Measuring the Impact of Sharing Abuse Data with Web Hosting Providers

Marie Vasek
The University of Tulsa & StopBadware
800 S. Tucker Dr.
Tulsa, Oklahoma
vasek@utulsa.edu

Matthew Weeden
The University of Tulsa
800 S. Tucker Dr.
Tulsa, Oklahoma
matthew-weeden@utulsa.edu

Tyler Moore
The University of Tulsa & StopBadware
800 S. Tucker Dr.
Tulsa, Oklahoma
tyler-moore@utulsa.edu

ABSTRACT

Sharing incident data among Internet operators is widely seen as an important strategy in combating cybercrime. However, little work has been done to quantify the positive benefits of such sharing. To that end, we report on an observational study of URLs blacklisted for distributing malware that the non-profit anti-malware organization StopBadware shared with requesting web hosting providers. Our dataset comprises over 28 000 URLs shared with 41 organizations between 2010 and 2015. We show that sharing has an immediate effect of cleaning the reported URLs and reducing the likelihood that they will be recompromised. Despite this, we find that long-lived malware takes much longer to clean, even after being reported. Furthermore, we find limited evidence that one-time sharing of malware data improves the malware cleanup response of all providers over the long term. Instead, some providers improve while others worsen.

CCS Concepts

•**Security and privacy** → **Economics of security and privacy;** Web protocol security;

Keywords

security economics, web-based malware, abuse data sharing

1. INTRODUCTION

The fight against cybercrime is chiefly waged by private actors who independently administer components of the Internet infrastructure. These operators, including Internet service providers (ISPs), hosting providers, and domain name registrars, regularly process reports that indicate one or more of their customers are facilitating abuse, such as hosting phishing sites, distributing malware, or participating in a botnet. Abuse reports are largely sent by private actors such as security companies, ISPs, or volunteer organizations.

WISCS'16, October 24 2016, Vienna, Austria

ISBN 978-1-4503-4565-1/16/10. . . $15.00

DOI: http://dx.doi.org/10.1145/2994539.2994548

In this paper, we investigate one aspect of this crime-fighting ecosystem: the remediation of web-based malware by hosting providers. Many hosting companies have dedicated technical staff who assist in the cleanup of customer websites infected with malware, but their efficacy can vary greatly. Some merely forward received reports to their customers. Some will reach out to blacklist operators directly, who might offer some general instructions for cleanup and the ability to request removal once the infection has been eradicated. Others have reached out to StopBadware, a non-profit anti-malware organization established in 2006. In addition to helping individual website owners who have been hacked, StopBadware occasionally shares lists of malware URLs to inquiring hosting providers. In this paper, we empirically examine what happens to organizations after they request such data from StopBadware.

We make the following contributions:

1. We report on an observational study where 41 web hosting providers received bulk reports from StopBadware on 28 548 URLs flagged as malicious by Google Safe Browsing between 2010 and 2015.
2. Directly sharing URLs with hosting providers is an effective strategy for cleaning up those particular URLs. We find that the median report to clean time for shared URLs is 101 days, compared to 153 days for a comparative sample of hosting providers that were not notified. Furthermore, the recompromise rate for shared URLs is 4%, compared to 10% for similar providers who did not receive information from StopBadware.
3. We observe a positive correlation between the time a URL has been blacklisted and the time required to clean it once reported. Long-lived malware URLs prove difficult to clean even after a provider is notified.
4. We find that for some providers, the long-term impact of sharing data is positive: 39% of providers demonstrably reduced the time from blacklisting to remediate malware in time periods after a malware report. But for 52%, no long-term improvement was observed, and for 9%, the response time actually got worse.
5. We find that providers that improved after receiving a report cleaned subsequent infections two months faster on average. Providers whose performance worsened after sharing cleaned URLs six months slower. Furthermore, worsened providers had, on average, more than eight times more recompromised URLs compared to providers that either improved or remained steady.

2. RELATED WORK

Provos et al. developed a mechanism to identify so-called "drive-by-downloads": websites that attempt to automatically download malware onto visitors' computers without any interaction [10]. This system ultimately became Google Safe Browsing, a real-time blacklist of websites crawled by Google that appear to be distributing malware. We use the Safe Browsing data in the study, as described in Section 3.

A number of researchers have been investigating the effectiveness of sharing abuse data. Moore and Clayton found that the removal of phishing websites was significantly slowed because takedown companies typically refuse to share data with their competitors [8]. While competitive concerns can sometimes inhibit sharing and hamper the cleanup process, the public response from law enforcement is often slower. Hutchins et al. interviewed public and private actors involved in website takedown, finding that law enforcement agencies are slower at removing malicious websites than commercial firms, in large part due to the expertise private firms obtain through specialization and learning [4].

In previous work, Vasek and Moore conducted an experiment in which they reported web-based malware infections to two entities: hosting provider and either webmaster or registrar [12]. They found that detailed abuse reports that articulated the website's malicious activity were effective, but that reports lacking such details were indistinguishable from doing nothing at all. In a follow-up study using a dataset of URLs used by the Asprox botnet, Cetin et al. confirmed that detailed reports shorten the time required to cleanup [2]. They also found no evidence that the email address of the abuse report sender affected the time to clean. Li et. al. analyzed data from Google on its interactions with webmasters when trying to remediate web-based malware [7]. They found that individual website operators reached who signed up for alerts from Google were more likely to clean up and do so more quickly. Nappa et al reported 19 malware exploit servers to hosting provider contacts [9]. Canali et. al. set up websites on 22 shared hosting providers and attacked them [1]. After 25 days of attack, they sent out abuse notifications to the hosting providers and measured the responses. They found that most of their compromises were never remediated fully.

A few other studies have transmitted vulnerability reports to intermediaries and website operators. Dumeric et. al. reported susceptibility to the Heartbleed OpenSSL vulnerability to the hosting provider or internet service provider contact [3]. Notified hosts increased the patching rate by 47%. Kührer et. al. notified operators of NTP servers vulnerable to DDoS amplification attacks, observing a drop from 1.6 million vulnerable hosts to 126 000 in just three months [5]. Li et. al. ran a study notifying different responsible parties (hosting provider WHOIS contacts, national CERTs) about three types of vulnerabilities and found that sending detailed notices to hosting provider WHOIS contacts was the most effective [6]. Stock et. al. notified different responsible parties (varying webmaster, host, and country-level contacts) about WordPress and client-side XSS vulnerabilities [11]. They found a small statistical effect of all their notification efforts, while despairing over the inefficacy of large-scale notification campaigns.

