The consequence of non-cooperation in the fight against phishing

Tyler Moore
Center for Research on Computation and Society
Harvard University
33 Oxford Street
Cambridge, MA 02138, USA
Email: tmoore@seas.harvard.edu

Richard Clayton
Computer Laboratory
University of Cambridge
15 JJ Thomson Avenue
Cambridge CB3 0FD, UK
Email: richard.clayton@cl.cam.ac.uk

***Abstract*—A key way in which banks mitigate the effects of phishing is to have fraudulent websites removed or abusive domain names suspended. This 'take-down' is often subcontracted to specialist companies. We analyse six months of 'feeds' of phishing website URLs from multiple sources, including two such companies. We demonstrate that in each case huge numbers of websites may be known to others, but the company with the take-down contract remains unaware of them, or only belatedly learns that they exist. We monitored all of the websites to determine when they were removed and calculate the resultant increase in lifetimes from the take-down company not knowing that they should act. The results categorically demonstrate that significant amounts of money are being put at risk by the failure to share proprietary feeds of URLs. We analyse the incentives that prevent data sharing by take-down companies, contrasting this with the anti-virus industry – where sharing prevails – and with schemes for purchasing vulnerability information, where information about attacks is kept proprietary. We conclude by recommending that the defenders of phishing attacks start co-operatively sharing all of their data about phishing URLs with each other.**

I. Introduction

Many security problems are dealt with by several distinct groups of people, who are taking individual action against a common threat. It is often the case that co-operation could improve their overall effectiveness, but at the same time, some of the groups may see co-operation as unfair because they would contribute more than others; or they may fear that their ability to sell specialist services is diminished; or the effect may be that the incentives for them to invest in improved techniques is much reduced. The balance between the positives and the negatives, and hence who has an incentive to change tactics, will determine whether co-operation occurs naturally, or whether it could only be imposed by an outside force.

In this paper we examine the phishing website take-down industry, where there is some measure of co-operation already. However, we can experimentally measure the impact of a lack of co-operation in sharing information about attacks, and give a robust estimate of the consequent increase in the risk of financial loss. The lessons we draw, and the conclusions we reach about how to foster more co-operation, go much wider than phishing. Many pressing threats to information security, from botnets to malware, are presently countered by piecing together disparate, incomplete, data sources. Defenders should instead arrange to work more closely together in tackling common threats.

A. Phishing website take-down

Phishing is the criminal activity of enticing people into visiting websites that impersonate the real thing, to dupe them into revealing passwords and other credentials, which will later be used for fraudulent activities. Although a wide range of companies are attacked in this way, from domain registrars, through auction sites and multi-user games to online merchants, the vast majority of attacks are against financial institutions. Hence for simplicity, within this paper we will use the term 'banks' for the firms being attacked.

One of the key countermeasures to phishing is the prompt removal of the imitation bank websites. The removal may be achieved by removing the web pages from the hosting machine, or in complex schemes where page requests are relayed by ever-changing armies of proxy machines, it may require that a registrar suspends a domain name from the DNS so it can no longer be resolved.

Although a small number of banks deal with phishing website take-down exclusively 'in-house', the majority hire specialist companies to remove either all of the sites they care about, or sometimes just the 'hard' cases. The 'take-down companies' that remove phishing websites are usually one arm of more generic 'brand-protection' companies that deal with counterfeiting and other intellectual property issues.

Whichever firm removes a particular website, quite clearly the process cannot start until the existence of the website becomes known. At present, almost all phishing attacks start with the sending of spoof emails. These emails contain links to the fraudulent websites, disguised as legitimate links to the genuine bank. There are other ways in which phishing can be done, such as pharming [1], and the links to the website may proceed by various indirections, but at present these attacks remain the exception.[1]

[1] It should also be noted that in some countries phishing is almost unknown and the main attack vector is the use of keyloggers or malware that hijacks banking sessions and converts legitimate transactions into fraudulent transfers of money. These attacks do not generally involve websites that impersonate banks, so we do not consider them in this paper.

This pre-proceedings version: October 9, 2008
3rd APWG eCrime Researchers Summit, Oct. 15–16, 2008, Atlanta, GA, USA

Therefore, for timely removal of the website, it is essential that banks rapidly become aware of the URLs advertised in phishing emails. Some of these URLs will be reported directly to the bank by customers who were not misled, and some will be identified because the emails are undeliverable and delivery failure reports are received by the bank. URLs identified in these two ways are of direct interest only to the particular bank, but there are other sources of information that are rather less specific.

Phishing emails are sent out indiscriminately, so many will be detected by random members of the public who take the initiative to report them. Several websites exist for making reports, operated by organisations such as CastleCops[2] and the Anti-Phishing Working Group (APWG)[3]. In addition, many of the browser 'toolbars' that block access to phishing websites incorporate reporting mechanisms which deliver URLs to a central clearinghouse. Also, almost all of the URLs that are in circulation will be detected by the anti-spam firms who are attempting to keep unsolicited email from reaching their clients. These firms collate the URLs found in the email spam and submit them to groups such as the APWG.

The data from these clearinghouses could be passed on directly to the appropriate bank, but they are usually distributed as a generic 'feed' of suspicious URLs. These feeds are examined by the banks, and more particularly by the take-down companies, to determine if the reports are accurate and hence that there are websites that need to be removed.

The take-down companies negotiate feeds of raw URLs from these clearinghouses, from individual ISPs, and from as many other sources they can manage. Additionally, they process their own incoming email spam and check what is arriving at any dormant domains that they own. The companies thereby create their own feed, which will generally have the desirable property that the URLs have been validated and the false positives removed. The feed will be used internally for their own take-down activities, but will also be supplied to client banks who wish to do their own take-downs, or who just wish to be informed as to the current levels of attack against their brands. The feeds are also occasionally sold to ISPs or domain name registrars, who wish to proactively police their part of the Internet.

B. Outline and contribution of this paper

In the course of our research into phishing activities, we have obtained proprietary feeds from two separate take-down companies. We have combined this data with four other feeds, from an owner of several major Internet brands, from the APWG, and from two public domain sources.

In earlier work [2] we examined phishing website lifetimes, and showed that prompt take-down of phishing sites made a significant contribution to reducing the number of stolen credentials. We measured average takedown times of 62 hours for compromised websites (and 95 hours for rock-phish domains). These times are significantly longer than is quoted by the take-down companies, who claim "less than 5 hours" [3] or "under 6 hours" for domestic providers and "less than 24 hours" for international sites [4]. This might be selective memory – the companies remembered the run-of-the-mill work they did, without realising quite how many sites hung around for many days.

This paper offers an alternative explanation. Much of the difference in take-down times can be ascribed to the companies simply being unaware of some of the websites that we were tracking. Given the multiple feeds available to us in this present study, we can demonstrate that we are aware of many more phishing websites than any one company on its own. Consequently, we are in an excellent position to test the conjecture that take-down times are unnecessarily long because of a lack of information sharing. Indeed, our research clearly demonstrates that non-cooperation among phishing defenders significantly slows the take-down process.

In Section II we set out the details of the feeds of URLs that we process and the way in which we collect data about website lifetimes. In Section III we compare the feeds of two take-down companies and unequivocally show that take-down times for phishing websites would be shorter if the two companies were to share the information in their feeds with each other. In Section IV we consider websites that are operated by the rock-phish gang, who use an innovative architecture that allows them to attack multiple banks in parallel. Taking down rock-phish websites is inherently co-operative because the process is already a 'sum of efforts'. Nevertheless, we show that once again there is a gain to be made from sharing feeds. In Section V we consider the incentives which prevent information sharing by take-down companies, while drawing parallels with other realms of computer security within which sharing is, or is not, the norm. We then make a clear recommendation for change within the anti-phishing community. In Section VI we survey related work on the pros and cons of information sharing and finally in Section VII we draw overall conclusions.

II. Data collection methodology

We have obtained a number of feeds of phishing website URLs. These include two volunteer organisations, 'PhishTank'[4] which specialises in the URLs of phishing websites, and 'Artists Against 419'[5] which mainly deals with sites that facilitate auction scams or complex advanced fee fraud conspiracies. We take a feed from a brand owner, which consists almost exclusively of URLs for websites attacking their company. Most significantly for the current work, we receive feeds from two brand protection companies who offer specialist phishing website take-down services. These companies amalgamate feeds from numerous other sources, including some we receive directly, and combine them with data from proprietary phishing email monitoring systems. Finally, we take the 'industry' feed that is collated from numerous sources by the APWG.

[2] http://www.castlecops.com/pirt

[3] http://www.antiphishing.org/

[4] http://www.phishtank.com/

[5] http://www.aa419.org/

Although by their nature all these feeds have substantial overlaps with each other, in practice each contains a number of URLs that we do not receive from any other source. The overall result is that we believe that our database of URLs is one of the most comprehensive available, and the overwhelming majority of phishing websites will come to our attention. In principle, we could use capture-recapture analysis to estimate what proportion of sites we were unaware of, as attempted by Weaver and Collins [5]. However, the lack of independence between the various feeds makes a robust estimate of coverage impractical to achieve.

