`. A half-way house to an actual domain name is the use of systems that provide domain names for dynamic IP address users, which results in the usage of domains such as `bankname.dyndns.org`.

Defence against phishing attacks is primarily carried out by the impersonated *targets* (banks etc.) themselves, with significant assistance from a number of technically-savvy volunteers, who often work at Internet Service Providers (ISPs). Suspicious emails will be reported by some of the users who received them, either to the targeted institution, or to one of several *collators* – entities that keep a record of reported phishing sites. Newer web browsers, such as Microsoft's Internet Explorer 7 and Mozilla's Firefox 2, contain single click reporting systems [7, 9] to make user reporting

[1]In February 2007 Google started to detect usage of their redirectors and provide a warning message [2], so it is likely that other redirectors will now be used in preference.

as simple as possible. In addition, spam filtering systems are increasingly picking out phishing emails by generic characteristics, and automatically generating reports where the link they contain was not previously known.

The recipients of the reports will then examine the site being linked to, in order to determine if it is illegitimate. Once a reported phish has been vetted, the URL will be added to the blacklists to block further email spam and to assist anti-phishing browser toolbars and other mechanisms in assessing the site's (in)validity. Meanwhile, the defenders will send a *take-down request* to the operator of the free webspace, or in the case of a compromised machine, to the relevant *ISP* who will temporarily remove it from the Internet or otherwise ensure that the offending web pages are disabled. Where a domain name has been registered by a phishing attacker, the defenders will ask the *domain name registrar* to suspend the offending domain. However, not all ISPs and registrars are equally co-operative and knowing that a phishing site exists does not automatically cause its removal. Some ISPs take down phishing sites immediately, while others do not act especially promptly. Responsiveness often varies by company and by country, as well as with the competence (and language skills) of the organisation requesting the removal.

3. DATA COLLECTION

The average duration for which phishing sites are accessible is an important measure of the state of phishing attack and defence. Most phishing sites are identified and removed within a few days, yet there must have been sufficient visitors during that period – because the attackers do not appear to be discouraged, but move on to new locations and continue their activities. We now describe a methodology for quantifying phishing site duration and determining the distribution of user-responses.

3.1 Phishing website availability

We gathered phishing reports from 'PhishTank' [11], one of the primary phishing-report collators. Comparison of their datasets with other public sources such as 'Castle Cops' and Google showed that their collection was by far the most complete and timely. The PhishTank database records the URL that has been reported to them, the time of that report, and sometimes further detail such as `whois` data or screenshots of the website. Volunteers use the URL to examine the website and determine whether it is indeed a phishing website or an incorrect report (perhaps of a legitimate bank).

Unfortunately, PhishTank does not provide an exact indication of when sites are removed, and its systems are regularly misled when phishing websites are not disabled, but replaced with generic advertising web pages. We therefore constructed our own testing system which, of necessity, became rather complex.

This system fetches reports of confirmed phishing websites from PhishTank and records exactly when PhishTank first learnt of the site. In order to track the existence of the website independently of whether its host name can be resolved, further records are constructed by replacing the host name part of the URL with the IP address it resolves to and the reverse DNS lookup of that IP address. These extra records also help to link together multiple reports of the same site. Additional canonicalisation is done to link together reports with or without trailing `/` characters, or when `index.html` (`index.php` etc.) are provided in some reports and not others.

We tested all of the sites in our database on a continuous basis, twice every hour, to determine if they were still accessible. The web page data fetched (along with its HTTP headers) was fingerprinted so that significant changes (anything apart from date-stamps, session-IDs, etc.) could be detected. Just prior to fetching the page, the host name was once again resolved (having ensured that there was no cached data in the DNS server) and if it had moved to a new IP address further records for that IP address (and its reverse DNS lookup) were added to the database as required. A website that returned a '404' error was removed from the database, but timeouts and other temporary failures were retried for at least 48 hours.[2]

This testing regime enables us to precisely (with an accuracy of about 30 minutes) determine when a phishing website is removed or changed, whilst remaining tolerant of temporary outages. Where multiple database entries pointed at the same web page, the fingerprinting enabled us to detect this and remove the duplicates. Also, for known malicious sites with identical fingerprints (and, in particular, the rockphish attacks described in Section 4), we immediately categorised the sites as malicious, without waiting to discover whether the PhishTank volunteers had correctly done so.

In practice, our observations showed that phishing websites were entirely static, and hence any change in fingerprint was sufficient to indicate that it had been removed, or further requests were showing a generic page. This simplified our monitoring considerably, but it was still necessary to view the first page we captured to determine which institution was being targeted or, as sometimes happened, whether it was already removed by the time we learnt of its existence.

3.2 Visitor statistics

We also wished to gain a better understanding of the distribution of user responses to phishing attacks, and were able to gather some limited information about how many visitors a typical website received, and how many filled in the web form and provided any data.

In a small number of cases (less than two dozen so far) the site recorded details of victims into text files that were stored on the site itself in such a way that we could retrieve them. Inspection of these files showed how many responses were received and whether or not they were likely to be valid. Some of the entries were clearly testing (random sequences of characters), or consisted of profanities directed at the recipient of the data. The remainder of the responses were counted as valid, although it is understood that some banks deliberately provide data on dummy accounts for their own tracing purposes, so our counts will to some minor extent overestimate the number of people actually compromised.

In other cases we have collected publicly available web page usage statistics collated by the sites where the phishing pages are residing. Webalizer [17] is a particularly popular package, which is often set up by default in a world-readable state on the type of web servers that seem to be regularly compromised. Indeed, it may be unpatched Webalizer vul-

[2]At present, we are excluding all sites that involve non-standard forms of redirection to reach the final phishing webpage. This avoids considerable complexity (some phishers even use Macromedia flash files to redirect traffic), at the expense of a lack of completeness.

nerabilities that permitted access in the first place. These statistical reports provide daily updates as to which URLs are visited, and these can be used to determine the total number of visitors and how many reached the 'thank you' page that is generally provided once personal data has been uploaded. By assuming that similar proportions of these 'hits' are valid occurrences of visitors compromising their identity information, it is possible to form a view as to the effectiveness of the phishing exercise and the distribution of visitors day by day. As new reports are obtained from PhishTank, we have automatically queried sites to determine whether Webalizer is running; if so, we returned daily to collect new reports. In all, we discovered over 700 phishing sites using Webalizer in this manner.

4. ROCK-PHISH ATTACKS

In Section 2 we described the way in which typical phishing websites were operated with web pages added to existing structures and the occasional use of misleading domain names. However, the 'rock-phish' gang operate (in early 2007) in a rather different manner. Having compromised a machine they then cause it to run a proxy system that relays requests to a back-end server system. This server is loaded with a large number (up to 20 at a time) of fake bank websites, all of which are available from any of the rock-phish machines. The gang then purchase a number of domain names with short, generally meaningless, names such as `lof80.info`. The email spam then contains a long URL such as: `http://www.volksbank.de.networld.id3614061.lof80.info/vr` where the first part of the URL is intended to make the site appear genuine and a mechanism such as 'wildcard DNS' can be used to resolve all such variants to a particular IP address.