The present work differs in a number of ways. Prior work has shared abuse reports for a single URL with the operators of websites. By contrast, we study the effect of reporting web-based malware URLs in bulk to a single requesting intermediary. Furthermore, our dataset spans a much longer period of time (6 years). With one exception [7], in the prior work the abuse reports are sent unsolicited. We also consider *compromises* rather than *vulnerabilities* (which may or may not later become compromised) making the incentive to clean much higher. In the present study, the intermediaries have asked for the information. This in turn could impact their willingness to take action. Finally, unlike the other studies, we do not provide the intermediaries resources to help guide the cleanup. This is because for bulk URL sharing, there can be significant heterogeneity in the types of malware and the ensuing cleanup strategy. Hence, the present work complements the prior work by investigating abuse data sharing in a new context.

3. DATA COLLECTION METHODOLOGY

3.1 Inquiries to StopBadware

StopBadware runs a manual, independent reviews process for three data providers: Google Safe Browsing, ThreatTrack Security, and NSFocus. When a webmaster searches for their URL through StopBadware[1], they can find all the times that the URL was blacklisted by any partner, and if it is currently blacklisted, they can request a review. Users can also search for Internet Protocol (IP) addresses and Autonomous System numbers (ASNs). Searching for IP addresses and ASN lists the number of URLs in any blacklist hosted and prompts the user, if responsible for that IP or ASN, to request a bulk data dump from StopBadware.

Note that by URL, we mean blacklisting entries. For example, perhaps malware was found on verybad.example.com/nogood/evil.php. This could potentially be the blacklisting entry. Other potential entries could be example.com, verybad.example.com, example.com/nogood/, or verybad.example.com/nogood/[2]. We refer to these blacklisting entries as URLs moving forward.

If the user is indeed responsible for that hosting space, URL information is shared as per agreements with the data partners. StopBadware shares as much data with hosting providers as their partners allow. For instance, Google Safe Browsing only permitted StopBadware to share a limited amount of data. In most cases, reports include data from at least two providers. All reports include specific URLs blacklisted and some include additional IP information. No further information is shared regarding particular compromises. However, some hosting providers later followed up on specific compromise by making requests to StopBadware's independent review process.

For the purposes of this study, we only investigate URLs blacklisted by Google Safe Browsing, even though StopBadware also shared URLs flagged by the other sources. We elected to do this because Google's blacklist contains relatively homogeneous malware URLs used to facilitate drive-by downloads, whereas their other partners' blacklists offer more heterogeneous lists of multiple forms of bad activity. These lists sometimes include content that is not necessarily malware or universally recognized as such. Consequently, for some URLs on these other lists a slow removal may some-

[1]https://www.stopbadware.org/clearinghouse/search
[2]https://developers.google.com/safe-browsing/v4/urls-hashing#suffixprefix-expressions

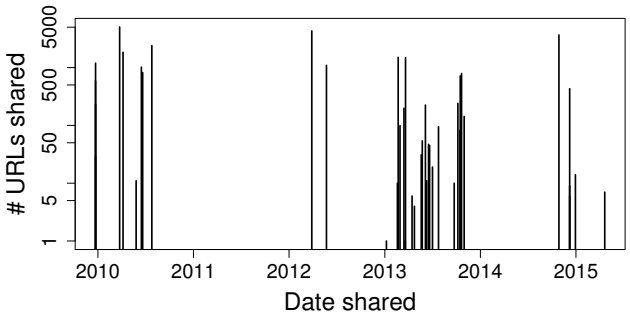

Figure 1: Bulk data reports sent from StopBadware to requesting hosting providers over time.

times reflect a policy decision as opposed to a hampered response. By sticking to Google reports for our study, we can measure the time to clean up a consistent, universally undesirable form of online wickedness.

We note that Google Safe Browsing also offers alerts for network operators of recently blacklisted URLs[3]. These alert network operators of new compromises on their network, whereas the reports in this study report **all** known compromises on their network.

From 2010 through 2015, StopBadware received 88 requests for bulk data from 69 different stakeholders ranging from country CERTs (Computer Emergency Readiness Teams) to AS (Autonomous System) owners to free domain services. For the purpose of this study, we wanted a more homogeneous group of entities to study; thus, we only considered hosting providers large enough to own their own AS, leaving us with 55 requests. A number of entities had multiple requests; removing those resulted in 42 remaining requests. Finally, we removed one AS whose request came a few weeks prior to the end of our study, since we could not reliably measure the long-term impact of sharing. Note that we refer to ASes and hosting providers interchangeably in the subsequent analysis. Figure 1 plots when StopBadware shared URLs with hosting providers over time. Each line indicates the number of URLs shared for each request.

3.2 Defining Malware Cleanup

We do not directly test websites for the presence of malware. Instead, we rely upon the outside judgments of blacklist providers to assess when a website is compromised, and therefore also when it becomes cleaned. The blacklists used by StopBadware are dynamic, and the operators strive to remove websites from the list as soon as they are believed to be clean. While we are not aware of any published studies of the accuracy of such blacklists, it is widely believed that these lists have very low false positive rates and modest false negative rates.

In most circumstances, it is straightforward to determine when a website is clean based on our data: it is simply the time that the website comes off the blacklist. Yet it is much less clear cut for a significant minority of websites that are placed back onto the blacklist shortly after being marked clean. In the extreme case, 0.05% of websites come on and off the blacklist 10 times or more. Some URLs rejoin the blacklist after a few hours, while others return years later.

[3]http://www.google.com/safebrowsing/alerts/

Re-blacklisting within a short period of time from the initial compromise could demonstrate that the attackers used the same vector of compromise or exploited a backdoor they left behind. It may also signal that malware managed to temporarily evade detection. To a first approximation, such websites have never really been cleaned. To that end, we attempted to distinguish between re-blacklisting events that signaled that the URL was never fully cleaned up and re-blacklisting events that signal that the URL was cleaned up and then was later compromised.

We check blacklist updates hourly. We consider a URL to be *clean* if it has been off the blacklist and stayed off for 21 days. If a URL is blacklisted after that 21-day period, we consider it to then be *recompromised*. If a URL falls of the blacklist and rejoins within 21 days, we consider the compromise to never have been cleaned up.

3.3 Measuring Secure Outcomes

We want to study both the direct impact of sharing URLs with the hosting providers and the indirect, longer term impact of sharing. To measure the direct impact of sharing, we looked at the time from when we reported the URLs to the time they were cleaned up, which we refer to as the *report-to-clean* time. Because many of the URLs shared had been compromised for a very long time before the hosting provider contacted StopBadware, measuring from the time of compromise to cleanup would not accurately measure provider effort compared to the report-to-clean time.