In all cases except PhishTank and APWG, the URLs are passed to us after they have been determined to be fraudulent sites. We deem the sites to have been 'up' at the time at which this determination was made or, if we are told when the site was initially reported, we use that (earlier) time instead. In practice, the take-down companies process URLs almost as fast as they are received.

In the case of PhishTank, we obtain the URL before the site has been validated as fraudulent – volunteers vote to determine this, sometimes triggering significant delays [6]. For our purposes, we simply use the time of the first appearance on the PhishTank website as the earliest time that the site is known to have existed, and so we are unaffected by how long voting decisions may take. For the APWG feed, we can use the time that the site appears in the feed and, because this data also flows to the take-down companies and to PhishTank, we can use their combined opinion to determine if it is a valid report.

A. Monitoring phishing websites

We run our own monitoring system which checks the websites reported to us for any changes to their content. Each phishing webpage is accessed some 15 to 20 times per day to determine what content is being served. A website that returns a '404' error is removed from testing, but other failures are retried for several days to ensure that minor outages are not mistaken for permanent removal.

We deem a phishing website to be 'up' while the content is unchanged, but any significant change (beyond session identifiers, cookies, etc.) is treated as a removal. In practice, the phishing attackers create fraudulent sites from 'kits' and do not change the pages once they are placed on a compromised machine – indeed they seldom change the kits from one week to the next, as they host the pages on a series of different hosts. The only manual step we apply is to ensure that the site has not been taken down and replaced with an innocuous page before we first monitor it. We measure the website lifetime from the earliest time we knew it was 'up' until the time of the last access to the fraudulent content.

Often multiple URLs refer to the same site, either because they turn up in different feeds or because we have canonicalised two URLs into the same basic format: we remove specious parameters, fix insignificant case changes, and convert non-standard representations of IP addresses from hex, octal etc. into a standard dotted quad format. We treat such multiple URLs as equivalent – and go further by removing the last component of the path as well. This solves two problems. First, it avoids double-counting equivalent URLs where one ends in a / and the other has an 'index.html' appended. Second, it overcomes the propensity of some feeds to include not only the initial webpage mentioned in the phishing email, but also the URLs of secondary pages that are encountered if credentials are filled in on the website. Wherever we have multiple URLs subsumed into one generic version, we determine the earliest of all the 'up' times and use that for our calculations.

False positives – where a non-phishing URL gets placed in one or more of the feeds – are rare, but must nonetheless be addressed. While not quantified, our qualitative impression is that the feeds from take-down companies contain very few false positives, especially for websites that impersonate their own customers. This is because the take-down companies carefully inspect suspected phishing websites before attempting to remove them. Ultimately, these companies stand behind the veracity of their feeds. By comparison, the feeds from APWG and PhishTank may include slightly more false positives.[6]

Nonetheless, we have taken additional steps to avoid false positives in computing our results. For each prospective phishing website, we search through the collected HTML for the name of the bank being targeted. For instance, we search for the phrase 'Bank of America' in the HTML of websites suspected of impersonating Bank of America. Consequently, our analysis only considers websites where a match has been found. The only exception occurs whenever the site has already been removed before we can visit to collect its published HTML. In this case, we trust the assessment of the take-down companies and assign a zero-lifetime to the sites which appear in their feeds. If the site only appears in PhishTank or the APWG feed, we remove it from consideration as a precaution.

B. Rock-phish and fast-flux websites

The monitoring we have described so far is used for phishing websites that are hosted directly on compromised web servers. However, there are two important types of phishing attack that operate using proxy machines, and we monitor these attacks somewhat differently.

The first form of attack we ascribe to the 'rock-phish' gang, a group of criminals who perpetrate phishing attacks on a massive scale [7]. Rather than ad hoc compromises of machines to host fake HTML, the gang purchases substantial numbers of domains with meaningless names such as `lof80.info`. Their spoof emails typically contain a long URL such as `http://www.bank.com.id123.lof80.info/vr`. Although the URL contains a unique identifier (to evade spam filters), all variants are resolved to a single IP address using 'wildcard DNS'. The IP address is of a machine that acts as an HTTP proxy, relaying web traffic to and from a hidden 'mothership' machine. If the proxy is removed, the DNS is

[6] In earlier work [6], we found only 39 URLs out of 176 654 submissions were subsequently identified as false positives on PhishTank.

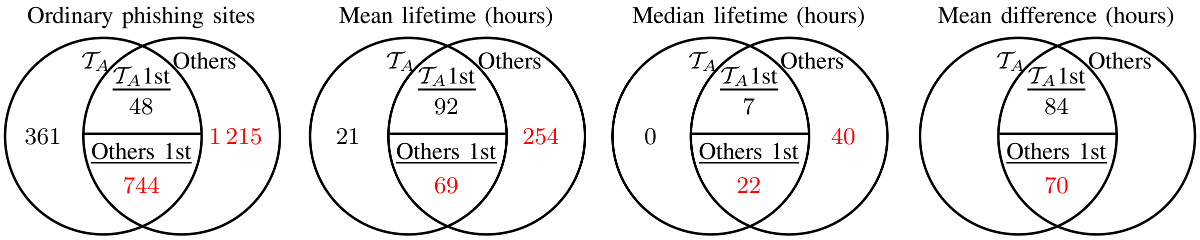

Fig. 1. Comparing take-down company $\mathcal{T}_A$'s awareness of phishing websites impersonating its client $A1$ to an amalgamated view constructed from all of our sources of data.

adjusted to use another proxy, and so the only practical way to remove the website is to get the appropriate registrar to remove the domain name from the DNS. A more complete description of the rock-phish gang's behaviour can be found in [2].

The second form of attack is dubbed 'fast-flux'. The mechanism is similar to the one that has just been described, except that the domain name is resolved to many IP addresses in parallel (typically 5 or 10) and the IP addresses used are rapidly changed (sometimes every 20 minutes). For these attacks the only practical approach is to have the domain name suspended. We have identified several disjoint fast-flux networks. Interested readers can find more details of fast-flux in [8] and about its use in phishing attacks in [2].

Besides using an innovative architecture, the rock-phish gang also attack multiple banks in parallel, with the URL path distinguishing between them. Since these bank 'micro-sites' generally appear and disappear together, we monitor the rock-sites generically, tracking whether the domain remains active. For convenience, we track fast-flux sites in a similar way, although they may attack only a single bank.

The data we discuss in this paper is taken from our record of the attacks that commenced in the period October 1, 2007 to March 31, 2008. To mitigate edge effects we monitored lifetimes for a further 50 days. Phishing attacks are continually evolving, but most of the attacks were relatively stable during this period – the exception being the fast-flux attacks which evolved from attacking one or two banks to attacking four or five at once. This makes them more like standard phishing sites at the beginning of the period and more like rock-phish at the end, which obscures many of the issues of co-operation we wished to examine. Hence, we completely exclude them from the analysis within this paper.

III. Evidence for non-cooperation in removing phishing websites

We start by considering the removal of standard phishing websites. We will consider rock-phish attacks in the next section, which means that in this section we are considering the take-down of sites that attack a single bank at a time. Take-down involves getting the offending web pages removed by the owner of the hosting machine, or sometimes suspending domains whose name has been chosen to be close to that of the real bank.

Some machines occasionally host attacks on several banks in parallel, and serial recompromise is very common, which in other research we attribute to the use of search engines to locate machines to compromise [9]. The main underpinning of high rates of recompromise is ineffectiveness at cleaning up compromised machines – and we believe that this makes it reasonable to view the disappearance of a particular set of web pages as being essentially independent of any other removal activity.

A. Motivating example

We begin by examining the phishing websites of one of take-down company $\mathcal{T}_A$'s clients, a bank called $A1$, which has hired $\mathcal{T}_A$ to remove phishing websites on its behalf.[7] While many of the websites impersonating $A1$ are identified by $\mathcal{T}_A$, $\mathcal{T}_A$ is not always aware of every site impersonating $A1$. Figure 1 presents Venn diagrams showing the sites impersonating $A1$. The left circle represents $\mathcal{T}_A$'s view of the sites impersonating $A1$, while the right circle represents the view from our own aggregated feed of phishing websites.

In $\mathcal{T}_A$'s view, 1 153 websites impersonated $A1$ during the sample period. In fact, information from other sources reveals that at least 2 368 websites did so. What is the effect of $\mathcal{T}_A$'s incomplete view? We can approximate the adverse effect by examining the lifetimes of the phishing sites for each category.

The sites $\mathcal{T}_A$ knows about exclusively are removed within 21 hours on average. Phishing website lifetimes are highly skewed [2], so a few long-lived websites can greatly impact the average lifetime. Hence, the median is a more robust measure of typical behaviour. The median lifetime for sites known exclusively to $\mathcal{T}_A$ is 0 hours – in other words, half the sites are already dead by the time we have started to monitor the URLs in the feed delivered to us from $\mathcal{T}_A$. By contrast, sites unknown to $\mathcal{T}_A$ have a much longer lifetime: 254 hours on average (40 hour median), nearly ten days longer than sites known exclusively to $\mathcal{T}_A$.

