Empirical Analysis of Factors Affecting Malware URL Detection

Marie Vasek
Department of Computer Science and Engineering
Southern Methodist University
Dallas, Texas, 75275, USA
Email: mvasek@smu.edu

Tyler Moore
Department of Computer Science and Engineering
Southern Methodist University
Dallas, Texas, 75275, USA
Email: tylerm@smu.edu

***Abstract*—Many organizations, from antivirus companies to motivated volunteers, maintain blacklists of URLs suspected of distributing malware in order to protect users. Detection rates can vary widely, but it is not known why. We posit that much variation can be explained by differences in the type of malware and differences in the blacklists themselves. To that end, we conducted an empirical analysis of 722 malware URLs submitted to the Malware Domain List (MDL) over 6 months in 2012–2013. We ran each URL through VirusTotal, a tool that allowed us to check each URL against 38 different malware URL blacklists, within an hour from when they were first blacklisted by the MDL. We followed up on each for two weeks following. We then ran logisitic regressions and Cox proportional hazard models to identify factors affecting blacklist accuracy and speed. We find that URLs belonging to known exploit kits such as Blackhole and Styx were more likely to be blacklisted and blacklisted quicker. We also found that blacklists that are used to actively block URLs are more effective than those that do not, and furthermore that paid services are more effective than free ones.**

I. Introduction

Cybercrime defenses are increasingly data-driven. Antivirus software updates frequently with the latest binaries, software patches that plug vulnerabilities are automatically disseminated, and web browsers automatically block newly reported phishing websites. Another important defense is the malware URL blacklist, which tracks a range of dangerous websites, from hacked websites that distribute fake antivirus software to maliciously registered websites that host payloads. In each case, the timeliness of the data underlying the mechanism is crucial. Criminals pollute the search results for trending topics with links to sites delivering malware [1], [2], where delays of just a few hours can lead to many thousands of infections. Security engineers have developed sophisticated mechanisms to rapidly identify web-based malware rapidly (e.g., [2], [3]). Yet some criminals still manage to evade detection, at least for a few hours or days.

In this paper, we describe the results of an empirical study into the factors that affect the timeliness and comprehensiveness of malware URL blacklists. Using URLs submitted to the Malware Domain List, we repeatedly test whether and when these URLs are added to dozens of blacklists. We present evidence that attributes of the malware itself can influence the accuracy and timeliness of the blacklist. Moreover, attributes of the blacklist itself (notably if it is fee-based and if it actively blocks URLs) also have a measurable effect.

Section II details the methodology we adopted for collecting the data as well as the factors influencing detection that we investigated. Section III describes summary statistics of the gathered data, while Section IV presents a regression and survival analysis to more rigorously investigate the relationship between the identified factors and detection rates. We discuss related work in Section V before wrapping up in Section VI.

II. Methodology

A. Data collection approach

Our goal is to study the blacklisting of malware URLs over time by different services in order to analyze the factors that affect malware blacklisting. Many malware URL blacklists are dynamic, so one measurement does not fully capture the nature of blacklisting. Furthermore, we anticipate that the greatest change in blacklisting status will appear within the first few days.

To that end, we examined the stream of malware URLs from the Malware Domain List (MDL) [4], a community-driven, publicly accessible blacklist set up by malware researchers. Anybody with an account on their website can submit a URL to the list, though all URLs are verified by the community before being officially added. When a URL is submitted to the MDL, the security professionals who are active in the MDL community try to take them down; to facilitate takedown, they include additional information (such as IP address, WHOIS information, and a description of the malware) along with the URL in their interface. Because we want to analyze the initial trend in blacklisting URLs, we restricted ourselves to the new URLs starting on the first day of our collection. So even though the MDL has blacklisted over 86 000 URLs since January 2009, we only included the 722 malware URLs (but not the one phishing URL which we discarded) that were added to the MDL between December 2, 2012 and May 29, 2013.