Transmitting unique URLs trips up spam filters looking for repeated links, fools collators like PhishTank into recording duplicate entries, and misleads blacklist users who search for exact matches. Since the numeric values are sent to the DNS server (which the gang also hosts) it is clear that tracking of responses is possible along with all kinds of customisation of responses. However, which bank site is reached depends solely upon the url-path (after the first `/`). Hence, a canonical URL such as `http://www.lof80.info/` is sufficient to fetch a top level web page and its fingerprint is sufficient to identify the domain and associated IP address as owned by the rock-phish gang.

The gang's methods have evolved over time – they originally placed all their websites into a `/rock` directory (hence their name), morphed later into `/r1` but now this directory name is dispensed with (although we found that `/r1/vr/` still works as a synonym for `/vr`). The gang's evolution has been tracked well enough, and their methods differ so much from other phishing websites, that it is useful to measure their activities separately for this study. In particular, their email spam, which has a characteristic section of random text followed by a GIF image containing the actual message, is estimated to account for between one third and one half of all phishing email. The rock-phish gang is believed to be extremely successful, and it is claimed that they have stolen in excess of $100m so far [6].

For traditional phishing sites, removing either the hosting website or the domain (if only used for phishing), is sufficient to remove a phishing site. However, rock-phish sites share hosts – so that if one is removed, the site automatically switches to working machines which are still hosting a copy of the proxy. This switching behaviour provides the strongest evidence that rock-phish sites collude. To verify this collusion, we selected a random rock-phish domain and examined each of the IP addresses associated with the domain. We tallied each domain that also used one of these IP addresses and recursively checked these domain's associated IP addresses. In this manner we identified every IP address associated with rock-phish sites starting from just one address.

It should be noted that our methodology meant that we were rapidly aware of DNS changes, where domain names were mapped to new IP addresses. Because we tended to make all of our name lookups over a short period of time we often recorded many names resolving to the same IP address, and the next time we accessed the rock-phish site we would see most of them resolving to another address. Users would not see the same effect because of caching by DNS servers (usually at their ISP). This caching would mean that their perception would be of a constant mapping between name and IP address until the cache entry expired, when the site would 'move'. This caching effect also means that the removal of a domain name does not lead to the instant disappearance of the website, provided that the machine at the relevant IP address remains 'up'. When another ISP customer has resolved the name already, the site will remain visible at that ISP for an extended period, and will often be reachable via the 'removed' domain name for most of a day.

4.1 'Fast-flux' phishing domains

While we were collecting data for this paper the gang introduced a new system dubbed 'fast-flux' by the anti-phishing community, with trials in February and wider deployment from March onwards.[3] They arranged for their domains to resolve to a set of five IP addresses for a short period, then switched to another five. This of course 'eats up' many hundreds of IP addresses a week, but the agility makes it almost entirely impractical to 'take down' the hosting machines. The gang is likely to have large numbers of compromised machines available, since if they are not used to serve up phishing websites they are available for sending email spam. For further obfuscation, the gang changed from using the url-path to select the target bank to using the `Host:` header from the HTTP connection. This makes it somewhat more complex for ISPs and registrars to understand the nature of the sites and to what extent they can be considered to be 'live'.

4.2 Rock-phish statistics

We analysed rock-phishing sites during a period of eight weeks between February and April 2007. During this time, we collected 18 680 PhishTank reports which we categorised as rock-phish (52.6% of all PhishTank reports for the time period). While these reports are intended to be unique, we identified many duplicates due to the use of unique URLs as described above. This yielded a significant saving in effort,

[3]We were able to identify several machines that were used for both the original rock-phish scheme and for the new fast-flux architecture, so we are confident the same gang is involved. Further, although there are currently (May 2007) two fairly distinct pools of fast-flux machines being used for phishing, there are a handful of overlaps which indicate to us that one gang is operating both of them.

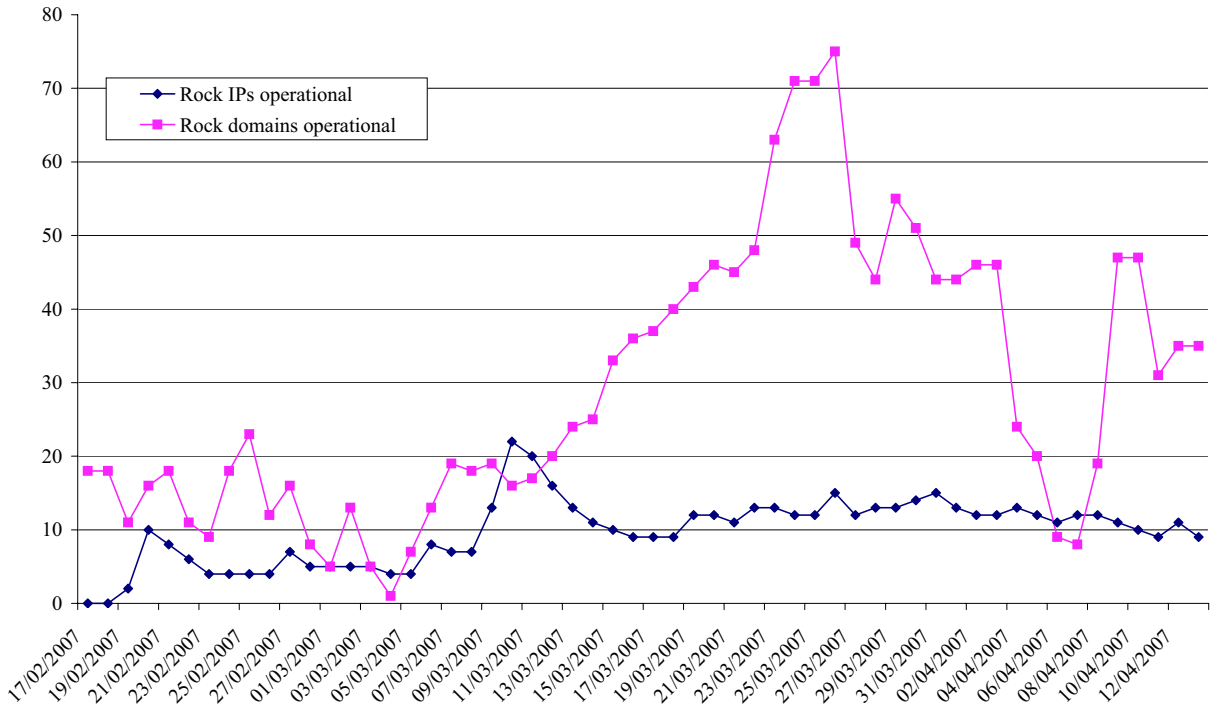


Figure 1: Rock-phish site activity per day.

since just 419 canonical rock-phish URLs were observed. Rock-phish sites used 122 IP addresses that were found to be operational for any duration. In all, the rock-phish sites impersonated 21 different banks and 3 other organisations.