It is reasonable to ask why phishing websites unknown to $\mathcal{T}_A$ are removed at all. There are several plausible explanations. First, $A1$ may also attempt to remove phishing sites directly without involving the take-down company. Second, $A1$ may have hired more than one take-down company – we have

[7] Unfortunately, we cannot disclose the names of the take-down companies or the identity of any of their clients.

	$\mathcal{T}_A$ only			Others only			$\mathcal{T}_A$ first					Others first				
	Lifetime (hours)			Lifetime (hours)			Lifetime (hours)			Delay (hours)		Lifetime (hours)			Delay (hours)	
	#	mean	median	#	mean	median	#	mean	median	mean	median	#	mean	median	mean	median
$\mathcal{T}_A$'s top 6 clients:																
$A1$	361	21	0	1215	254	40	48	92	7	84	0	744	69	22	70	18
$A2$	1 526	21	0	890	66	14	182	36	19	31	16	997	41	19	45	16
$A3$	552	11	0	556	73	15	99	26	12	163	23	697	56	20	24	13
$A4$	364	5	0	981	68	0	51	132	0	0	0	862	12	0	19	8
$A5$	387	8	0	252	44	0	87	10	0	39	2	296	15	0	46	10
$A6$	106	38	0	1 248	63	18	41	80	43	74	29	445	130	46	68	25
Combined totals for the 54 clients of $\mathcal{T}_A$ that were attacked:																
$A*$	4 118	17	0	5 962	112	12	577	44	12	67	17	4 313	56	18	50	15
	$\mathcal{T}_B$ only			Others only			$\mathcal{T}_B$ first					Others first				
	Lifetime (hours)			Lifetime (hours)			Lifetime (hours)			Delay (hours)		Lifetime (hours)			Delay (hours)	
	#	mean	median	#	mean	median	#	mean	median	mean	median	#	mean	median	mean	median
$\mathcal{T}_B$'s top 6 clients:																
$B1$	522	14	0	84	55	0	299	29	11	37	11	66	44	28	25	5
$B2$	176	3	0	0	0	0	35	12	0	22	22	1	0	0	0	0
$B3$	99	45	0	23	26	0	41	76	21	6	6	9	48	38	5	0
$B4$	99	0	0	6	0	0	34	0	0	0	0	3	0	0	1	1
$B5$	112	49	0	14	15	8	32	29	13	4	3	6	24	26	4	1
$B6$	113	13	0	16	190	7	63	13	2	17	11	14	26	21	4	4
Combined totals for the 66 clients of $\mathcal{T}_B$ that were attacked:																
$B*$	2 225	18	0	199	91	0	722	21	4	53	11	120	37	22	25	3

TABLE I

Phishing-website lifetimes for the clients of take-down companies $\mathcal{T}_A$ and $\mathcal{T}_B$, broken down according to whether the websites are identified by their respective take-down provider or by outside sources. When websites appear in more than one feed, the time difference between appearances is also given.

entirely excluded from our analysis clients we were told were employing multiple take-down companies, but the companies may have been unaware of all of their clients' arrangements. Another explanation is that the owner of the machine where the website was hosted may have become aware of the site themselves, either through examining logs, or because of other abusive activity, from sending out spam to hosting IRC 'bots'. Reports may also be received from individuals who contact the website owner directly, or work through one of the many third parties (such as the CastleCops team) who work on a voluntary basis to remove phishing websites.

In addition to the websites impersonating $A1$ that $\mathcal{T}_A$ was unaware of, many websites were identified by $\mathcal{T}_A$ some time after other sources knew of their existence. Of the 1 153 websites $\mathcal{T}_A$ did know about, 744 were identified by the other sources first, at an average of 70 hours (18 hrs median) before $\mathcal{T}_A$ learnt of them. The impact of such delays can be seen in the lifetime figures: these websites remain for 69 hours on average, 48 hours longer than sites known only to $\mathcal{T}_A$.

Clearly, the bank $A1$ stands to gain if phishing website lifetimes are reduced because $\mathcal{T}_A$ has learnt of their existence in a timely manner by obtaining feeds from other take-down companies. $\mathcal{T}_A$ would also benefit, not only from the increased revenue opportunities of having more work to do, but also from being able to market a better overall service.

However, we have only examined one client bank so far, and policy should be based on more than a single example, so we now conduct a more comprehensive analysis, examining all of $\mathcal{T}_A$'s clients, along with the clients of a second take-down company called $\mathcal{T}_B$.

B. Comparing phishing website coverage of two take-down companies

Having just demonstrated that incomplete knowledge of phishing websites has harmed one bank and one take-down company, we now show that the effect applies more broadly. We study the phishing feeds provided to us by two take-down companies, called $\mathcal{T}_A$ and $\mathcal{T}_B$. We cannot simply compare the complete coverage of each feed to ascertain whether their feeds are comprehensive. This is because disparities between the feeds will inevitably result from take-down companies being more concerned about obtaining comprehensive lists of websites that impersonate their client banks than in overall completeness. Instead, we examine the extent to which their feeds omit websites impersonating their respective clients.

Table I breaks down website lifetimes for the six most frequently attacked client banks for each take-down company. We refer to company $\mathcal{T}_A$'s clients as $A1$ (this is the same $A1$ as in Section III-A above), $A2$, $A3$, $A4$, $A5$ and $A6$, and similarly we refer to company $\mathcal{T}_B$'s most frequently attacked clients as $B1$, $B2$, $B3$, $B4$, $B5$ and $B6$.

We study $\mathcal{T}_A$'s performance first. Notably, the website lifetimes of every client $A2$–$A6$ is consistent with the intuition established for $A1$ in the previous section: sites missing from $\mathcal{T}_A$'s feed take much longer to be removed, while sites appearing in $\mathcal{T}_A$'s feed only after being discovered elsewhere are also longer-lived. We can also see that $\mathcal{T}_A$'s feed is quite incomplete when compared with our amalgamated view. Combining all of $\mathcal{T}_A$'s 54 clients, 5 962 phishing websites, or 40% of the total, were completely missed by $\mathcal{T}_A$. Another 4 313 websites, or 29%, were identified by other sources before $\mathcal{T}_A$ identified them. Hence, $\mathcal{T}_A$'s clients stand to gain substantially if the company can arrange to obtain more feeds.

In particular, there exists a large gap of over three and a half days between the lifetimes of phishing websites known exclusively to $\mathcal{T}_A$ and those known only to others. This stark difference can help us understand why there are conflicting views of take-down performance, and it offers an explanation why take-down companies often boast of the speed with which they remove targets when the actual times measured by outside observers can be worse [2].

Having established that $\mathcal{T}_A$ stands to gain from greater sharing of phishing website feeds, we now study whether $\mathcal{T}_B$ also needs to obtain more information. Studying additional companies is important because it helps determine whether the case for reciprocation can be made. In other words, do $\mathcal{T}_A$ and $\mathcal{T}_B$ both stand to gain from sharing their feeds with each other?

Alas, the case for $\mathcal{T}_B$ is less clear-cut than for $\mathcal{T}_A$. $\mathcal{T}_B$'s feed appears substantially more comprehensive than $\mathcal{T}_A$'s. The benefit $\mathcal{T}_B$ might obtain from outside sources is smaller. For client $B2$, for example, there is no unique contribution from other sources, just one site is discovered by others before $\mathcal{T}_B$.

$\mathcal{T}_B$'s most attacked client $B1$ exhibits lifetime figures that are consistent with the results for $\mathcal{T}_A$: the 84 websites discovered exclusively by others remain for two days longer than those identified only by $\mathcal{T}_B$, and the 66 sites picked up by others before $\mathcal{T}_B$ remain for about one day longer. $B6$'s lifetimes are also consistent with intuition. However, the lifetimes for clients $B2$–$B5$ do not nicely match up to the expected outcomes as $\mathcal{T}_A$'s top clients did. This could be due to the small sample size of the other contributions, the scarcity of which would reflect well on $\mathcal{T}_B$, or it could merely arise because $\mathcal{T}_B$ receives much the same feeds as we do. Nonetheless, the overall contribution from others, while small, is not entirely trivial. Of the 3 266 websites impersonating $\mathcal{T}_B$'s clients, 199, or 6%, were identified by others and missed by $\mathcal{T}_B$. An additional 120 websites, or 4%, were picked up by others first. When we combine the results from all 66 of $\mathcal{T}_B$'s clients, the outcome once again becomes consistent with that found for $\mathcal{T}_A$'s clients. We conclude that although the effect is smaller than for $\mathcal{T}_A$, almost all of $\mathcal{T}_B$'s clients still stand to gain something from a shared data feed.