To measure the indirect impact of sharing, we compare hosting provider performance in the period before StopBadware shared URLs to the period afterwards. The idea is that some providers, upon receiving information about compromised websites in their network, will make improvements to the detection and remediation process that benefits their long-run security.

For measuring both the direct and indirect impact of sharing, we use survival analysis, a technique that works with *censored* data. Survival analysis allows us to include all data points, even the URLs reported which never came off the blacklist during the measurement interval.

We assign all URLs blacklisted before the reporting date to the *pre-contact* group. These are then compared with URLs blacklisted after the reporting date in the *post-contact* group. For both groups, we compute the *blacklist-to-clean* time, that is, the time from when each URL is added to the blacklist to the time it is marked clean. We censor the URLs blacklisted on the reporting date (since any still compromised at that time were shared with hosting providers). To make these time periods more comparable, we only consider URLs blacklisted within two years of the report date. If we sent the report in the last two years, we consider the same length of time before and after the report. For the post-contact group, any URLs still compromised at the end of the period are also censored.

In addition to survival probability, we also compute the overall recompromise rates for the pre- and post-contact groups. Recompromise rates, aggregated over a hosting provider during an extended period of time, offer a good indication of how effective the provider's efforts to clean up compromised websites are.

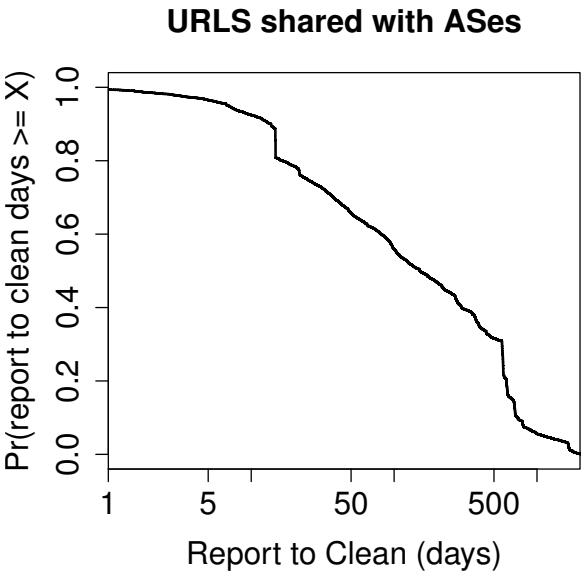

Figure 2: Survival probability for the number of days from a bulk report to clean (*reported* URLs)

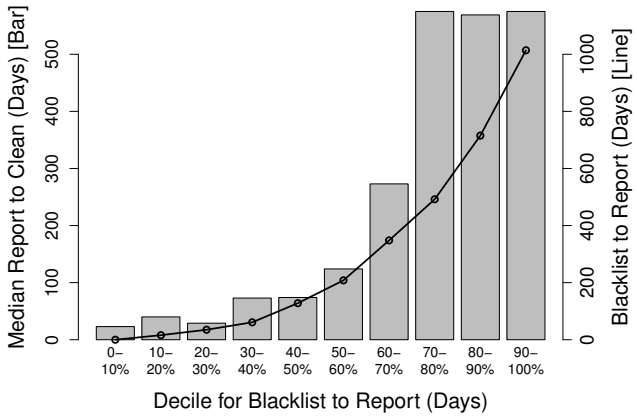

Figure 3: Measuring long lived malware: comparing report to clean times (bar) with blacklist to report times (line) for all reported URLs.

4. RESULTS

4.1 Direct impact of sharing

We first look at the rate at which URLs that StopBadware directly shared with hosting providers come off the blacklist. Figure 2 plots the survival probability for all 28 541 reported URLs. We find that within two weeks of receiving a report from StopBadware, about 20% of URLs come off the blacklist. Half the URLs come off the blacklist within 100 days of the report. Nonetheless, a significant minority of reported URLs remain compromised long after sharing. Approximately 40% remain blacklisted one year after StopBadware reported the URLs to the hosting provider.

We observe significant variation in the report to clean times by hosting provider. Figure 4 plots the survival probabilities for individual hosting providers (solid black line), along with the overall survival probability from Figure 2 in dashed lines. Note that we include only the 33 providers that experienced at least 10 malware reports both before and after sharing occurred. Mult indicates that the AS received multiple bulk reports from StopBadware.

Summary statistics for these providers is also provided in Table 1. The median number of total compromised URLs for the hosting providers receiving reports was 2 212, whereas the median number shared by StopBadware is 225, approximately 10% of the total. We note substantial variation in the number of malware URLs shared, ranging from only 1 to over 5 000.

Some providers clearly reacted more quickly to the information provided by StopBadware than others. For example, providers 2, 3, and 4 clearly outperformed the rest, cleaning up more quickly than the other providers. By contrast, providers 8 and 9 lagged substantially, waiting over a year to clean up the vast majority of URLs shared. For many other providers, the differences were not so clear cut. For instance, provider 22 lagged for the first couple months, but then cleaned up most URLs and eliminated most long-lived infections at above average speed.

As noted in Section 3.3, shared URLs have a longer lifespan than other URLs. This might be because they are more likely to be maliciously registered than compromised and attackers are good at hiding. This could also be from particularly pernicious bits of malware. While measuring the report-to-clean time mitigates this effect to some degree, long-lived malware is nonetheless harder to clean. We now attempt to quantify this effect.

Figure 3 demonstrates the relationship between long-lived malware URLs and the time required to clean them after reporting. The line plots the number of days from blacklist to when the report is sent by decile. For example, 20% of reported URLs had been blacklisted for 35 days or less when they were shared, but 60% had been blacklisted for 348 days or less. Meanwhile, 10% of shared URLs had already been on the blacklist for at least 1 014 days when StopBadware shared them. The bar plot indicates the median number of days from reporting to cleanup for each decile of the blacklist to report. The plot clearly demonstrates a strong positive relationship between long-lived infections and a slower time to clean once reported.

We note that one-off bulk reports like those shared between StopBadware and hosting providers are most helpful if they reflect a more systematic strategy to reduce abuse on a network, rather than a temporary fix. Thus, we spend the rest of this paper analyzing the lasting effects of bulk reports by studying what happens to URLs reported after sharing has taken place.

4.2 Long-term impact of sharing

We want to find if sending bulk reports to hosting providers creates a lasting impact on efforts to counter abuse. To this end, we consider all the hosting providers we report to. We then compare the URLs added to the blacklist before we shared with the hosting provider (pre-contact) with all the URLs added to the blacklist after (post-contact), as described in Section 3.