Meanwhile, fast-flux sites triggered 1 803 PhishTank reports during the collection period. These reports pare down to 67 unique domains which resolve to 2 995 IP addresses. Observed fast-flux sites have targeted 18 banks and 10 other organisations.

Rock-phish sites continue to work for a particular domain that is mentioned in a spam email, provided that they can be resolved to at least one working IP address. Figure 1 tracks the average number of operational rock-phish domains and IP addresses on a daily basis. Sites or domains were removed constantly, but they were replenished frequently enough to keep a number of sites working every day. Only once, right at the start of our data collection period, did the sites fail to work entirely, because the IP addresses being used for DNS resolution all failed. Otherwise, between 1 and 75 domains and between 2 and 22 IP addresses were always available.

Notably, the number of operational domains steadily increased during the month of March, before falling steadily in late March and early April. This is primarily attributed to a large number of `.hk` domains bought from a single registrar, which was slow to remove the offending domains. But why would the rock-phish gang continue to buy new domains when their earlier ones still worked? One reason is that the domains may lose effectiveness over time as they are blocked by spam filters. Indeed, comparing the number of domains added per day to the number removed (see Figure 2-top) reveals only a weak correlation between domain addition following removal. This suggests the rock-phish gang are motivated to purchase new domains even when registrars are slow to take action.

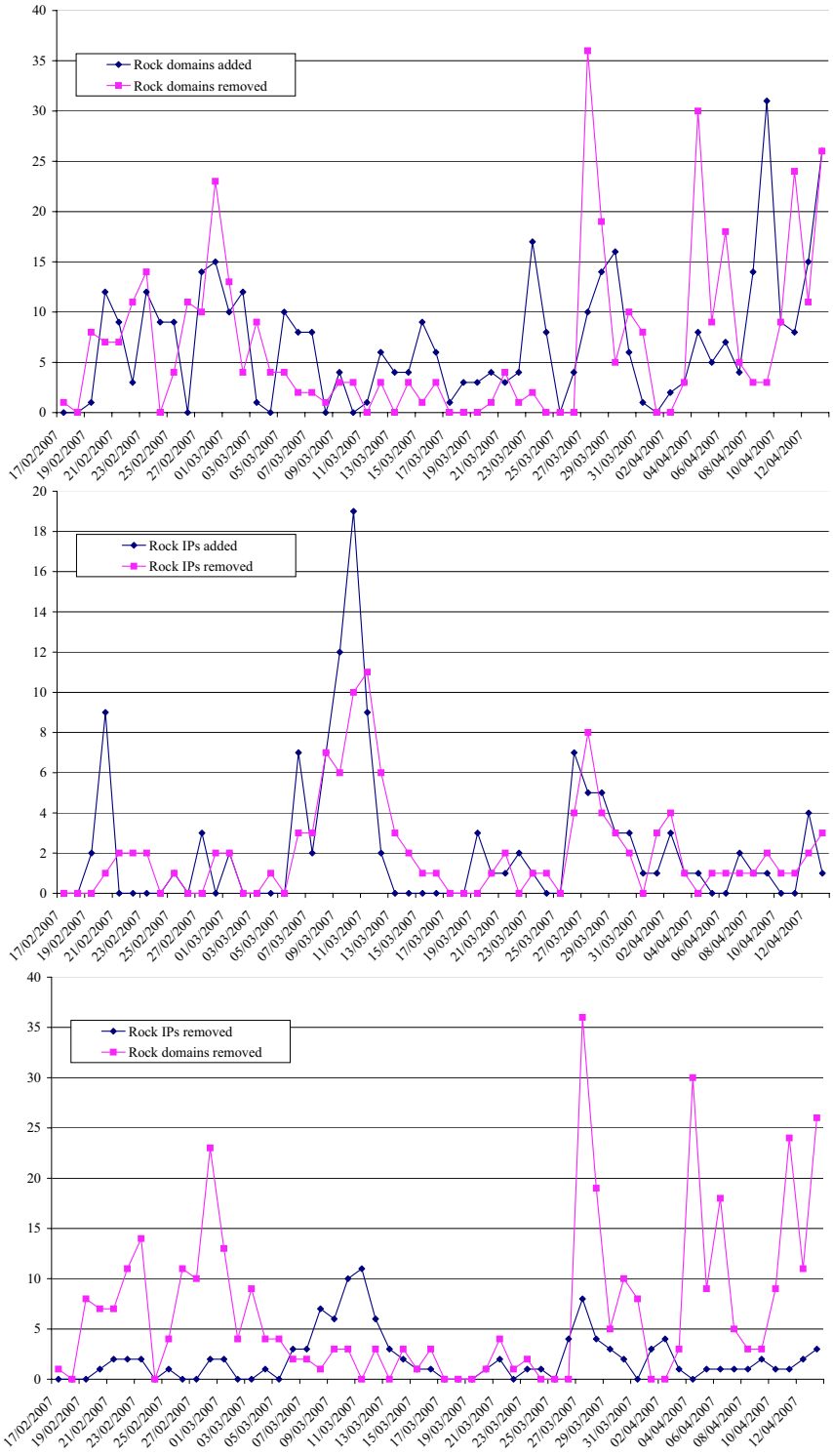
The story is rather different for the machines that rock-phish domains resolve to. Figure 2-middle plots the day-by-day addition and removal of compromised machines used. Here the correlation is strong: as soon as machines are removed, new ones replace them. The correlation coefficient of 0.738 implies that 54% of the total variance is explained by the correlation between adding and removing machines. Perhaps the rock-phish gang have automated IP replacement; automating domain acquisition, by contrast, is more difficult and costly – so it is not surprising that the data suggests that manual selection prevails when adding domains.

Finally, we can infer whether co-ordination between rock-phish domain and machine removal takes place by comparing daily takedown rates for both (Figure 2-bottom). There is almost no correlation between the number of domains removed on a given day and the number of machines removed. This suggests that very little co-operation between registrars and ISPs is taking place. Furthermore, the lack of correlation implies that either banks and other removal entities are not communicating convincingly to both ISPs and registrars, or they do not fully understand the rock-phish gang's use of domains and compromised machines.