Table I also lists the average time lag in reporting between phishing websites that are detected by the relevant take-down company and by someone else as well. Whenever $\mathcal{T}_A$ is slower to identify phishing websites, it is 50 hours slower than the first reports on average. This corresponds to the average 39-hour gap between sites identified by others first and sites only found by $\mathcal{T}_A$. Similarly for $\mathcal{T}_B$, the average 25-hour difference whenever its reports are slower matches the 19-hour gap between lifetimes when $\mathcal{T}_B$ finds sites alone and when other feeds pick up the websites first. Hence, the data on differences reinforce the connection between delays in reporting and longer phishing-website lifetimes.

C. Non-cooperation harms big targets more

Table I also reveals substantial differences between the composition of $\mathcal{T}_A$'s and $\mathcal{T}_B$'s clients. Slightly more of $\mathcal{T}_B$'s clients were attacked during the period of our study (66 to $\mathcal{T}_A$'s 54), yet $\mathcal{T}_B$'s clients are impersonated much less frequently (3 266 to $\mathcal{T}_A$'s 14 970). Many of $\mathcal{T}_B$'s clients are smaller banks and credit unions, whereas $\mathcal{T}_A$'s client base includes several large national banks, which provide more attractive criminal targets.

52% of $\mathcal{T}_B$'s clients were impersonated fewer than 10 times during the sample period, compared to 37% of $\mathcal{T}_A$'s clients. Notably, 13% of $\mathcal{T}_A$'s client banks were impersonated more than 1 000 times, while none of $\mathcal{T}_B$'s were. These highly-targeted clients account for much of the difference in the total number of sites removed.

Given such wide disparity, it is worth examining whether the number of phishing websites per client affects how likely outside sources are to contribute. Figure 2 plots the proportion of client phishing sites detected by $\mathcal{T}_A$ (left) and $\mathcal{T}_B$ (right) compared to other sources as the number of impersonating websites varies (note the logarithmic x-axis binning).

As a bank is targeted more frequently, a single feed becomes less complete. For $\mathcal{T}_A$'s client banks targeted fewer than 10 times, $\mathcal{T}_A$ was the only source for 68% of the phishing sites. However, $\mathcal{T}_A$'s feed is found lacking as brands are targeted more. The proportion of phishing sites identified by $\mathcal{T}_A$ steadily decreases, finally to 24%, for clients impersonated more than 1 000 times. For these highly targeted clients, other feeds contributed 40% of the total websites detected, alongside another 31% of websites picked up by others before $\mathcal{T}_A$.

In other words, for a bank impersonated 10 times, $\mathcal{T}_A$ might be the sole source for 7 phishing websites, with 1 also picked up by others and 2 missed by $\mathcal{T}_A$ altogether. For a big bank hit 1 000 times, though, $\mathcal{T}_A$ might uniquely contribute 240 sites. Additionally, 50 sites might be spotted by $\mathcal{T}_A$ before others, 310 identified by others before $\mathcal{T}_A$, and 400 phishing websites completely missed by $\mathcal{T}_A$. In terms of both the absolute number of sites and proportions, more-frequent targets suffer most from non-cooperation.

$\mathcal{T}_B$'s clients do not appear to follow the same distinctive pattern, but there is insufficient data to make a definitive judgement as to why this might be. Although it looks as if $\mathcal{T}_B$'s effectiveness is independent of the size of the attack, it has no clients that are attacked over 1 000 times, and 63 of its 66 clients are attacked less than 200 times. In other words, when considering numbers of attacks, $\mathcal{T}_B$'s client base is less

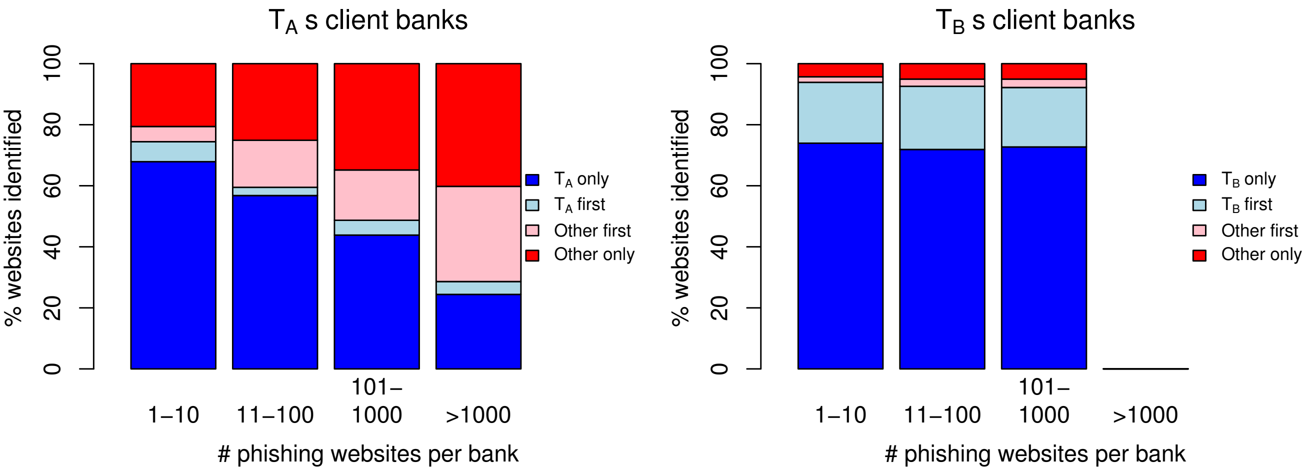

Fig. 2. Proportion of client phishing sites identified by take-down companies and other sources (company $\mathcal{T}_A$ – left, company $\mathcal{T}_B$ – right). For $\mathcal{T}_A$, as the number of observed phishing sites per client increases, outside sources identify a larger proportion of the sites.

diverse than $\mathcal{T}_A$'s. Nevertheless, even though $\mathcal{T}_B$'s feed is more complete and faster than $\mathcal{T}_A$'s at present, there would still be a slight benefit from outside assistance. What we cannot be sure of is whether the impact would remain marginal if $\mathcal{T}_B$ was to take on more highly-targeted clients.

There are several reasonable explanations for the effect we have just observed. Banks and credit unions will send the details of websites they learn about to the take-down companies for removal. Since only the take-down company they have hired can be expected to take any action, the banks are unlikely to inform anyone else, which explains why the sites are missing from other sources. In addition, when only a handful of phishing sites are created, the bank may be able to identify all of them, so the more general phishing-site detection mechanisms used by the take-down companies are not as helpful. When there are many sites, by contrast, it is also likely that there has been a stronger spam effort. In these circumstances, a bank-provided list is unlikely to be sufficient or timely. The take-down companies do use spam traps and other proprietary methods of identifying phishing websites. However, these techniques are unlikely to be comprehensive, and they are likely to miss more sites whenever many are being created.

D. Non-cooperation causes long-lived sites

In earlier work [2], we found that the distribution of phishing-website lifetimes corresponds to a highly skewed lognormal distribution. This means that most websites are removed quickly, but there is a 'long tail' of websites that remain for much longer, even for many weeks. There are several explanations for the existence of long-lived phishing websites. One is that the sites are hosted in places with unresponsive owners and ISPs. Another is that the take-down company or bank is unaware of the website entirely. Examining the feeds for both $\mathcal{T}_A$ and $\mathcal{T}_B$, we find evidence that not knowing of a website increases the chances that it will remain for a longer period.

We studied the proportion of websites impersonating $\mathcal{T}_A$'s and $\mathcal{T}_B$'s client banks that remain alive for more than one week. Our findings are presented in Figure 3. Overall, 6.8% of websites impersonating $\mathcal{T}_A$'s clients remain for at least one week. However, just 3.4% of the websites that $\mathcal{T}_A$ was aware of remained for that long. Strikingly, 12.1% of websites missed by $\mathcal{T}_A$ but identified by others remain for more than a week. This higher proportion once again suggests that knowing about a website has a great impact on whether it will be removed.

If the other sources had shared their feeds with $\mathcal{T}_A$, then $\mathcal{T}_A$ would have tried to take them down, and so only 3.4% might be expected to remain after a week. Thus, $5\,962 \times (12.1 - 3.4)\% = 519$ websites impersonating $\mathcal{T}_A$'s clients might be removed quickly instead of hanging on for much longer.

Similar results can be seen for $\mathcal{T}_B$, albeit on a smaller scale. Overall, 2.0% of websites impersonating its client banks remain for more than one week, but if $\mathcal{T}_B$ knows about them the proportion is 1.7%; whereas 5.7% last a week or more if $\mathcal{T}_B$ is ignorant of their existence.

E. What is the cost of non-cooperation?

Thus far, we have established that not sharing phishing feeds slows down the removal of some phishing websites. This inevitably exposes more customer details. We now attempt to quantify the added risk imposed by non-cooperation. Exposure is most simply measured by the number of hours phishing websites are accessible. While introducing many caveats, translating risk into dollars may nonetheless aid decision makers, and we can build on existing work to do so.