To check for blacklisting, we ran each URL through VirusTotal [5] within an hour of it being put on the MDL. VirusTotal is a Google-owned service that evaluates suspected malware binaries and URLs against multiple antivirus (AV) engines as well as non-AV malware/phishing blacklists. VirusTotal gets updates every 15 minutes from the services that is collates. We did not use the malware binary features of VirusTotal, since our paper concentrates purely on malware URLs. We then checked each URL every other hour through VirusTotal for the first 48 hours and every day thereafter for two additional weeks (for a total of 16 days' coverage). We do this to check against our hypothesis that more action takes place during the first 48 hours after reporting. We stopped our collection after 16 days because the blacklisting state seemed to be constant after that.

All of the URLs submitted to VirusTotal are shared with participants; if a URL has at least one positive result then it is sent to all participants but if a URL is not marked positive by any service, it is sent out to premium participants. Thus, submitting these URLs to VT might have side effects which would lead to faster blacklisting by these services than normal. This suggests the detection rates we recorded represent an upper bound for the effectiveness of blacklists.

B. *Factors affecting malware URL detection*

We hypothesize that much of the variation in malware URL detection accuracy and speed is caused by factors other than what particular blacklisting service is used. We devise two categories of explanatory variables that might affect the blacklisting of a URL by a specific service. We group these variables into those that are characteristics of a URL and those related to the malware blacklisting service.

Note that we seek to analyze the different factors that lead URLs to be picked up by AV services. Our paper does not seek to compare the *effectiveness* of AV services against each other; to do so is completely out of the scope of this paper.

1) URL variables: We look at a range of different properties of a URL that could influence whether it would likely be blacklisted or not.

IP address: This variable is true if the URL has no domain name, but rather just consists of an IP address, e.g. `http://78.110.62.95/jentrate.php`. We hypothesize that these URLs are less likely to be detected and detected later than other URLs because blacklisting an IP could potentially also blacklist legitimate URLs hosted on the same IP address.

Has a Path: This variable reflects whether the URL has a path, e.g. `http://askmeaboutcctv.com/wmiq.html`. We would expect URLs that don't have paths are more likely to be set up by malicious actors and thus would be more likely to be blacklisted, since there would be no incidental damage to legitimate websites.

Executable: This variable is true if the file ends with .exe, presumably an executable file, e.g. `http://euxtoncorinthiansfc.co.uk/1689.exe`. Since executable files are generally not infected URLs and easier to check for maliciousness, we would expect them to be blacklisted earlier and more likely to be blacklisted at all.

Fake AV This variable is true if the URL is found to be part of a rogue security software campaign (Fake AV). These URLs have paths which follow patterns such as `/index/two/` and `/?affid=*&promo_type=*&promo_opt=*` [6]. We would expect these URLs to be detected earlier but less likely to be detected, since Fake AV exhibits more cloaking behaviors compared to other forms of malware.

Styx: This variable is true if the URL has a hex-encoded path directory in its path, e.g. `http://masterpeaceloves.com:8888/L7kU3T0ZD6X04aKh0UDer0by3Q0F0JX0HaJQ0Giot0hL7K04R7k0Rg0O08DW00jK5j/`. All the URLs with this feature in our MDL collection are associated with the Styx Exploit Kit [7]. We would expect these URLs to be detected earlier than other types of malware because this exploit kit has been around since mid 2012 and the URLs look the same.

Blackhole Landing Page: This variable marks URLs which are landing pages for Blackhole exploit kits, such as `http://gimiinfinfal.ru:8080/forum/links/column.php`. We identified these by looking at the structure of the URL since these URLs are constructed formulaically. We also corroborated with http://urlQuery.net/, a free URL scanner that cross references various intrusion detection systems on top of their own analysis, and the metadata given from the MDL. As with the Styx exploit kit, we would expect these URLs to be more likely to be detected and be detected earlier because this exploit kit has been around for 3 years. We also expect these URLs to be more likely to be detected than Styx URLs since the Blackhole exploit kit is more popular and has been around longer.

Other: We compare executables, fake AV URLs, Styx URLs and Blackhole landing pages to 282 "other" URLs. These include

- domains with no paths,
- malware receipt spam URLs,
- generic-looking malware URLs,
- other low-frequency exploit kits (Impact, Propack, Sweet Orange,...) and traffic detection system URLs.