5. WHO IS WINNING THE ARMS RACE?

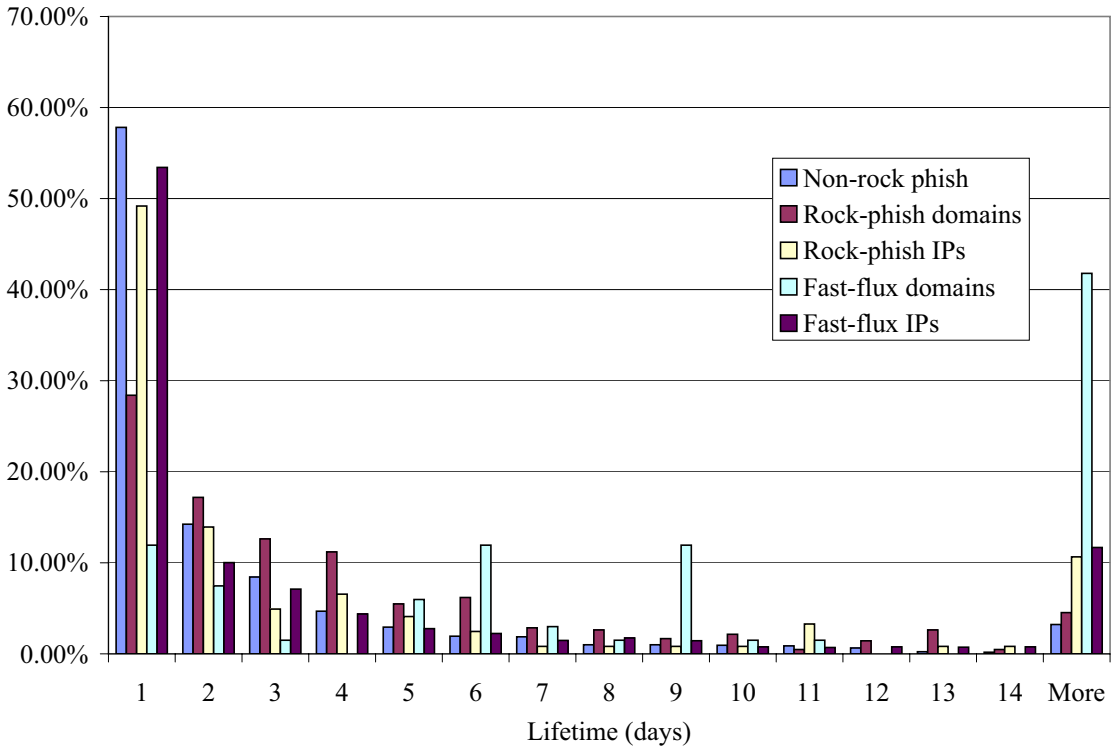
Phishing targets invest significant resources in removing phishing sites. In this section we present data on the duration of phishing sites and on user response to these sites to determine the effectiveness of the take-down strategy.

In addition to the collection of rock-phish sites, we also examined reports of regular phishing sites targeting a number of banks and other sites. From 15 030 reports gathered over the same 8-week period from February to April 2007, we



	Correlation coefficient r	r^2
Rock domains added–Rock domains removed	0.368	0.135
Rock IPs added–Rock IPs removed	0.738	0.544
Rock IPs removed–Rock domains removed	0.0629	0.00359

Figure 2: (Top) new and removed rock-phish domains per day; (Middle) new and removed rock-phish IPs per day; (Bottom) rock-phish domain and IP removal per day. Also included is a table of the respective correlation coefficients.



	Sites	Mean lifetime (hours)	Median lifetime (hours)
Non-rock	1 707	58.38	20
Rock domains	419	94.26	55
Rock IPs	122	124.9	25
Fast-flux domains	67	454.4	202
Fast-flux IPs	2 995	124.6	20

Figure 3: Histogram of phishing site lifetimes with table of sample size and mean and median lifetimes.

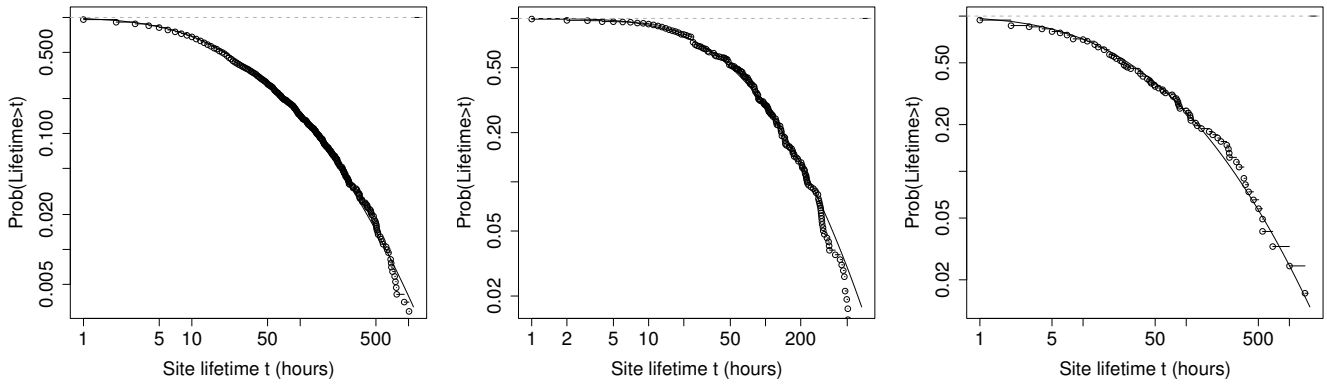
identified 1 707 unique non-rock-phish sites that were alive upon initial inspection. Because ordinary phishing sites do not follow a consistent pattern, establishing uniqueness is difficult. We considered two sites to be duplicates if they were hosted on the same domain, impersonate the same bank and were reported to PhishTank within two days of each other. However, removing duplicates does not account for the entire reduction of 14 062 reports. Many sites had already been removed by the time they have been verified and promulgated by PhishTank. Because we cannot evaluate whether dead-on-arrival-sites are in fact a phishing site or simply a malformed URL, we exclude them from our lifetime analysis. Thus, the lifetimes discussed below do not account for the many sites that are removed immediately.

5.1 Phishing site lifetimes

The site lifetimes for each type of phishing attack are given in the table in Figure 3. The mean lifetime of a normal phishing site is 58.38 hours, while for rock-phish domain the mean lifetime is 94.26 hours. Notably, for all phishing types, the median takedown time is much less than the average time. The reason why can be seen in the histogram of phishing site lifetimes in Figure 3. Each bin represents one day, and the histogram covers two weeks, which is long enough for most samples we collected (sites lasting longer are indicated by the 'More' column). 58% of non-rock-phish sites are removed within 24 hours of reporting, while the remainder do not survive much longer. Only 28% of non-rock-phish sites last more than 2 days, though notably the tail carries on for several weeks. For instance, the longest-lived ordinary phishing site from our sample stuck around for over six weeks!