In earlier work, we devised a rough measure for the overall cost of phishing [2]. First, we estimated the number of victims a typical phishing website snagged over time by examining visitor logs from several real phishing websites. Next, we combined the victim estimate with average phishing-website

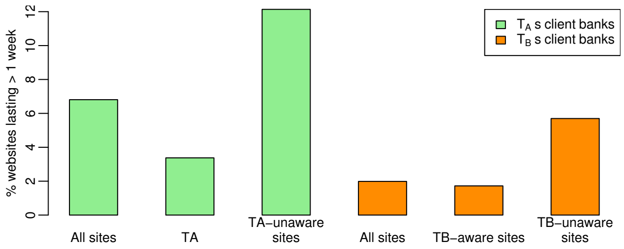

Fig. 3. Proportion of websites lasting more than one week depending on who knows about them.

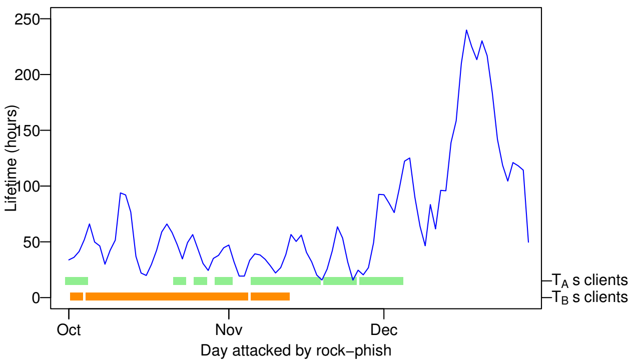

Fig. 4. Timeline for rock-phish attacks.

	Rock-phish sites		Lifetime (hours)	
	#	days	mean	median
All rock-phish	2 458	92	73.1	34.0
Neither $\mathcal{T}_A$ nor $\mathcal{T}_B$ attacked	739	30	141.2	76.4
Only $\mathcal{T}_A$ attacked	590	22	43.0	21.3
Only $\mathcal{T}_B$ attacked	553	24	47.8	33.5
$\mathcal{T}_A$ and $\mathcal{T}_B$ attacked	576	16	40.8	27.6

TABLE II
VARIATION IN ROCK-PHISH WEBSITE LIFETIMES WHEN $\mathcal{T}_A$'S AND $\mathcal{T}_B$'S CLIENTS ARE TARGETED.

lifetimes and a $572 per victim loss calculation from Gartner [10]. Florêncio and Herley arrived at a remarkably similar estimate of phishing victims obtained using a very different technique [11]. We apply our earlier method here to quantify the financial exposure to banks caused by not sharing URL lists.

Given that the average lifetime of all 14 970 websites impersonating $\mathcal{T}_A$'s banks is 67 hours, the total hours exposed is 1 005 000. This translates, using the formula from [2], to a total financial exposure of $276 million for $\mathcal{T}_A$'s banks. But what portion of the $276 million is caused by not sharing feeds? To determine this, we must first estimate the lifetime of sites whenever $\mathcal{T}_A$ knows about them, and then calculate the difference in time for the sites missed or identified later by $\mathcal{T}_A$. To compute the lifetime of sites $\mathcal{T}_A$ knows about, we subtract the average difference for sites identified by others first, arriving at an average of 13.5 hours.

If the 5 962 websites missed by $\mathcal{T}_A$ had instead been identified by $\mathcal{T}_A$, we would expect their lifetimes to shorten from 112 hours to 13.5 hours. This difference represents a timed exposure of 587 000 hours, and the formula yields a financial exposure of $119 million. But that is not all. We also have to account for the 4 313 websites identified by $\mathcal{T}_A$ after other sources were aware of them, since this caused an average delay of 50 hours. This translates to a timed exposure of 216 000 hours and a financial exposure of $44 million. Therefore, the financial exposure to $\mathcal{T}_A$'s client banks caused by not sharing feeds is $163 million over six months, or $326 million at an annualised rate. The table presents the complete figures, along with the lower results for $\mathcal{T}_B$'s banks:

Exposure figures (6-month totals)	**$\mathcal{T}_A$'s client banks**	**$\mathcal{T}_B$'s client banks**
Actual values	1 005k hrs ($276m)	78k hrs ($32.0m)
Effect of not sharing	587k hrs ($163m)	17k hrs ($3.5m)
Expected if sharing	418k hrs ($113m)	61k hrs ($28.5m)

Combining the results, we conclude that the overall lifetimes of phishing websites has increased by a factor of 2.3 as a direct result of not sharing URLs, and the financial exposure of the banks to phishing has been increased by a factor of 2.2 (the non-linear effects in our earlier formula almost cancelling out).

We cannot extrapolate an industry-wide measure of exposure due to non-cooperation from these figures alone. Together $\mathcal{T}_A$ and $\mathcal{T}_B$'s clients had 18 236 phishing websites removed, approximately 18% of all sites removed during the period. We anticipate that removal could also be speeded up for the other 82% of websites. Note further that these figures do not include the cost of not co-operating when tackling rock-phish attacks, which we discuss in the next section.

IV. CO-OPERATION AND ROCK-PHISH

As we explained in Section II, because of the proxy scheme being used, the only effective way to take down phishing websites created by the rock-phish gang is to persuade a registrar to suspend a domain name.

The rock-phish attacks represent a common threat to the banking industry – up to 25 banks are impersonated simultaneously within each domain, and all currently impersonated banks may be reached from any live domain. Hence, there is automatically implicit co-operation in the removal of rock-phish domains because whoever gets the domain suspended stops all of the attacks on other banks simultaneously. Nonetheless, we have found no evidence that any explicit co-operation is occurring at present. While many banks and take-down companies may not fully understand the nature of rock-phish attacks, the take-down companies $\mathcal{T}_A$ and $\mathcal{T}_B$ certainly do, as their clients have been targeted for some time. They track rock-phish websites, even when their clients are not currently attacked. However, the companies tell us that they wait until their own clients are targeted before starting to actively remove rock-phish domains.

We studied the clients of $\mathcal{T}_A$ and $\mathcal{T}_B$ that were targeted by

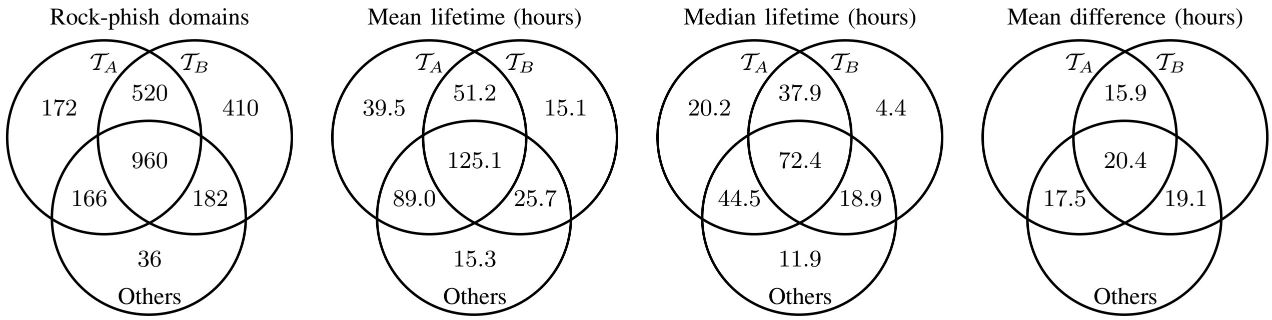

Fig. 5. Comparing awareness of rock-phish domains for take-down companies $\mathcal{T}_A$ and $\mathcal{T}_B$.

the rock-phish gang between October and December 2007.[8] Five of $\mathcal{T}_A$'s clients and two of $\mathcal{T}_B$'s were attacked. Figure 4 shows the days when new rock-phish domains appear and their clients are one of the targets. $\mathcal{T}_B$'s clients were attacked throughout October until mid-November, while $\mathcal{T}_A$'s clients were attacked briefly in early October and again beginning in late October through early December. For most of December, neither $\mathcal{T}_A$'s nor $\mathcal{T}_B$'s clients were attacked. Figure 4 also plots (as the blue line) the average lifetimes of rock-phish domains depending on the day on which the domains were launched. It is immediately apparent from this plot that activity by $\mathcal{T}_A$ and $\mathcal{T}_B$ significantly shortens the lifetime of rock-phish domains. When $\mathcal{T}_A$ and $\mathcal{T}_B$ are not actively removing rock-phish domains, as happened during December, the domains remain up for much longer.

Table II provides further insight into the impact of $\mathcal{T}_A$ and $\mathcal{T}_B$ on defending against rock-phish attacks. We observed 2 446 rock-phish domains over 92 days. On average, these domains are removed after 73 hours, about three days. However, rock-phish domains launched on days when neither $\mathcal{T}_A$'s nor $\mathcal{T}_B$'s clients are attacked last 141 hours, nearly twice as long as average. On days when only $\mathcal{T}_A$'s clients are attacked, domains are removed within 43 hours, considerably faster. The story is similar for days when only $\mathcal{T}_B$'s clients are attacked (48 hours), and domains are removed fastest (40 hours) on days when clients of both $\mathcal{T}_A$ and $\mathcal{T}_B$ are attacked.