2) Malware Blacklist variables: We also identified two different characteristics of malware blacklists that could affect both their quality and their tolerance of false positives.

Blocks Users: This variable is true if the malware blacklist blocks users from accessing the URLs on its blacklist in some form (e.g. Google Safe Browsing, AV blacklists). We expect malware blacklists that block users to be less aggressive compared to other blacklists, since a false positive harms a domain more if it is blocked.

Costs Money: This variable reflects whether the malware blacklist service in the same form that it is provided to VirusTotal costs money to use. This is true for services that offer an *n*-day free trial and services that only offer access to the blacklist wrapped in paid services, such as antivirus software. We would expect services that cost money to have more resources, thereby blacklisting more URLs and blacklisting them quicker.

III. Summary Statistics

A. *Coverage*

As a first step, we can examine each blacklists' coverage in ever detecting reported URLs as malicious. Figure 1 plots the percentage of MDL URLs that are ever detected as malicious by each list, shown in sorted order from most to least comprehensive. A few things stick out. First, we observe that six services do not detect a single URL. Second, at the

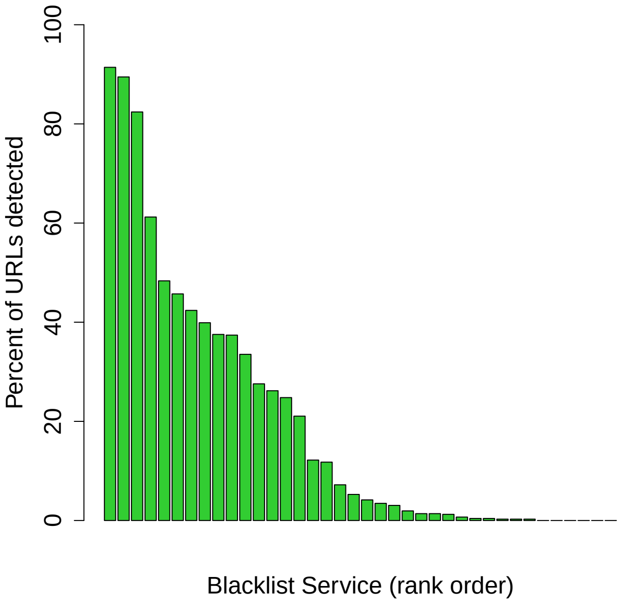

Fig. 1: Percentage of URLs ever detected by service.

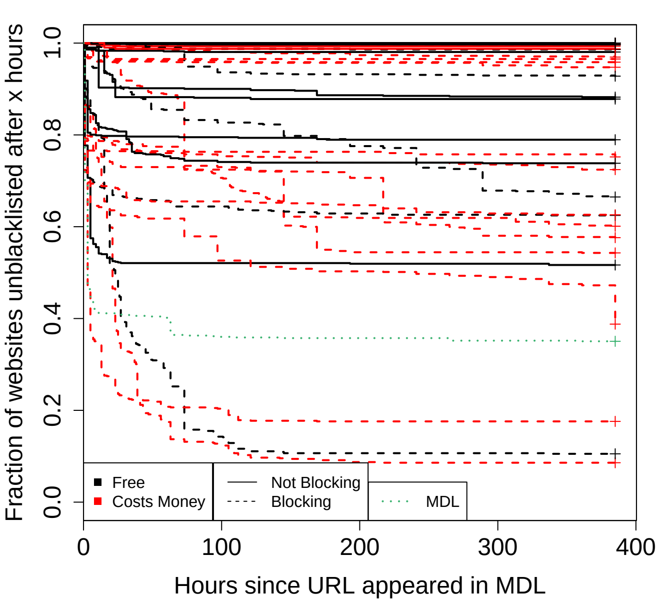

Fig. 2: Survival probability by service.

other extreme, three malware blacklists are so successful at blacklisting MDL URLs (greater than 80% detection) that we suspect that they use MDL submissions as input to their own blacklists. Other researchers have also found evidence that the MDL is used by some services to flag malware URLs [8].