Figure 5 compares the cleanup rate for all ASes we reported to before and after we contacted them. We notice that cleanup is slightly faster for the post-contact URLs –

ASN	# URLs	Direct Reporting			Pre-Contact			Post-Contact			Long-Term Impact		
		# Shared	Day Shared	Report to Clean	# URLs	Blacklist to Clean	Recomp. Rate	# URLs	Blacklist to Clean	Recomp. Rate	Δ Blacklist to Clean	Δ Recomp. Rate	Survival Prob.
1	91	1	2013-03-20	393 days	12	46 days	0	34	14 days	0.029	31.5 days	-0.029	improved
2	576	82	2013-10-15	88 days	282	112 days	0.103	179	39 days	0.039	73.2 days	0.064	improved
3	3 740	198	2013-03-15	89 days	532	56 days	0.041	231	44 days	0.087	12.3 days	-0.045	unclear
4	44	10	2013-09-23	38 days	23	35 days	0.087	16	34 days	0.188	0.9 days	-0.101	improved
5	93 102	3 664	2014-10-28	247 days	43 061	52 days	0.061	11 325	69 days	0.032	-17 days	0.03	unclear
6	2 730	99	2013-02-28	151 days	353	99 days	0.142	2164	50 days	0.081	49.3 days	0.06	improved
7	2 663	241	2013-10-07	125 days	1 239	92 days	0.084	1 278	91 days	0.058	0.6 days	0.026	unclear
8	21 984	2 392	2010-07-26	696 days	6 616	169 days	0.114	2 457	288 days	0.073	-119.1 days	0.041	worsened
9	3 980	1 018	2010-06-16	799 days	456	322 days	0.094	976	334 days	0.04	-12 days	0.054	unclear
10	1 797	143	2013-10-31	143 days	866	89 days	0.07	421	80 days	0.076	8.6 days	-0.006	unclear
11	58	31	2013-05-20	102 days	33	81 days	0	25	76 days	0.04	5.5 days	-0.04	unclear
12	9 679	5 046	2010-03-25	244 days	5 577	122 days	0.042	2636	52 days	0.039	70.2 days	0.003	improved
13	13 383	792	2013-10-21	232 days	5 207	88 days	0.083	4 522	63 days	0.028	24.4 days	0.055	improved
14	26 716	4 322	2012-03-28	242 days	2 077	132 days	0.087	12 012	56 days	0.13	76.2 days	-0.043	improved
15	454	47	2013-06-17	121 days	151	84 days	0.066	155	53 days	0.071	31.6 days	-0.005	improved
16	76	4	2013-04-24	46 days	11	149 days	0	47	15 days	0	133.2 days	0	unclear
17	698	234	2009-12-23	226 days	500	121 days	0.036	60	120 days	0.1	1.5 days	-0.064	unclear
18	624	95	2013-07-25	110 days	261	123 days	0.111	241	68 days	0.095	54.7 days	0.016	improved
19	33 909	1 491	2013-03-21	122 days	12 924	239 days	0.077	16 323	152 days	0.039	86.6 days	0.038	improved
20	2 212	820	2010-06-21	650 days	725	164 days	0.113	185	509 days	0.054	-344.9 days	0.059	worsened
21	3 739	1 193	2009-12-23	149 days	1 403	155 days	0.066	413	86 days	0.099	68.7 days	-0.034	unclear
22	3 927	225	2013-06-05	167 days	812	72 days	0.087	266	68 days	0.064	3.6 days	0.024	unclear
23	1 554	590	2009-12-23	207 days	530	78 days	0.066	274	62 days	0.088	15.9 days	-0.022	unclear
24	2 294	431	2014-12-08	22 days	1 282	101 days	0.035	293	114 days	0.02	-13.3 days	0.015	unclear
25	317	45	2013-06-21	214 days	82	91 days	0.146	43	71 days	0.047	19.6 days	0.1	unclear
26	186	9	2014-12-08	271 days	85	111 days	0.141	48	110 days	0 days	0.9 days	0.141	unclear
27	69	14	2014-12-30	163 days	45	58 days	0.156	15	121 days	0	-63.6 days	0.156	worsened
28	16 057	1 844	2010-04-07	296 days	3 596	186 days	0.058	1 102	193 days	0.089	-7 days	-0.031	unclear
29	617	19	2013-07-02	68 days	124	51 days	0.032	288	33 days	0.038	18.6 days	-0.006	unclear
30	249	11	2010-05-27	59 days	134	126 days	0.067	19	17 days	0.158	108.9 days	-0.091	improved
31	318	10	2013-02-18	104 days	29	96 days	0.172	216	27 days	0.056	68.7 days	0.117	improved
32	105 586	1 092	2012-05-23	248 days	9 252	141 days	0.094	51 504	60 days	0.044	80.6 days	0.05	improved
33	6 688	717	2013-10-16	93 days	3 344	58 days	0.051	2871	70 days	0.055	-12.5 days	-0.004	unclear
							*M*edian values for the 33 providers listed above:						
	2 212	225		151 days	530	99 days	0.077	274	68 days	0.055	15.9 days	0.015	

Table 1: Summary statistics for 33 hosting providers who had at least 10 URLs blacklisted before and after StopBadware shared data.

about 80% of URLs are cleaned up within 100 days whereas only about 70% of URLs before StopBadware contact are cleaned up within 100 days.

Despite the overall improvement, significant variation exists between individual hosting providers. We now look further into the effects of sharing on each individual AS. Figure 6 plots the per-AS survival plots, while Table 1 reports summary statistics split between pre- and post-contact. As before, we only include survival plots for the 33 ASes we reported to who also have at least 10 URLs blacklisted both before and after contact.

We notice that AS 12 took 15 days on average to clean URLs from the time they were blacklisted, good for the shortest time in our study. This is especially impressive given the provider's relatively large size (nearly 10k malware URLs observed). Before StopBadware shared URLs with AS 12, they averaged 122 days from blacklist to clean and 4.2% of cleaned URLs were later recompromised. However, after sharing, blacklist-to-clean time improved to 52 days. We note we shared URLs with AS 12 in 2010; we hypothesize the time of sharing could affect the response. By contrast, AS 20 did not improve their cleanup after sharing with them – they had an average of 509 days from blacklist to clean and were the worst AS in our study. We also shared data with AS 20 in 2010. The blacklist to clean time was larger after our report. However, the recompromise rate decreased by 6 percentage points after sharing data.