We also tested the feeds' coverage of rock-phish domains. Figure 5 shows Venn diagrams for $\mathcal{T}_A$, $\mathcal{T}_B$ and others. Despite considerable domain re-use when targeting several banks simultaneously, there remain significant gaps in coverage. $\mathcal{T}_A$ knows about 1 818 domains, but is unaware of 628 domains, 26% of the total. Similarly, $\mathcal{T}_B$ is aware of 2 072 domains but missed 374, 15% of the total. If $\mathcal{T}_A$ and $\mathcal{T}_B$ exchanged rock-phish feeds both could defend their clients more effectively. The only potential impediment to sharing is that take-down revenue would be more evenly spread whenever they both have clients that are being attacked, and this might be a disincentive for a company that thought it had a more extensive list of domains.

While $\mathcal{T}_A$ and $\mathcal{T}_B$ stand to gain from exchanging rock-phish feeds, other companies and banks with less complete feeds stand to gain even more. One possible explanation of the much longer-lived rock-phish domains in December (see Figure 4) is that the other firms being targeted do not know about the domains $\mathcal{T}_A$ and $\mathcal{T}_B$ have discovered. Sharing their feeds with less-informed banks and take-down companies might greatly strengthen efforts to tackle rock-phish.

The middle two Venn diagrams in Figure 5 give the domain lifetimes based upon when different groups became aware. Surprisingly, it appears that the more feeds a domain appears in, the higher the average lifetime. Domains only appearing in the feeds from $\mathcal{T}_A$, $\mathcal{T}_B$ and elsewhere last for 39.5, 15.1, and 15.3 hours, respectively. The figures increase if the domain is picked up by one of $\mathcal{T}_A$, $\mathcal{T}_B$ and others, and increase even more to 125.1 hours on average (72.4 hours median) when the domain is identified by all three sources. This result runs counter to intuition – we might expect that more widely-known domains would be more likely to be removed, since at any given time at least one motivated defender will have identified the domain and will be trying to get it suspended.

So what might explain this effect? The first explanation is that our monitoring is likely to become aware of domains at an earlier time if several organisations are picking them up, and some detect the domain's usage faster than others. The Venn diagram in Figure 5 (right) examines the overlap and shows the difference between the first organisation to identify a site and the last. Whenever two groups identify the domain, the lag between the first discovery and the second ranges from 15.9 to 19.1 hours. Whenever a domain is rediscovered three times, the difference between the first and last identification is slightly higher, averaging 20.4 hours. This difference accounts for a portion of the difference in lifetimes.

Another reason why rediscovered websites last longer on average is that there may be a selection bias. Domains that remain around for a long enough period to be rediscovered several times are more likely to be difficult to remove, whereas domains that can be removed shortly after the first organisation finds it may be removed before the other detectors get a chance to notice them. It is difficult to determine the magnitude of this bias.

It should also be carefully noted that any selection bias

[8] We use the period October to December 2007 for rock-phish attacks because this is the period for which have reliable information as to whether the banks attacked were clients of $\mathcal{T}_A$ and $\mathcal{T}_B$ at the pertinent time.

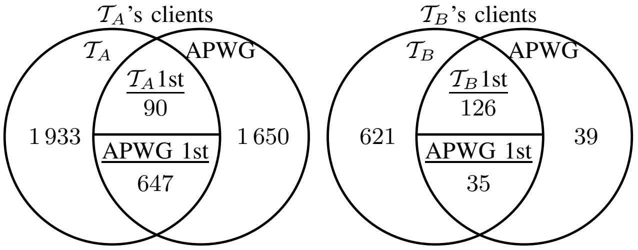

Fig. 6. Phishing URLs appearing in the feeds from $\mathcal{T}_A$, $\mathcal{T}_B$ and the APWG.

that is present will have had a similar influence on our earlier results for the lifetimes of ordinary phishing websites described in Section III above. However, because the effect is to inflate the lifetimes we measure for multiply detected websites, it serves to strengthen our finding that websites appearing in just one single feed have longer lifetimes.

V. Can information sharing work in the anti-phishing industry?

On the face of it, the anti-phishing industry co-operates quite a bit. There is the APWG, which conducts regular meetings where members share tips on the latest attacker innovations. The organisation also collects and disseminates a feed of phishing websites. But who exactly contributes to this feed? It seems to primarily come from third parties who have access to spam data, or who build browser toolbars, but are not directly interested in removing sites. Once aggregated, this feed is then distributed to APWG members who are in the business of removing phishing websites.

However, the take-down companies have put themselves in a curious position. They will happily accept phishing feeds from any organisation willing to share. But there is seldom reciprocity, or a willingness to share with anyone who asks. Take-down companies often market themselves as having a unique and valuable insight into phishing, which other companies do not have. Some companies are proactive in selling their feed, or at least those URLs which target a given client. Hence, their own contributions serve as a differentiator and these URLs are not shared, as evidenced by the substantial number of websites uniquely identified by both $\mathcal{T}_A$ and $\mathcal{T}_B$ in earlier sections of this paper.

Comparing the feeds from $\mathcal{T}_A$ and $\mathcal{T}_B$ to the APWG's feed demonstrates how one-sided sharing via the APWG can be. Figure 6 compares the phishing URLs identified by $\mathcal{T}_A$, $\mathcal{T}_B$ and the APWG from January to March 2008.[9] Both $\mathcal{T}_A$ and $\mathcal{T}_B$ are active members of the APWG, and both claim to incorporate the APWG's feed into their own feeds.

$\mathcal{T}_B$ adequately processes most phishing URLs appearing in the APWG list. 80% of the 200 phishing websites impersonating $\mathcal{T}_B$'s clients picked up by the APWG were also identified by $\mathcal{T}_B$. It is unclear why the remaining 39 websites, or 20%, did not appear in $\mathcal{T}_B$'s feed. Perhaps the websites were removed before $\mathcal{T}_B$ processed them, or they were determined by $\mathcal{T}_B$ to be classified incorrectly. Notably, however, 621 websites identified by $\mathcal{T}_B$, or 76% of the total, did not appear in the APWG feed. This suggests that $\mathcal{T}_B$ does not contribute phishing URLs to the APWG feed.

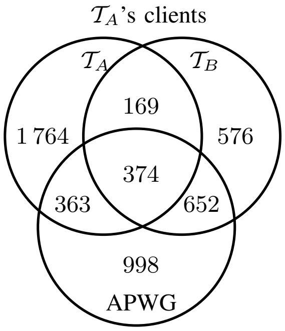

Fig. 7. Overlap in phishing URLs between $\mathcal{T}_A$'s feed and the feeds from $\mathcal{T}_B$ and the APWG.

$\mathcal{T}_A$ does not appear to contribute to the APWG feed either. 1 933 websites, or 45% of the total, are found by $\mathcal{T}_A$ and missed by the APWG feed. Notably, however, $\mathcal{T}_A$ has also failed to incorporate a significant number of sites impersonating its clients from the APWG feed into its own. 1 650 websites, or 38%, appeared in the APWG feed but not in $\mathcal{T}_A$'s own feed. The APWG feed is large, and subscribers must sift through a number of false positives and sites impersonating other banks, so processing it is far from trivial, and unfortunately, $\mathcal{T}_A$ seems to have missed a number of websites affecting its own clients.

Given that, during the period we studied, $\mathcal{T}_B$ does a better job processing the APWG feed than $\mathcal{T}_A$ does, we must also ask whether the missing websites for $\mathcal{T}_A$'s clients discussed in Section III are found in the APWG feed or from the contributions of others. Figure 7 shows a three-way Venn diagram for websites impersonating $\mathcal{T}_A$'s clients. Both $\mathcal{T}_B$ and the APWG make unique contributions to the sites missed by $\mathcal{T}_A$. $\mathcal{T}_B$ and the APWG both identify 652 sites missed by $\mathcal{T}_A$, while the APWG uniquely finds another 998 and $\mathcal{T}_B$ finds 576 missed by $\mathcal{T}_A$ and the APWG. This provides further evidence that take-down companies could contribute positively to their competitors' feeds.

Given that the feeds could better inform take-down companies and banks, we now weigh the merits and drawbacks of companies sharing their phishing feeds. We envision a third party collating the feed and distributing it to the industry, including the competitors of those who supplied the data. This third party could be the APWG, expanding their existing feed to include contributions from all possible producers and specifically the take-down companies. The APWG is in a strong position to encourage its members to contribute more widely to the feed. It could even consider making contribution a condition of receiving the feed.

We anticipate that take-down companies would share raw, unverified feeds of phishing URLs. As we explain in Section V-B below, anti-virus companies exchange virus samples,

[9] We lack APWG data from earlier periods.

but each verifies the sample's legitimacy and develops its own signatures. Similarly, take-down companies add value by cleaning up phishing feeds and standing behind their assessments. Take-down companies could continue to charge for processed feeds to clients who want to receive this type of validated data.

A. Incentive analysis

Banks targeted by phishing attacks would benefit most from sharing feeds. Any reduction in the lifetimes of phishing websites would be welcome. From the discussion in Section III-C, larger banks that are targeted more often stand to gain the most.