For rock-phish sites, the distribution is slightly different. While clearly skewed toward shorter times, the distribution has a heavier tail: a small but substantial number of rock-phish domains remain operational for longer periods. 28% are removed on the first day, 17% on the second, and 55% remain for 3 days or longer.

The slightly longer survival time of rock-phish sites may be partially explained by the persistence of usable hosting machines (see the final histogram in Figure 3). Recall that rock-phish spam always uses a domain name in the linked URL. This allows the gang to cycle through IP addresses as they fail. Several rock-phish domains resolve to the same IP address at any given time; when the machine is removed, they switch to another IP address in their pool. Figure 3 suggests that they do not have to switch all that often: IP addresses work for an average of 124.9 hours. While many are removed within one day, some remain for months before being removed.



	Lognormal				Kolmogorov-Smirnov	
	μ	Std err.	σ	Std err.	D	p-value
Non-rock	3.009	0.03553	1.468	0.02512	0.03878	0.1996
Rock domains	3.924	0.05942	1.216	0.04202	0.06426	0.4158
Rock IPs	3.314	0.1656	1.829	0.1171	0.08945	0.7007

Figure 4: Cumulative probability distributions with lognormal curve fit: non-rock-phish lifetimes with $\mu = 3.01$,$\sigma = 1.47$ fit (Left); rock-phish domain lifetimes with $\mu = 3.92$, $\sigma = 1.22$ fit (Centre); rock-phish IP lifetimes with $\mu = 3.31$, $\sigma = 0.166$ fit (Right).

Another explanation for the longer lifetimes of rock-phish sites is that their attack method is not widely understood, leading to sluggish responses. Splitting up the components of the phishing attack (domains, compromised machines and hosting servers) obfuscates the phishing behaviour so that each individual decision maker (the domain registrar, ISP system administrator) cannot recognise the nature of the attack as easily when an impersonated domain name is used (e.g., `barclaysbankk.com`), or HTML for a bank site is found in a hidden sub-directory on a hijacked machine.

Fast-flux sites exhibit markedly different behaviour. Domains last much longer: nearly 19 days on average, and there are half as many as are used for rock-phish. This may be largely attributed to the lack of any response by a few of the targets, whose good name has been borrowed to assist in recruiting mules. Interestingly, the average lifetime of fast-flux IP addresses (124.6 hours) is almost identical to the IPs used for rock-phish attacks (124.9 hours). Hence the practice of burning through many IP addresses does not yield any longer lifetime for the fast-flux hosting machines.

The skewed distribution of site lifetimes shows that while most sites are removed promptly, a substantial number remain for a very long time. These long-lived sites cause the average lifetime to be much longer than the median lifetime. We have managed to fit some of the takedown data to match the lognormal probability distribution. To do so, we first estimated the parameters μ and σ which specify the distribution using maximum likelihood estimation. To test the fit, we computed the Kolmogorov-Smirnov test 1000 times to compute the average maximum difference D between the model and data.

The lognormal distribution turns out to be a good fit for the distribution of ordinary phishing sites as well as rock-phish domains and IP address lifetimes. However, it is not a good fit for fast-flux sites. The table in Figure 4 gives the relevant attributes for each fitted distribution, and the plot show the lognormal cumulative probability distributions and the observed data points. Note that both axes are logarithmic in scale to demonstrate the goodness-of-fit in the tail of the distribution. It is significant that the takedown times for these three different categories of phishing attack can each be modelled by the same family of heavy-tailed distribution, particularly since the actors responsible for the takedown are different (domain registrars, ISPs and system administrators).

5.2 User responses to phishing

Having established how long phishing sites remain operational, we now estimate user-response rates to phishing sites. We analysed the site usage statistics from 83 phishing sites, from which we obtained daily snapshots of hit rates broken down according to URLs. From this list of popular URLs, we identified the phishing entry and completion pages and cross-referenced its PhishTank report to establish the earliest report date. Note that these were all ordinary phishing sites; the rock-phish gang do not leave logging data visible.

Webalizer also provides a rank ordering of entry pages. An entry page is the first one that a site visitor views. By tracking entry pages, we can readily distinguish between hits to the phishing page and the rest of the site. Each time we discovered a live site publishing Webalizer reports, we automatically returned daily to obtain updated reports until the site was taken offline. Thus, we ended up with a time sequence of reports used to estimate the distribution of victim responses for the days surrounding the phishing report.[4]

For most phishing scams, when someone enters their details on the site, they are taken to a fake confirmation page. We picked out these confirmation pages and noted the num-

[4]Our system was not alone in visiting these websites to determine if they were still operational. We took steps to exclude these automated monitors from our datasets.

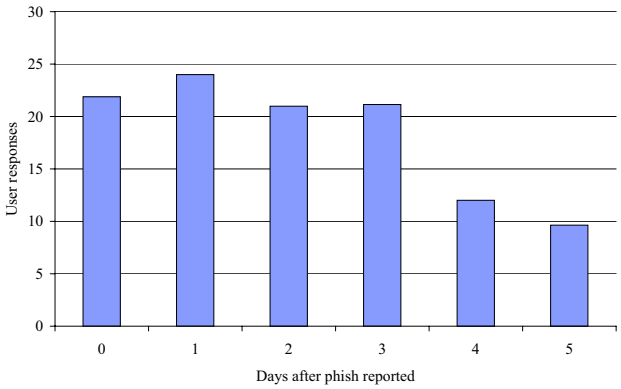


Figure 5: User responses to phishing sites over time. Data includes specious responses.

ber of hits they received, which was small compared with the number of hits on the initial page. Regrettably, Webalizer does not record the number of unique visits for all URLs, so we could seldom obtain the number of unique visits to the more popular entry pages. Instead, we estimated the number of unique visits to each site's confirmation page by taking the number of hits, and assuming the same fraction of visits to hits that we saw for the entry page.

Unfortunately, from the point of view of collecting good data, in many cases the site statistics presented difficulties: we could only obtain one reading before the site was removed, it could be unclear which were the confirmation pages, or the Webalizer values were not fetched until several days after the site went live. For these reasons, we were only able to obtain usable day-by-day statistics from twenty sites. An average of these results is given in Figure 5.

We estimate that 22 unique users continue through to the confirmation page on the same day that the phish is reported. On the next day, another 24 responses are expected. Somewhat surprisingly, the user responses continue at a fairly high level until the site is removed. We cannot say whether this is caused by ongoing spamming activity, or by users catching up with email backlogs in their in-boxes. This ongoing activity was demonstrated to an extreme by the usage statistics for a PayPal phishing site loaded onto a web page for the Niger Water Basin Authority. This site remained alive into March 2007 and received a steady stream of phishing responses over a month and a half, so the failure to take it down more quickly caused ongoing problems. Thus it does appear that take-down, even when it is slow, is always going to have some positive effects.