The difference between the three top-performing services is further demonstrated by examining the time delays between reporting in the MDL and detection by the service. Figure 2 plots the survival probability for a given URL to be detected by various services since the URL first appears in the MDL. Most services detect a substantial fraction of websites shortly after being reported to the MDL (from 10–40%). However, the top three services are in a league of their own, detecting 60–80% of URLs as malicious within a day or so of appearing in the MDL. Furthermore, we observe that VirusTotal reports these three services as having better coverage of the URLs than the MDL itself! While puzzling, this makes sense considering the timing of it all. VirusTotal checks the MDL every 15 minutes. We check the MDL every hour. If we check the MDL before VirusTotal, it makes sense that a URL is not intitially found bad. VirusTotal only uses the "active" URLs on MDL. If the next time a URL is rescanned by VirusTotal MDL has already marked it inactive, then the URL will not be reported as malicious. We also note that the MDL's reports of activity do not line up with the reports of activity from the any other service in the collection. Some of the noise can also be attributed to the fact that VirusTotal is not perfect and does not scan every URL through every service every time.

Given the evidence that some services track the MDL, we elected to remove reports from these blacklisting services from further analysis. Additionally, we remove the 6 blacklisting services that did not blacklist any of the URLs on the MDL at any time during our study.

Figure 2 also shows how the survival probabilities differ by encoding characteristics of the services. We can see that, overall, the services that cost money and those that block tend to perform better. However, there are some notable exceptions, such as the free non-blocking service Clean MX that detects around a third of the URLs on the first day.

After looking at each services' coverage of the URLs, we next take the perspective of the URLs being blacklisted by the services. Figure 3 plots the fraction of URLs that are detected by a number of services as a function of time. For example, the red line in the graph shows that 80% of URLs are detected by at least one malware blacklist service within a few hours. The trouble is figuring out which one! No single service performs nearly that well at detecting malware URLs. Moreover, even after a week or more around 5% of URLs are never detected.

The other lines show the corresponding reduction in detection as more blacklists detect the URLs. For example, around 60% of URLs are detected within 24 hours by at least two blacklists, 50% by three and 20% by six. One final point worth noting is that detection rates tend to plateau. After a week or so, diminishing returns set in for detecting malware URLs.

Finally, Figure 4 plots the average number of blacklist hits for a URL from MDL over time. Again, this shows diminishing returns over time. However, what is especially noteworthy is that the number of detections is not strictly monotonically increasing. This hints at the dynamic nature of blacklisting, as URLs deemed malicious one day may be cleaned (or hidden from a malware checker) the next.

B. Malware URL factors

In Section II-B we described several of the factors hypothesized to affect detection by URL blacklists. We now report on

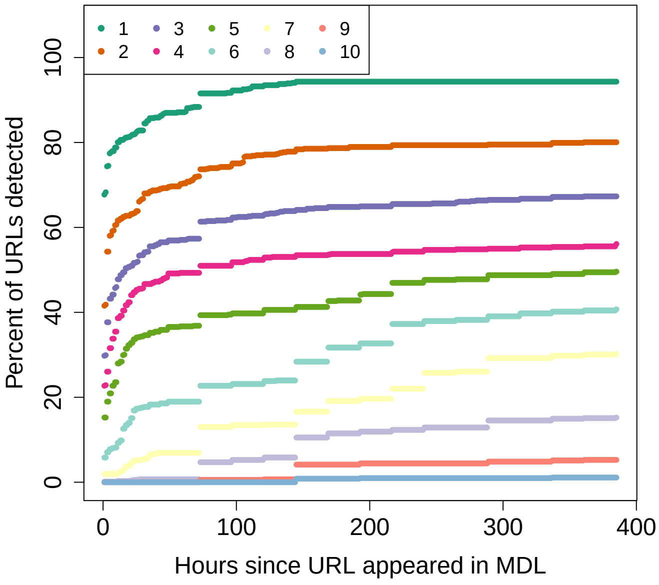

Fig. 3: Percentage of URLs detected by at least n services.