We believe that take-down companies would also benefit, through improved service for their clients and increased revenue per client as more websites are identified. Inevitably, the benefit to a particular company would vary greatly, but what matters is not only the speed and effectiveness of their existing detection, but also the types and sizes of clients they have (see Section III-C). It is true that our analysis has shown that company $\mathcal{T}_B$ stands to gain less from sharing than company $\mathcal{T}_A$, due to both $\mathcal{T}_B$'s more comprehensive feed and its large base of small clients. However, $\mathcal{T}_B$ would still be able to offer marginal improvements to its clients, and it seems entirely likely that contributions from the other take-down companies we have not studied would also be beneficial.

Of course, there are disadvantages to sharing as well – which helps to explain why it is not happening already! Phishing feeds do have inherent value. Some take-down companies emphasise the selling of feeds to clients, so they are more likely to take a dim view of 'giving them away for free'. For other companies, though, reduced sales from feeds should be dwarfed by the prospect of increased take-down activity. It would require some changes in marketing stance: take-down companies compete on a number of factors, including price, customer service, speed of removal, and feed completeness. Sharing feeds would eliminate one competitive aspect, and well-established companies with comprehensive feeds (such as $\mathcal{T}_B$) may view their feed as giving them a major advantage over less-informed firms when selling services to a client. In such a case, it would be in their interest to refuse sharing with weaker competitors.

Were widespread sharing of phishing feeds to happen, it would have significant competitive implications. Most importantly, sharing would substantially lower barriers to entry for prospective take-down firms. This would be good for competition within the take-down industry, and consequently helpful to banks, take-down companies' primary clients. By contrast, reduced barriers to entry would be viewed negatively by established take-down companies.

Widespread sharing could allow some companies to 'free-ride' by taking the feed without investing any effort into finding new sites to contribute. While this is a significant concern, the existence of thriving community feeds such as PhishTank and the APWG list suggests that a significant proportion of sites are already detected completely independently of the take-down companies. Furthermore, take-down companies should remain incentivised to continue operating their own detection systems because they will wish to identify phishing sites that impersonate their own clients; passing along irrelevant sites detected at the same time is unlikely to significantly increase their costs.

Another form of free-riding is possible as banks face common threats like rock-phish attacks. Section IV revealed diminishing lifetimes for rock-phish domains as more defenders get involved. This is hardly surprising. Defending against rock-phish attacks can be considered a 'sum-of-efforts' problem, where total protection depends on the aggregate contributions from each defender. Game-theoretic analysis of similar circumstances [12] reveals the potential for less efficient players to free-ride off the efforts of the higher-motivated. Take-down companies are highly motivated to remove phishing sites whenever their clients are targeted, since they are compensated for removing sites. However, banks who do not outsource website removal may be tempted to free-ride off the efforts of the more motivated take-down companies so long as they mitigate the threat from rock-phish more than the expense of removing sites directly.

B. Information-sharing examples from elsewhere in information security

At this point it is helpful to consider other threats to information security where sharing has or has not happened. We first consider the anti-virus industry. In its early days, anti-virus companies did not share virus definitions; instead, they differentiated themselves by the comprehensiveness of their lists. Trade magazines published head-to-head comparisons of competing products, testing whether 'Dr. Solomon' caught more viruses than 'Norton'. However, this produced significant biases in the results depending upon who supplied the virus samples that were tested. In 1993, a series of press releases from each the major companies promoted the fact that some new virus was being overlooked by the competition, and it became clear that the overall effect was damaging to the industry. A meeting at that year's EICAR conference led to an agreement to stop hoarding viruses, and the firms began to share virus samples with their competitors [13].

Today, whenever a virus is identified, it is first published to a common list so that each company can develop its own detection 'signature' as quickly as possible. Consumers benefit from more comprehensive virus detection, and the companies compete on other factors (such as price or levels of support). In fact, it is now viewed as extremely bad manners to refuse to share a virus sample, as evidenced by the industry's recent uproar over a newcomer's reticence to pass on information about a mobile phone virus [14]. To keep potentially harmful information from reaching outsiders (and forestall free-riding), the group remains quite exclusive and shares only between established members who have demonstrated value to the group. Such clubbiness is occasionally railed against by newcomers, such as when the organizers of a malware repository publicly pleaded with the industry to share [15].

A second lesson about sharing can be drawn from the world of vulnerability disclosure. Some security researchers advocate full and immediate disclosure: publishing details (including exploit code) on mailing lists such as Bugtraq [16]. While undoubtedly prompting vendors to publish a patch, full and immediate disclosure has the unfortunate side effect of leaving consumers immediately vulnerable. A more balanced alternative is 'responsible disclosure' as pioneered by CERT/CC in the US. CERT/CC notifies vendors to give them time to develop a patch before disclosing the vulnerability publicly. When the vulnerability is finally disclosed, no exploit code is provided.

Empirical analysis comparing patch-development times for vulnerabilities reported to Bugtraq and to CERT/CC revealed that CERT/CC's policy of responsible disclosure led to faster patch-development times than Bugtraq's full disclosure policy [17]. This is because CERT/CC has developed a more constructive relationship with software vendors, working with them to fix vulnerabilities.

Some firms, led by iDefense and Tipping Point, have gone a step further by actively buying vulnerabilities. Their business model is to provide vulnerability data simultaneously to their customers and to the vendor of the affected product, so that their customers can update their firewalls before anyone else. However, the incentives in this model have been shown by Kannan and Telang to be sub-optimal: users who do not participate in the closed circle of subscribers cannot protect their systems in time [18].

C. Recommendation

The anti-phishing industry has an important choice to make: whether to increase sharing, following the anti-virus industry's example, or to continue leveraging their feeds as a competitive edge, as is currently the case with some vulnerability brokers. In 2006, the anti-phishing industry appeared to be at the same point as the early days of the anti-virus industry, arguing over the completeness and accuracy of each other's anti-phishing toolbars [19].[10] Today, the APWG feed shows that some cooperation is occurring, and it has proved to be an effective way of getting disinterested third parties to contribute the URLs they learn about.

We believe that the evidence is strongly in favour of choosing to evolve beyond the current arrangements, much as the anti-virus industry did, and start viewing all phishing feeds as public goods rather than keeping some of the information private. Stopping short of a fully public arrangement, by instituting a 'sharing club', might appeal to the take-down companies by addressing issues of market entrance – but as we have already noted, this reduces competition, and so the banks may pay more than otherwise. Additionally – since feeds are also used by the anti-phishing toolbars – it may not be in the wider consumer interest either.

Whatever the minutiae of the change, in our view, sharing feeds is a winning proposition for most take-down companies: better coverage can lead to increased revenue and improved customer service. For the banks the issue is a 'no brainer': sharing feeds means that phishing websites that attack their brands will be removed more quickly. Only the few companies which specialize in producing feeds, and do little take-down of their own, can have any reasonable objection to the change of approach.

It is our recommendation that the take-down companies start sharing feeds immediately. Significantly, the banks, who pay the take-down companies for their services, can use their financial clout to encourage this change to happen. We contrast this relatively small number of clients, each with significant purchasing power, with the much broader spectrum of mainly individual customers that the anti-virus industry sells to. We would suggest this concentration of purchasing power means that comparatively rapid change could occur in the anti-phishing industry. Should this change not occur, then a regulator might intervene – but regulations that mandated sharing, while in our view economically justifiable in light of the data analysis presented in this paper, are unlikely to be practical – or indeed timely enough to be effective.

VI. Related Work

The academic work on phishing has been diverse, with a useful starting point being Jakobsson and Myers' book [21]. However, there has been only limited examination of the take-down process employed by the banks and specialist companies, even though it is the primary defence employed today. In earlier work [2], we estimated the number and lifetimes of phishing websites using data from PhishTank and demonstrated that timely removal reduced user exposure. This paper covers many more phishing sites (PhishTank contributes just 35% of the websites identified by our amalgamated feeds) and the wider view has meant that we have been able to explain how incomplete feeds contribute to the substantial number of long-lived sites that we reported earlier. Weaver and Collins examined two phishing feeds, and found that sharing of URLs was not taking place. They then computed the overlap and applied capture-recapture analysis to estimate the total number of phishing attacks that must be occurring [5].

Information sharing has been recognized as an important way to improve information security. Gordon and Ford discussed early forms of sharing in the anti-virus industry and contrasted it with sharing when disclosing vulnerabilities [22]. Worried about protecting critical infrastructures owned by private industry, the US government has encouraged data exchange via closed industry groups known as Information Sharing and Analysis Centers (ISACs). However, making information sharing occur in practice has often proven difficult. The development of ISACs has yielded mixed results. While some industries responded quickly, others took several years to comply. Many firms have expressed concerns over sharing security information with competitors and with the government [23]. Another worry about information sharing explored

[10] Meanwhile, academic research by Zhang et al. has contradicted the industry's sponsored research, showing none of the toolbars to be satisfactory [20].

in the academic literature is that firms might free-ride off the security expenditures of other firms by only 'consuming' shared security information (e.g., phishing feeds) and never providing any data of their own [24]. Even the threat of such free-riding occurring can stymie sharing.