Looking at each AS individually shows the heterogeneity in the efficacy of reporting and the dimensions of what makes an effective cleanup strategy. ASes like AS 1 took fewer days to clean a URL (from blacklist time) after reporting, but had a higher recompromise rate. ASes like AS 8 took significantly more days after reporting, but had a lower recompromise rate. On one hand, it is quicker to clean a URL if additional time is not spent to additionally protect the URL; furthermore, recompromise rates might be outside the reach of the AS (for instance, if their customers use WordPress or other popular content management software). On the other hand, on a statistical level, the shorter time to clean the URL, the longer potential time to recompromise the URL. Furthermore, some ASes might achieve quicker cleanups by cutting corners and not eradicating the root cause of the compromise, e.g., by simply deleting files but not updating software or closing backdoors.

Table 2 shows summary statistics for these ASes in aggregate – 13 of these ASes had consistently improved cleanup after receiving a report from StopBadware, whereas 3 ASes worsened their cleanup trend. We labeled providers as *improved* if the survival probability is lower post-contact for more than 85% of the days observed. Similarly, a provider is labeled *worsened* if the survival probability is lower post-contact for fewer than 10% of the days. The progress for all other providers are labeled *unclear*. It is heartening that more ASes seem to improve than worsen, though most ASes do not exhibit a statistically clear trend. Particularly, we notice that the improved ASes have the best improvement in average blacklist to clean time (shortening by two months), whereas the worsened ones get half a year worse.

Furthermore, we found that improved hosting providers generally cleaned the URLs StopBadware reported to them faster. As can be seen in Figure 4, of the 13 improving providers, five clearly performed better than average on reported URLs, and none performed clearly worse. Of the three worsening providers, two performed clearly worse than average. Of the remaining 17 providers, for whom data sharing had an unclear long-term effect, six clearly performed better than average and one did worse on cleaning up reported URLs. From this analysis, we conclude that there may be a link between the performance of providers in cleaning up data from bulk reports and their long-term performance. We explore that possibility in greater detail next.

Figure 7 examines the interactions between our metrics on how hosting providers clean up malware. We hypothesized that these all would be correlated: that hosting providers

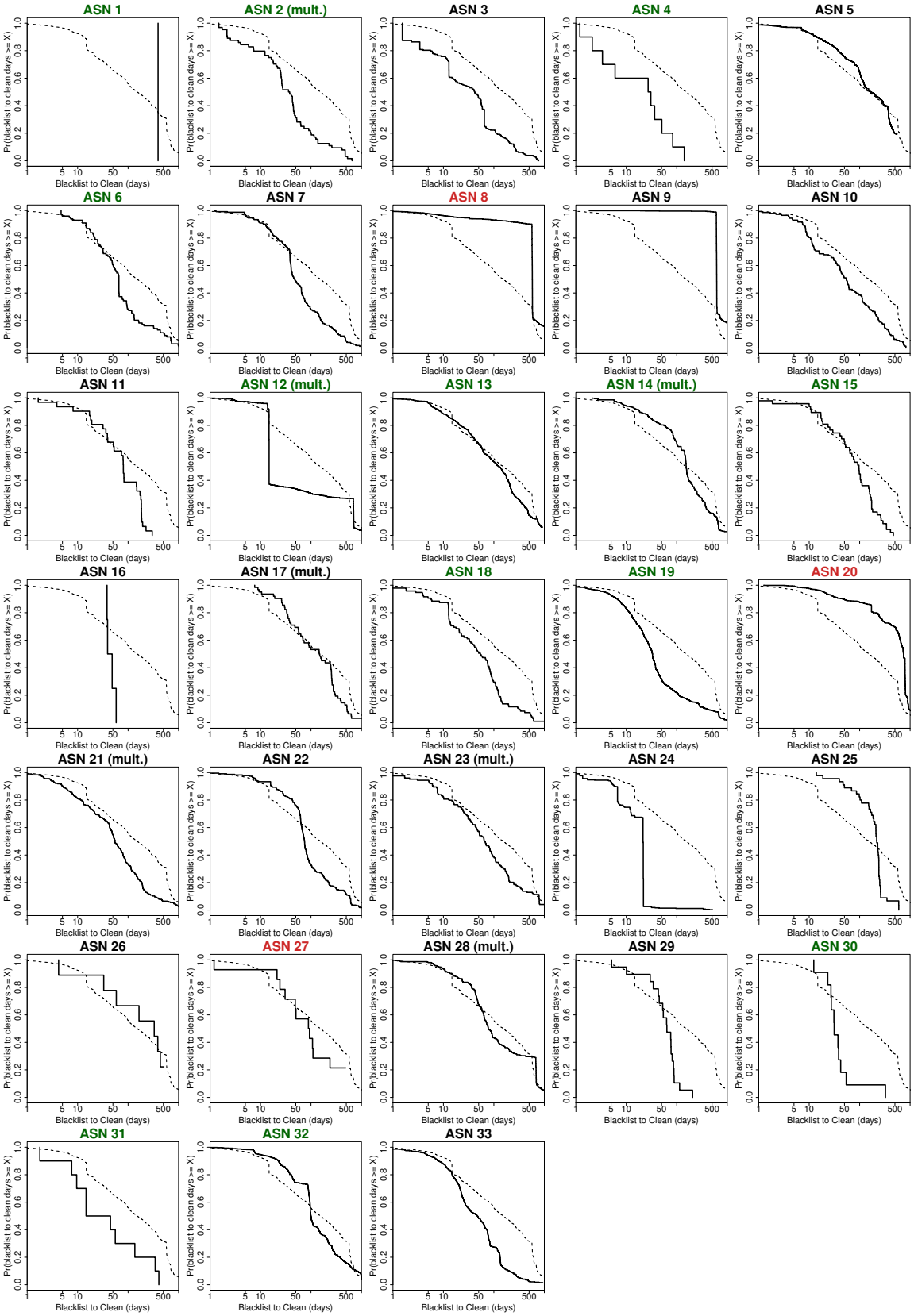

Figure 4: Survival probability for the time from reporting to clean per hosting provider. Figures are titled green if the hosting provider improved after contact, and red if they worsened. Dotted line indicates the average.