We also note that there is some noticeable variation in the number of responses received. One site (excluded from the average presented in Figure 5 because of missing data) drew over 500 responses in one day. Hence a small number of sites may draw significantly larger numbers, so the data presented here should be viewed as a conservative estimate.

But how accurate is the confirmation rate as a measure of successful attack? Just because the confirmation page is visited, this does not necessarily mean that every hit corresponds to a theft of personal details. To arrive at a more accurate success rate, we have also gathered 414 user responses with personal information published on phishing sites in what the attacker believed to be an obscure location. We examined each response by hand to determine whether the responses appeared plausible. Many responses were obviously fake, with names and addresses like 'Die Spammer' and '123 Do you think I am Stupid Street'. In fact, the responses were evenly split: 214 responses were obviously fake, while 200 appeared real. Hence, albeit from a small sample, we can estimate that half the responses to a phishing site represent actual theft of details.

So how does this user-response data relate to the phishing site lifetimes we described in Section 5.1? Of the sites we sampled, we might expect around 25 victims per site if they are removed within one day of reporting, and rising by 10 victims for each successive day. This is a substantial number, and it is unclear whether the phishing targets can act sufficiently quickly to reduce it by very much.

5.3 Estimating the cost of phishing attacks

We can now use our empirical data to estimate the cost imposed by phishing attacks. We must of course qualify our calculations by noting that we are using a number of rather fuzzy estimates, so that substantial refinement will be possible in the future as better figures come to light.

We first consider the cost imposed by ordinary (i.e., not rock-phish or fast-flux) phishing sites. We collected data for eight weeks and confirmed 1 448 banking phishing sites (we exclude eBay phishing scams for the purpose of this calculation). Extrapolating, we might expect 9 437 sites per year. These particular sites remain operational for around 57 hours on average, which yields approximately 33 victims based on the analysis in Section 5.2. Gartner has estimated the cost of identity theft to be $572 per victim [4].[5] Hence, the estimated annual loss due to ordinary phishing sites is $9\,437 * 33 = 311\,449$ victims $* \$572 = \$178.1$m. Gartner estimates that 3.5 million Americans give away their details annually, which leads to an estimated loss of $2bn.

We cannot reliably provide an estimate for the costs of rock-phish and fast-flux phishing scams since we do not have similar response data. However, given that the rock-phish gang send a large proportion of all spam [6], which drives visitor numbers, it is fair to assume that they steal at least as much money as ordinary phishers. Thus, we estimate, at an absolute minimum, that at least $350m is lost annually due to phishing scams. The disparity with Gartner's total of $2bn is doubtless due to the extremely rough approximations used, both by ourselves and Gartner. But the difference will also be accounted for by the other ways in which personal data can be stolen, for example the theft of merchant databases, and the activities of malware that scans files or operates keyloggers.

6. DISCUSSION

6.1 Do weekends affect take-down?

Defenders working for targets of phishing attacks often speculate that attackers deliberately wait to advertise phishing sites until just before the weekend to maximise site uptime, since many system administrators will be away. Upon examining the data, we find that sites launched before the weekend are no more likely to last longer.

[5]Gartner also gives a value of $1 244 per victim, but reports that over half of this is subsequently recovered.

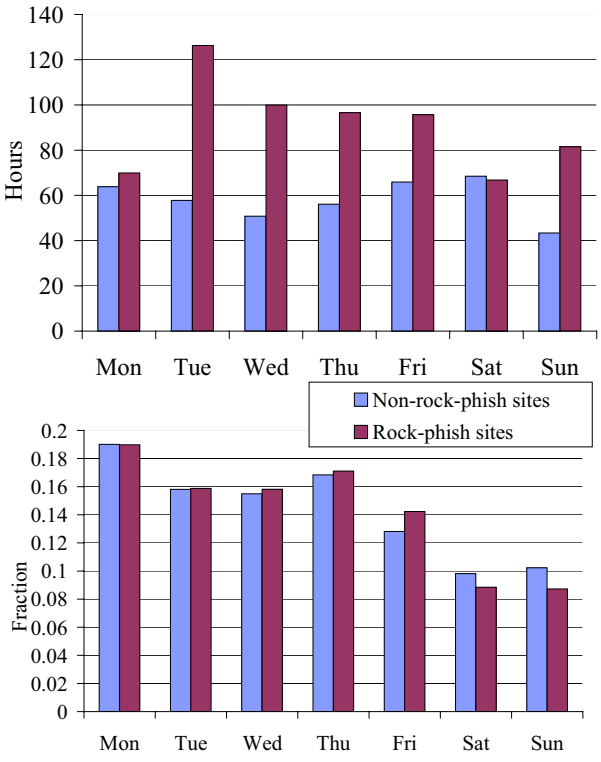


Figure 6: Phishing-site lifetimes (Top) and counts (Bottom), collated by weekday they were first reported.

We first examine whether sites reported near the weekend stay around longer than those reported earlier in the week. The upper graph in Figure 6 shows the average duration of phishing sites based upon the day of the week the site was first reported. Rock-phish sites reported on Tuesday last longest, while those reported on Monday and Saturday are removed quickest. It is unclear whether there is any significance to these differences. Non-rock-phish sites launched on Saturday last around one day longer than those reported on Sunday, so it seems as if reports from both Saturday and Sunday are actioned at much the same time.

The next question we address is whether some days are more popular for launching phishing sites than others. The lower graph in Figure 6 measures the fraction of sites reported on each day of the week. The most striking conclusion to be drawn from this graph is that the weekend is the least popular time for both rock-phish and ordinary phishermen to set up sites. More accurately, fewer *reports* of new phishing sites are created over the weekend. It is impossible to tell whether there are fewer sites appearing, or fewer people looking for them, on Saturday and Sunday.

6.2 Comparing bank performance

There are 120 banks and other institutions targeted from our sample of ordinary phishing sites. However, some banks are targeted a lot more than others: PayPal was impersonated by 400 of the 1 707 sites, while 50 banks were only spoofed once. A pie chart showing the proportion of targeted banks is given in Figure 7.