There can also be positive economic incentives for sharing security information. Gal-Or and Ghose developed a model of where sharing can work [25]: they argue that information sharing can encourage additional security investment. In many circumstances, the providers of security services stand to gain by sharing information, which can drive up demand. Where there is a lack of industry awareness to threats, sharing information can certainly foster broader investment. This tendency to simultaneously share information and spend more on security has a more profound effect on fiercely competitive industries, such as take-down companies, where product substitutability is high. Gal-Or and Ghose also found that formal sharing organisations are more effective (in terms of information sharing and investment spurred) when members join sequentially. By joining first, market-leading firms bootstrap the alliance and demonstrate their commitment to share information, which encourages others to subsequently join. So if the more established take-down companies take the initiative and share feeds, others are likely to follow.

Often, overall security levels depend on the efforts of many interdependent principals. Hirshleifer told the story of *Anarchia*, an island whose flood defences were constructed by individual families and whose defence depends on the weakest link, that is, the laziest family; he compared this with a city whose protection against missile attack depends on the single best intercepting shot [26]. Varian extended this to three cases of interest to the dependability of information systems – where performance depends on the minimum effort, the best effort, or the sum-of-efforts [12]. Program correctness can depend on minimum effort (the most careless programmer introducing a vulnerability), while software validation and vulnerability testing may depend on the total of everyone's efforts. Similarly, defence against common threats like rock-phish attacks rely on the sum of efforts from all banks targeted. Further, constructing the most complete phishing feeds requires aggregating everyone's contribution. Varian's analysis predicts that the principals who stand to gain the most will carry out most of the effort, while some others free-ride. Since take-down companies are compensated for taking action they will all tend to find themselves in the former category, contributing when it helps their clients – especially if those clients are insisting upon the best possible service.

VII. Conclusions

In order to counter phishing attacks, banks put significant effort into removing the fraudulent websites that capture customer credentials. A specialist take-down industry has arisen to perform this function for many banks. We have obtained the feeds of phishing website URLs – their view of where attacks occur – from two take-down companies and have amalgamated this data with a number of other feeds. Over a six month period we have monitored how long all of the websites remain available, and we have then analysed these site lifetimes to see how effective the take-down companies are.

We examined data for the bank clients of the two take-down companies, and we found that websites had consistently longer lifetimes when the take-down company was either completely unaware they existed, or when they belatedly learnt of them. This effect was most apparent for banks that were frequently attacked, whereas it was less obvious, but still non-trivial, for small credit unions that might only be attacked on a handful of occasions. We also showed that websites were far more likely to last for more than a week if the take-down company was unaware of their existence.

We calculated how website lifetimes could more than halve if the take-down companies were to share information with each other. There is a direct linkage between longer take-down times and the funds put at risk by the compromise of visitor credentials. So we also translated these lifetimes from hours into dollars, finding that for these two companies alone – on some fairly rough estimates – around $330 million a year might be made safe.

We also examined take-down times for rock-phish domains and found that their lifetimes were higher when no client of the two take-down companies was being attacked. We demonstrated that once again each company was only seeing a part of the overall picture, and hence that lifetimes would be reduced by sharing information.

We considered the reasons why the take-down companies might not wish to share information, and concluded that in almost every case they would benefit, to a greater or lesser extent, from data sharing. We have therefore recommended that the industry change its practices as soon as possible. We noted in passing that the banks uniformly benefited from universal sharing and – since they are paying the bills – they are in a strong position to force change upon the industry.

Although our data analysis and results are specific to the take-down of phishing websites, we believe that the conclusions reached about the value of co-operation (and the real dollar cost of failing to do so) have application to other computer security scenarios as well, most notably in how the community handles knowledge of security vulnerabilities.

Acknowledgements

We are deeply grateful for the assistance we have received from the anti-phishing community generally and from the companies we have called $\mathcal{T}_A$ and $\mathcal{T}_B$ in particular. Without their data, the information they have provided about their clients, and their ongoing advice about how the industry works, this paper would not have been possible. Tyler Moore was supported in this work by the UK Marshall Aid Commemoration Commission and by US National Science Foundation grant DGE-0636782.

References

[1] B. Violino, "After phishing? Pharming!," *CSO Magazine*, October 2005, http://www.csoonline.com/read/100105/pharm.html.

[2] T. Moore and R. Clayton, "Examining the impact of website take-down on phishing," in *Anti-Phishing Working Group eCrime Researcher's Summit (APWG eCrime)*, 2007, pp. 1–13.

[3] Cyveillance Inc., "Cyveillance anti-phishing datasheet," 2007, http://www.cyveillance.com/web/docs/AntiPhishing.pdf.

[4] InternetIdentity Inc., "PowerShark phishing site deactivation," 2008, http://www.internetidentity.com/phishing-site-takedown.html.

[5] R. Weaver and M. Collins, "Fishing for phishes: applying capture-recapture to phishing," in *Anti-Phishing Working Group eCrime Researcher's Summit (APWG eCrime)*, 2007, pp. 14–25.

[6] T. Moore and R. Clayton, "Evaluating the wisdom of crowds in assessing phishing websites," in *12th International Financial Cryptography and Data Security Conference (FC08)*, Springer Lecture Notes on Computer Science (LNCS), vol. 5143, 2008, pp. 16–30.

[7] R. McMillan, 'Rock Phish' blamed for surge in phishing," *InfoWorld*, 12 December 2006, http://www.infoworld.com/article/06/12/12/HNrockphish_1.html.

[8] Honeynet Project and Research Alliance, "Know your enemy: fast-flux service networks, an ever changing enemy," July 2007, http://www.honeynet.org/papers/ff/fast-flux.pdf.

[9] T. Moore and R. Clayton, "Evil Searching: compromise and recompromise of Internet hosts for phishing," in submission, 2008.

[10] Gartner Inc., "Gartner says number of phishing e-mails sent to U.S. adults nearly doubles in just two years," Press Release, 9 November 2006, http://www.gartner.com/it/page.jsp?id=498245.

[11] D. Florêncio and C. Herley, "Evaluating a trial deployment of password re-use for phishing prevention," in *Anti-Phishing Working Group eCrime Researcher's Summit (APWG eCrime)*, 2007, pp. 26–36.

[12] H. Varian, "System reliability and free riding," in *Economics of Information Security*, Vol. 12, *Advances in Information Security*, L. J. Camp and S. Lewis, Eds. Boston: Kluwer Academic Publishers, 2004, pp. 1–15.

[13] M. Hypponen (private communication), 23 April 2008.

[14] R. Lemos, "Anti-virus groups fight over Crossover sharing," *The Register*, 6 March 2006, http://www.theregister.co.uk/2006/03/06/crossover_ruling_angers_industry/.

[15] B. Krebs, "Defcon speakers team up to fight 'queen bots'," *Washington Post*, 9 August 2006, http://blog.washingtonpost.com/securityfix/2006/08/defcon_speakers_team_up_to_fig.html.

[16] Security Focus Inc., "Bugtraq mailing list," http://www.securityfocus.com/archive/1.

[17] A. Arora, R. Krishnan, R. Telang and Y. Yang, "An empirical analysis of vendor response to disclosure policy," in *Workshop on the Economics of Information Security (WEIS)*, 2005.

[18] K. Kannan and R. Telang, "Market for software vulnerabilities? Think again," *Management Science*, vol. 51, no. 5, 2005, pp. 726–740.

[19] E. Montelbano, "Mozilla: Firefox antiphishing tool better than IE 7," *InfoWorld*, 15 November 2006, http://www.infoworld.com/article/06/11/15/HNmozillaphishing_1.html.

[20] Y. Zhang, S. Egelman, L. F. Cranor and J. Hong, "Phinding phish: evaluating anti-phishing tools," in *Proceedings of the 14th Annual Network & Distributed System Security Symposium (NDSS)*, 2007.

[21] M. Jakobsson and S. Myers, Eds., *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft*. New York: Wiley, 2006.

[22] S. Gordon and R. Ford, "When worlds collide: information sharing for the security and anti-virus communities," IBM research paper, 1999.

[23] R. Dacey, "Information security: Progress made, but challenges remain to protect federal systems and the nation's critical infrastructures," US General Accounting Office (GAO), GAO-03-564T, April 2003, pp. 1–75.

[24] L. Gordon, M. Loeb and W. Lucyshyn, "Sharing information on computer systems security: An economic analysis," *Journal of Accounting and Public Policy*, vol. 22, no. 6, 2003, pp. 461–485.

[25] E. Gal-Or and A. Ghose, "The economic incentives for sharing security information," *Information Systems Research*, vol. 16, no. 2, 2005, pp. 186–208.

[26] J. Hirshleifer, "From weakest-link to best-shot: the voluntary provision of public goods," *Public Choice*, vol. 41, 1983, pp. 371–386.