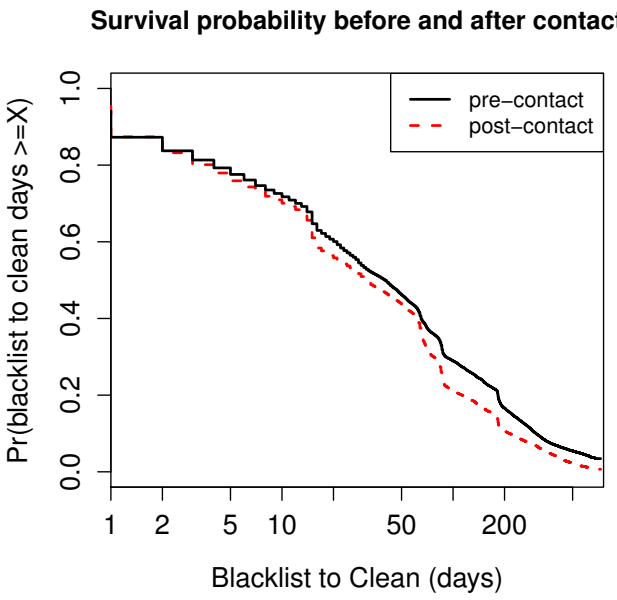

Figure 5: Survival probability for the number of days from blacklisting to clean for URLs pre- and post-contacting the host (all reported-to hosting providers).

	#	Δ days to clean	Δ recomp. rate
Improved	13	58	0.010
Worsened	3	-176	0.085
Unclear	17	13	0.008

Table 2: Comparing cleanup rate and recompromise rate on the 33 reported-to ASes with 10 or more URLs ever blacklisted on them.

who cleaned up the malware we sent them would also improve their cleanup rate and lower their recompromise rate after a report.

Figures 7[a] and 7[b] explore the relationship between efforts to clean shared URLs and the long-term metrics just presented. The vertical axis in both graphs is the median report to clean time for the shared URLs. The point size is proportional to the number of URLs shared. We change the point icon and color based on the year the report was sent, since trends in sharing (as well as size of malware infections) could change over time. In Figure 7[a] the horizontal axis is difference in median blacklist to clean time before and after sharing, where each time period is cut to no more than two years before/after a malware report. In Figure 7[b] the horizontal axis is the difference in median recompromise rate before and after sharing. In both cases, positive numbers indicate improvement (i.e., the median blacklist to clean time has gotten shorter).

Figure 7[a] exhibits an approximately linear downtrend. This suggests that there is a correlation between the speed of cleanup in response to shared URLs and the improvement (or lack thereof) in cleaning up malware for the period after sharing takes place. By contrast, Figure 7[b] shows no discernible trend. This (discouragingly) shows that the recompromise rate is largely orthogonal to a provider's responsiveness to cleaning URLs that have been shared.

Figure 7[c] compares the interaction of both long-term metrics, in hopes of demonstrating whether or not hosting providers improved their malware incident response over the long haul after they received assistance from StopBadware. The horizontal axis plots the change in the median survival blacklist to clean time from the period before to the period after sharing has taken place. Positive numbers indicate improvement (i.e., the median blacklist to clean time has gotten shorter). The vertical axis plots the change in the recompromise rate after sharing. Again, positive numbers indicate improvement. Points are scaled by the number of URLs shared, and they are color-coded according to their relative performance in the report to clean time for shared URLs. Overall, it is striking that most hosting providers improve on at least one of the two measures – only two providers appear in the lower left quadrant, indicated that their performance worsened on both measures. Many providers improved on one measure but not both (top left and bottom right quadrants). The top-performing hosting providers appear in the top right quadrant.

We can see that those providers that responded most quickly to the reports from StopBadware also tended to improve their time to clean long afterwards (most of the top quartile appears in the right quadrants). The top performers over the long term tended to process a smaller number of reports, whereas those processing more reports were more likely to reduce either the recompromise rate or the median blacklist to clean time, but not both.

4.3 Matched Pair Analysis

Ideally, we would be able to compare the world where we reported to an AS against the world where we did not in order to isolate the direct effect of reporting bulk URLs to ASes. However, this is obviously not possible. Instead, we attempt to replicate this approach by matching each reported-to AS to a sister AS. These matched pairs have a similar level of compromise on each AS's reporting date and are located in the same country[4]. We assume that without our direct intervention, each AS in the pair would have the same compromise level. Thus, significant differences in trends between the reported-to ASes and their sister (matched pair) ASes indicates an effect from reporting.

First we compare the report to clean time from URLs blacklisted in our reported-to ASes with the URLs blacklisted in their sister ASes on the day we reported to the ASes (Figure 8). For the sister ASes (where no report was actually issued), we compute the report to clean time by identifying all URLs that would have been shared had a report been requested and measuring the time to clean starting from the day sharing would have occurred.

We see that URLs that were reported (black line) were cleaned up quicker than URLs that were not (dashed blue line). Half of the reported URLs were cleaned up within 100 days whereas half of the matched pair URLs were cleaned up in over 500 days. Additionally, the recompromise rate for shared URLs is 4% whereas for matched pair URLs have a recompromise rate of 10%. This provides clear evidence that

[4]In one case, the AS is the only one in the country with malware; here we use the AS with a similar level of malware that is based in the United States instead.