While banks cannot control the number of fake sites that

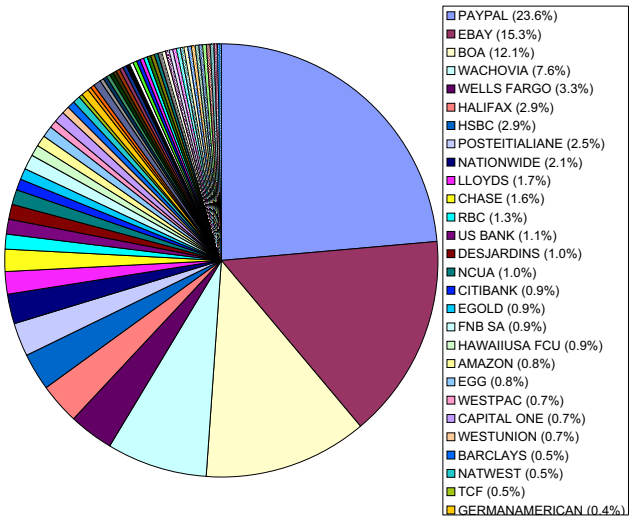


Figure 7: Number of phishing sites per bank (only banks with seven or more sites in the legend).

appear, they certainly can help determine how long they stick around. Here there is also significant disparity. Figure 8 presents a rank-ordering of the average site lifetimes for banks impersonated more than five times during the sample period. TCF Bank, eGold and Citibank are slowest at taking down sites (over four days), while Capital One, NatWest and Flagstar Bank are quickest (around twelve hours). Note that the results should be treated with caution because the differences will, at least in part, result from different choices by the attackers as to where their sites are hosted.

6.3 Comparing free-hosting performance

We identified a number of providers of 'free' webspace that regularly hosted phishing websites. Two organisations had been particularly exploited: `yahoo.com` based in the US and `pochta.ru` sited in Russia. Yahoo! had hosted 59 phishing websites on their free 'GeoCities' webspace during the period we analysed, attacking a wide range of different institutions. The Russian site (operating a number of domain names, but all providing free space) had hosted 65 eBay sites, and one each of PayPal and Posteitaliane. As can been seen from the table, there was a wide disparity in take-down times between the two organisations, with Yahoo! being faster than one would generally expect for any institution, and the Russian sites (about one third of all eBay sites) being removed at a similar speed to eBay sites hosted elsewhere.

	Sites	Mean lifetime	Median lifetime
`yahoo.com`	59	11.27 hours	5 hours
`pochta.ru`	67	82.24 hours	31 hours

6.4 Collusion dividend for the rock-phish gang

Collusion has enabled the rock-phish gang to pool its resources to its advantage. First, co-operation has strengthened its defence by swapping between compromised machines as they are removed by ISPs. Second, the gang can impersonate many banks on each domain.

Such overt co-operation creates additional risks, however. Notably, collusion increases the site's value as a take-down

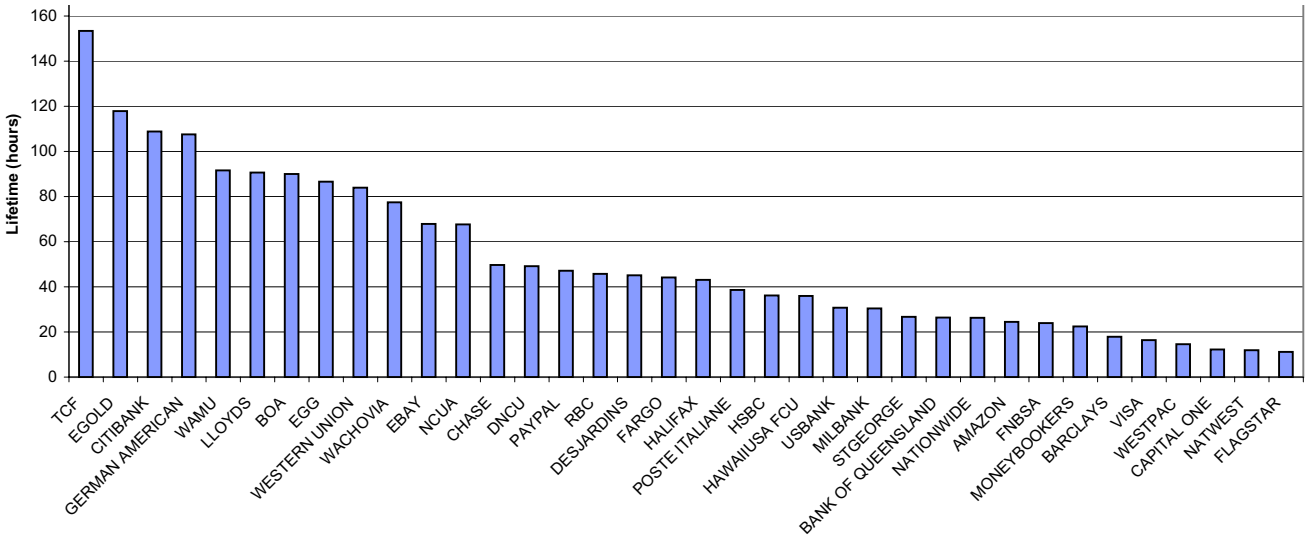


Figure 8: Phishing-site lifetimes per bank (only banks with five or more sites are presented).

target. All the banks whose sites are present on the rock-phish servers ought to be motivated to remove a site, not just one bank as for regular phishing sites. The effectiveness of phishing defence will be the sum of the banks' efforts, so if they are fully co-operating, then one might expect faster take-down times. However, we were told (off the record) that banks tend not to worry about rock-phish sites until their brand is mentioned in spam emails. It is also possible that some of the banks targeted by rock-phish sites are not co-operating at all, but are instead free-riding on the efforts of a few more capable organisations [16]. Given the longer take-down times for rock-phish sites, it appears that currently the benefits to the gang from collusion outweigh the costs – at the present level of co-operation by the defenders.

6.5 DNS trade-offs

When phishing first became widespread it was commonplace to see fake domain names which were minor variations on the real site's identity. This is now rather less common. One of the reasons for this will be that it gives the defenders the option of getting either the site removed or having the domain name suspended – with the latter being rather simpler since it requires co-operation by relatively 'clued-up' registrars who are already experienced in dealing with the branding implications of too-similar domain names; rather than seeking help from ISPs who might not be very familiar with phishing attacks.

The rock-phish gang use nondescript domain names and avoid this issue of branding, leaving the registrar with the problem of breaking their contract for the supply of the name on the word of a third-party who claims that it is being used for phishing. That registrars are now prepared to suspend the names is apparent from our data – though it is interesting to note that at present no systematic attempt is being made to suspend the names that are being used for the DNS servers associated with the rock-phish domains. This is despite these names being created solely for the purpose of providing an indirection for the DNS servers used to resolve the rock-phish URLs. The argument that these too are entirely fraudulent is not yet won – though as can be seen from Figure 1, when the rock-phish DNS system is disrupted the effect can be dramatic. Of course, when these name service names are regularly suspended the gang will use absolute IP addresses to locate their DNS servers, thereby continuing to operate, albeit with slightly less flexibility.