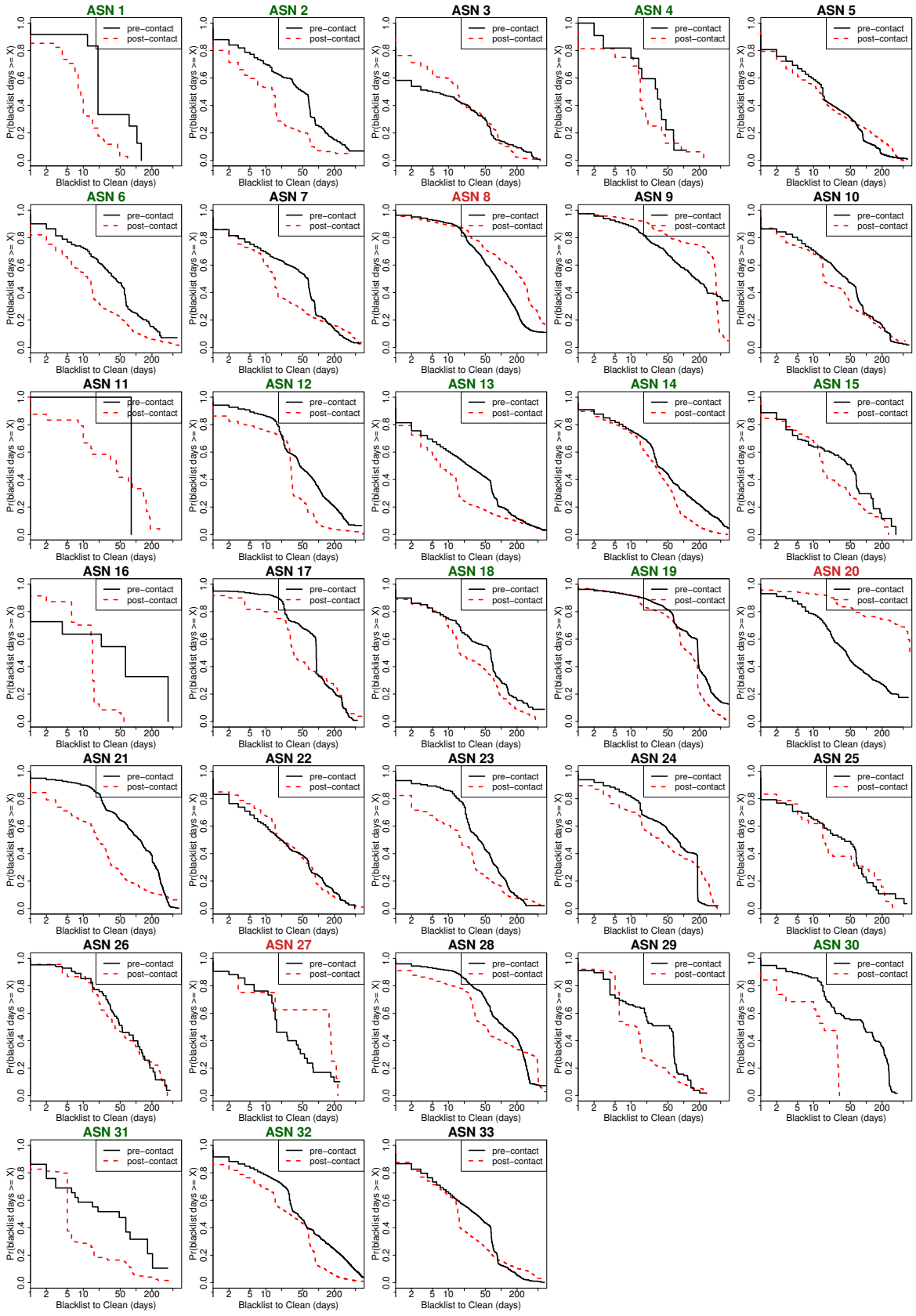

Figure 6: Survival probability for the time from blacklist to clean, for the two years before and after contact with StopBadware. Figures are titled green if the AS improved after contact, and red if they worsened.

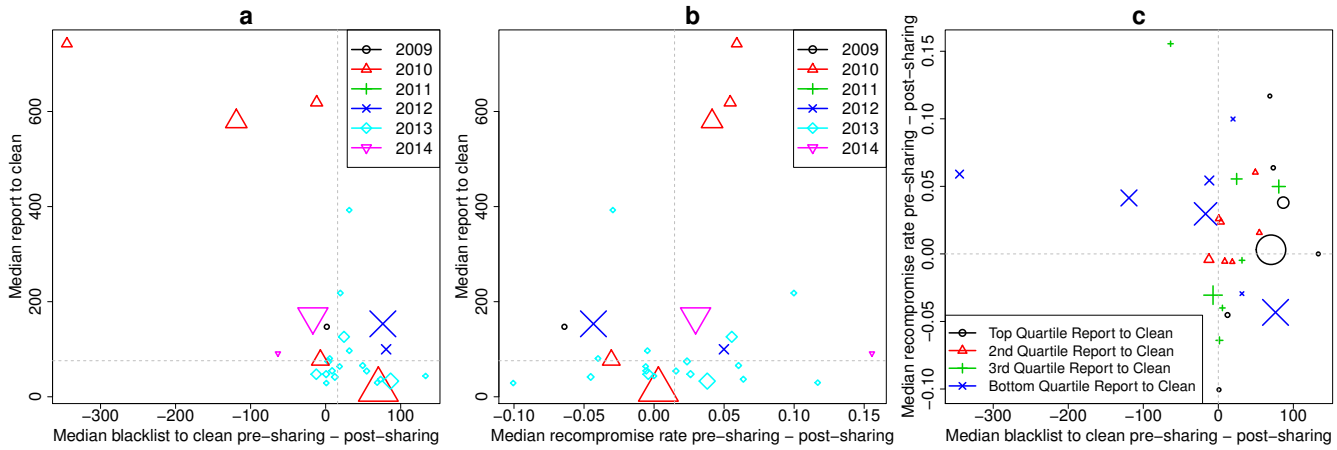

Figure 7: Interactions between clean measures: comparing change in the blacklist to clean time, change in the recompromise rate, and the report to clean time for StopBadware reported URLs. Points are scaled by the number of URLs shared (more URLs shared for larger points).

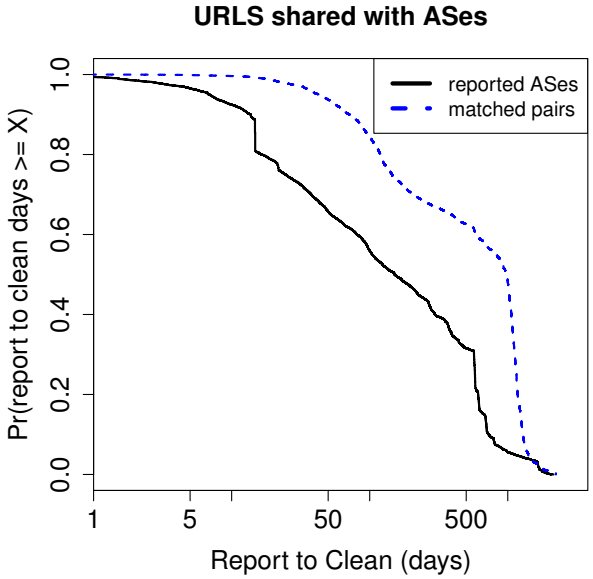

Figure 8: Survival probability for the number of days from report to clean for the reported URLs compared to the survival probability for URLs in the matched pair ASes

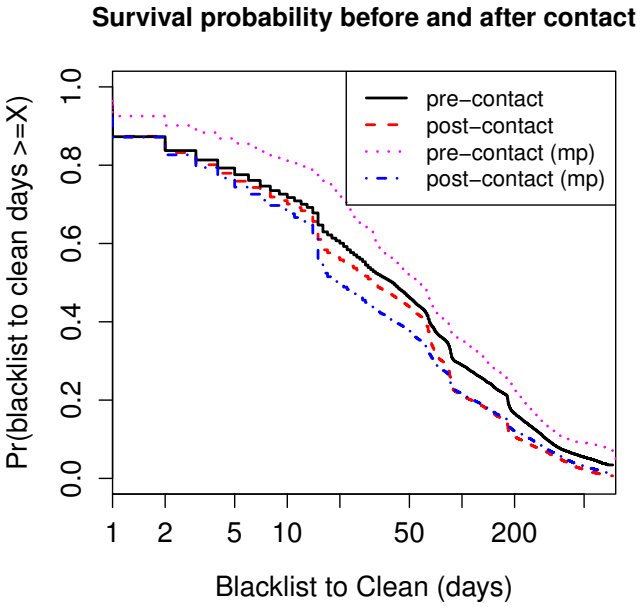

Figure 9: Survival probability for the blacklist to clean time for URLs pre- and post-contacting the host (all hosting providers).

reporting has a large direct effect on improving the cleanup times for the URLs that are shared.

But what about the long-term indirect effects of sharing? Figure 9 compares the blacklist to clean time before and after contact for both the reported to ASes (exactly like Figure 2) and their sister ASes. We see that the pink dotted line (sister ASes pre-contact) is much higher than the black solid line (reported-to ASes pre-contact). This indicates that before the date StopBadware sent reports to the ASes, their sister ASes took a longer time to clean up a URL. In the period of time after the notification, the reported-to ASes (red dashed line) are about as effective at cleaning malware as their sister ASes (blue dot dashed line).

This analysis suggests that the long-term improvement observed after contact may be caused by some other factor besides receiving malware reports from StopBadware. Nonetheless, individual variation by AS (e.g., among the ASes that significantly improved performance like those appearing in the top right quadrant of Figure 7[c] may be affected by sharing data. Further investigation is needed.

5. CONCLUDING REMARKS

The analysis in this paper helps quantify the impact of sharing abuse data with interested providers. It demonstrates how the responses even from well-intentioned, proactive operators can vary widely. We found that providers

cleaned up most of URLs that we directly reported. Even though the cleanup could still take weeks or many months, we are confident that these reports helped the remediation of the shared URLs. However, the evidence that these one-off reports helped improve security over the long term or reduce the prevalence of malware is weak. On the one hand, we should be encouraged that sharing data can clearly improve outcomes for some providers. On the other hand, it is concerning that many providers do not improve after receiving actionable abuse data. Thus, the paper at once demonstrates the potential and the limits of sharing security data between private actors.

But, should anything be done to change this? Abuse can be hard to find. Furthermore, many operators are skeptical when others claim to have found problems on their networks. This wariness combined with attackers' effective cloaking techniques makes it hard for outsiders such as StopBadware or Google to repeatedly verify that an infection is present.

Web-based malware affects the involved stakeholders quite differently. Operators who spend money hiring technical staff or professionals to clean abuse have vastly different capabilities, incentives and experiences with malware than the general public who visit the offending page. This means that the costs of cleaning up malware are often borne by those not directly harmed by it, prompting less abuse to be cleaned than would be optimal for society. Informing customers and providers without internalizing the costs has not been shown to be effective.

Much of the pernicious abuse studied in this paper and shared with operators were malware URLs that have remained operational for many months or even years. A misalignment of incentives among operators has allowed such malware to stick around far longer than even the criminals would have hoped for. Many current efforts in cleaning networks concentrates on the most recent infections, ignoring such long-standing abuse.

One final takeaway from this paper should be a call to investigate ways to clean up this long-lived abuse. In addition to concentrating efforts on recently compromised URLs, operators should also attempt to clean websites that have been compromised for years. New approaches are needed for these hard cases. Perhaps the dearth of technical information in languages other than English or the lack of resources for non-technical website operators on low cost shared web hosts is finally catching up to us. At the very least, more investigation into why some compromises persist is needed.

We found that sharing bulk data on blacklisted malware URLs with hosting providers was, on average, helpful. Yet bulk malware lists like those that StopBadware shares include lots of long-lived abuse. In order to make bulk data sharing more effective, we need to figure out how to eradicate all compromises, not only the new ones.

Acknowledgments

We thank StopBadware for sharing data on its sharing practices. We also thank StopBadware's Clearinghouse data providers (Google, ThreatTrack Security, and NSFocus).

This publication was supported by a subcontract from Rutgers University, DIMACS, under Award No. 2009-ST-061-CCI002-06 from the U.S. Department of Homeland Security and by a grant from the Netherlands Organisation for Scientific Research (NWO), under project number 628.001.022.

6. REFERENCES

[1] Canali, D., Balzarotti, D., and Francillon, A. The role of web hosting providers in detecting compromised websites. In *International World Wide Web Conference* (2013), pp. 177–188.

[2] Cetin, O., Jhaveri, M. H., Gañán, C., van Eeten, M., and Moore, T. Understanding the role of sender reputation in abuse reporting and cleanup. In *Workshop on the Economics of Information Security* (2015).

[3] Durumeric, Z., Kasten, J., Adrian, D., Halderman, J. A., Bailey, M., Li, F., Weaver, N., Amann, J., Beekman, J., Payer, M., and Paxson, V. The matter of Heartbleed. In *Internet Measurement Conference* (2014), ACM, pp. 475–488.

[4] Hutchings, A., Clayton, R., and Anderson, R. Taking down websites to prevent crime. In *APWG Symposium on Electronic Crime Research* (June 2016).

[5] Kührer, M., Hupperich, T., Rossow, C., and Holz, T. Exit from Hell? Reducing the Impact of Amplification DDoS Attacks. In *Proceedings of the 23rd USENIX Security Symposium* (August 2014).

[6] Li, F., Durumeric, Z., Czyz, J., Karami, M., Bailey, M., McCoy, D., Savage, S., and Paxson, V. You've got vulnerability: Exploring effective vulnerability notifications. In *USENIX Security Symposium* (Aug. 2016).

[7] Li, F., Ho, G., Kuan, E., Niu, Y., Ballard, L., Thomas, K., Bursztein, E., and Paxson, V. Remedying web hijacking: Notification effectiveness and webmaster comprehension. In *International World Wide Web Conference* (Apr. 2016).

[8] Moore, T., and Clayton, R. The consequence of non-cooperation in the fight against phishing. In *eCrime Researchers Summit* (2008), IEEE, pp. 1–14.

[9] Nappa, A., Rafique, M. Z., and Caballero, J. Driving in the cloud: An analysis of drive-by download operations and abuse reporting. In *International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment* (July 2013), Springer.

[10] Provos, N., Mavrommatis, P., Rajab, M. A., and Monrose, F. All your iFRAMEs point to us. In *USENIX Security Symposium* (2008).

[11] Stock, B., Pellegrino, G., Rossow, C., Johns, M., and Backes, M. Hey, you have a problem: On the feasibility of large-scale web vulnerability notification. In *USENIX Security Symposium* (Aug. 2016).

[12] Vasek, M., and Moore, T. Do malware reports expedite cleanup? An experimental study. In *USENIX Workshop on Cyber Security Experimentation and Test* (Aug. 2012).