The final trade-off of note that relates to DNS is the caching mentioned in Section 4. Setting a high value for 'time-to-live' will ensure that domain names may be resolved, particularly at larger ISPs, for some time after the domain is suspended by a registrar. However, lower values offer more agility as compromised machines are reclaimed by their owners.

6.6 Countermeasures

So if take-down strategies are not completely mitigating phishing attacks, what else can be done?

One important advance would be to reduce the information asymmetry for the defenders. Phishers obfuscate their behaviour and make sites appear independent and thereby phishing appears to many to be an intractable problem. Security vendors are happy to accept inflated statistics to make the problem seem more important. Law enforcement will not prioritise investigations if there appear to be hundreds of small-scale phishing attacks, whereas their response would be different if there were just a handful of people involved. Hence, improving the measurement systems, and better identifying patterns of similar behaviour, will give defenders the opportunity to focus their response upon a smaller number of unique phishing gangs.

Other entirely obvious countermeasures include reducing the availability of compromised machines, rate-limiting domain registration, dissuading users from visiting the sites, and reducing the damage that disclosing private information can do. Unfortunately, these strategies, are either infeasible or are being attempted with limited impact so far. What does seem to be working, at least to some extent, is for the sites that are attacked to improve their back-office controls. The incentives to go phishing are much reduced if miscre-

ants cannot use the account numbers and passwords they steal to transfer money out of accounts; or if they cannot get money out of the banking system in such a manner that the transfers cannot be clawed back.

7. CONCLUSION

In this paper we have empirically measured phishing site lifetimes and user response rates, to better understand the impact of the take-down strategies being employed by institutions targeted by phishing attacks. While take-down certainly hastens the fraudsters' movement from one compromised site to another, many users continue to fall victim. Furthermore, the data reveals that sophisticated attackers can extend site lifetimes. Indeed, the rock-phish gang has already demonstrated techniques for adapting to regular removal. They have clearly invented (or stumbled upon) a relatively successful formula, and with 'fast-flux' are experimenting with another, but it is far from clear that the defenders currently understand precisely what the mechanisms are, and how best to disrupt them.

Removing phishing websites is often perceived of as a sisphyean task, but our analysis shows that even when it is done slowly, it does reduce the damage that is done. We have also demonstrated wide disparities in reaction time between comparable organisations. We have shown that these disparities extend across borders, some banks work faster than others and some webhosting companies do a better job at removing sites. Improving the transparency of attacker strategy and defender performance is key to reducing the success of phishing scams.

There is still much work to be done to better understand attack behaviour and the extent to which defenders are pulling their weight. Much more analysis can be carried out on the data we are collecting to show how well they are doing. For instance, we could compare site lifetimes categorised by hosting country in order to estimate the externality impact different countries impose on others. We would also like to study how size and perceived security practices impact the way in which attackers select particular organisations as targets, it may be that a brief display of competence will send the attackers to another target, much as burglar alarms protect you and not your neighbours.

Acknowledgements

Removed for anonymous submission.

8. REFERENCES

[1] L. Jean Camp: Reliable, Usable Signaling to Defeat Masquerade Attacks. The Fifth Workshop on the Economics of Information Security (WEIS 2006), 2006.

[2] Ionut Alex. Chitu: Google Redirect Notice. 16 Feb 2007. http://googlesystem.blogspot.com/2007/02/google-redirect-notice.html

[3] Christine E. Drake, Jonathan J. Oliver, and Eugene J. Koontz: Anatomy of a Phishing Email. First Conference on Email and Anti-Spam (CEAS), Mountain View, CA, USA, 2–3 Aug 2004.

[4] Gartner Inc: Gartner Says Number of Phishing E-Mails Sent to U.S. Adults Nearly Doubles in Just Two Years, Press Release, 9 Nov 2006. http://www.gartner.com/it/page.jsp?id=498245

[5] Markus Jakobsson and Steven Myers (Eds.): Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft. Wiley, Nov 2006, ISBN: 978-0-471-78245-2.

[6] Robert McMillan: 'Rock Phish' blamed for surge in phishing. InfoWorld, 12 Dec 2006 http://www.infoworld.com/article/06/12/12/HNrockphish_1.html

[7] Microsoft Inc.: Phishing Filter: Help protect yourself from online scams. 28 Oct 2006. http://www.microsoft.com/athome/security/online/phishing_filter.mspx

[8] Daisuke Miyamoto, Hiroaki Hazeyama, and Youki Kadobayashi: SPS: a simple filtering algorithm to thwart phishing attacks. Asian Internet Engineering Conference (AINTEC), 13–15 Dec 2005.

[9] Mozilla Corp.: Phishing Protection. 2006. http://www.mozilla.com/en-US/firefox/phishing-protection/

[10] Ying Pan and Xuhua Ding: Anomaly Based Web Phishing Page Detection. 22nd Annual Computer Security Applications Conference (ACSAC'06), IEEE, 2006, pp. 381–392.

[11] PhishTank: http://www.phishtank.com/

[12] John S. Quarterman: PhishScope: Tracking Phish Server Clusters. Digital Forensics Practice, 1(2), 2006, pp: 103–114.

[13] Blake Ross, Collin Jackson, Nick Miyake, Dan Boneh and John C Mitchell: Stronger Password Authentication Using Browser Extensions. 14th USENIX Security Symposium, 2005.

[14] Stuart E. Schechter, Rachna Dhamija, Andy Ozment and Ian Fischer: The Emperor's New Security Indicators: An evaluation of website authentication and the effect of role playing on usability studies. 2007 IEEE Symposium on Security and Privacy.

[15] Rob Thomas and Jerry Martin: The underground economy: priceless. USENIX ;login: **31**(6), Dec 2006.

[16] Hal Varian: System Reliability and Free Riding. In Economics of Information Security, L. J. Camp, S. Lewis, eds. (Kluwer Academic Publishers, 2004), vol. 12 of Advances in Information Security, pp. 1–15.

[17] Webalizer: http://www.mrunix.net/webalizer/

[18] Liu Wenyin, Guanglin Huang, Liu Xiaoyue, Zhang Min and Xiaotie Deng: Detection of Phishing Webpages based on Visual Similarity. Proc. 14th International World Wide Web Conference, ACM Press, 2005, pp. 1060–1061.

[19] Min Wu, Robert C. Miller and Simson L. Garfinkel: Do security toolbars actually prevent phishing attacks? Proceedings of the SIGCHI conference on Human Factors in computing systems (CHI'06), ACM Press, 2006, pp. 601–610.

[20] Yue Zhang, Serge Egelman, Lorrie Cranor, and Jason Hong: Phinding Phish: Evaluating Anti-Phishing Tools. In Proceedings of the 14th Annual Network & Distributed System Security Symposium (NDSS 2007), 2007.