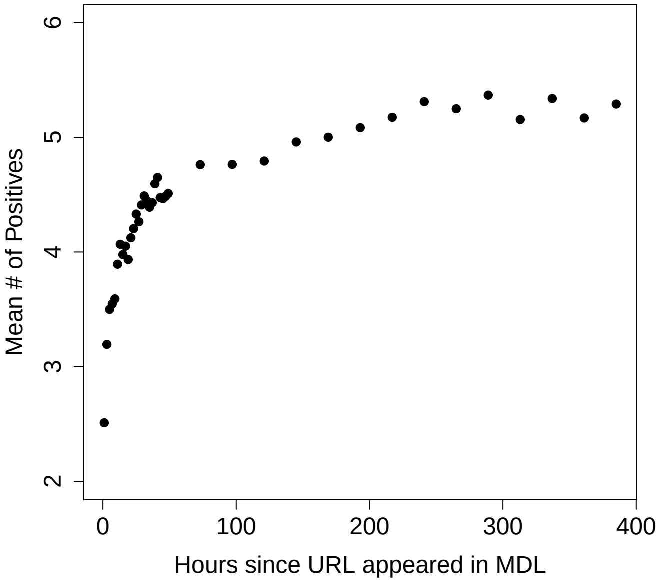

Fig. 4: Blacklisting over time.

Malware Type	#	%	IP/Domain	#	%	Path?	#	%
Executable	175	24%	IP Address	124	17%	Has Path	675	93%
Fake AV	65	9%	Domain	598	83%	No Path	47	7%
Styx	51	7%						
Blackhole Lnd.	149	21%						
Other	282	39%						

TABLE I: Frequency of different malware URL characteristics.

Blocks?	#	%	Costs?	#	%
Blocks Users	22	58%	Costs Money	17	45%
Doesn't Block	16	42%	Free	21	55%

TABLE II: Frequency of different blacklist characteristics.

their relative incidence in the MDL dataset. Table I reports the URL characteristics we identified. From the first column, we see that around a quarter of submitted URLs are executables, while 9% are clearly fake antivirus and another 28% are one of the two leading exploit kits on the MDL. This leaves 39% of URLs unclassified, either because they are unrelated or (frequently) because it can be difficult to identify the type of malware from the URL patterns alone. The vast majority (83%) of URLs include domains, while 93% also include paths in the URL rather than the second-level domain name alone.

Table II, by contrast, reports on two key observable characteristics of the blacklists themselves. First, we see that the majority of blacklists actively block suspected malware URLs to protect their customers. Second, a slight majority of blacklists are given away as free services.

IV. Empirical Analysis

We now describe the results of two related statistical models. First, we present a logistic regression examining the influence of factors on *whether* an URL is ever blacklisted. Second, we present a Cox proportional hazards model examining the influence of factors on *when* a URL is blacklisted.

A. Logistic Regression

We use logistic regression to look at the influence of our variables on whether a URL is blacklisted or not by a given blacklist. We use a logistic model because the dependent variable is a boolean (blacklisted or not blacklisted). The model takes the following form:

$$\begin{aligned}\log\frac{p_{BL}}{1-p_{BL}} = &\beta_0 \; + \beta_1 \text{IP address} + \beta_2 \text{ Has Path}\\ &+\beta_3 \text{ Executable} + \beta_4 \text{ Fake AV}\\ &+\beta_5 \text{ Styx} + \beta_6 \text{ Blackhole}\\ &+\beta_7 \text{ Blocks?} + \beta_8 \text{ Costs Money} + \varepsilon\end{aligned}$$

In this model, $\log\frac{p_{BL}}{1-p_{BL}}$ is the log odds that a URL will be blacklisted by a given blacklist, β_0 is the constant interval rate, $\beta_1, \beta 2, \beta_3, \beta_4, \beta_5, \beta_6$ are best-fit constants for the URL variables, β_7, β_8 are the best-fit constants for the blacklist variables, and ε is the error term. Table III shows the best-fit coefficients, odds ratios and significance.

IP addresses are blacklisted at no different rate compared to domains. We can see that the type of URL matters. Executables are 2.8 times more likely to be blacklisted compared to uncategorized URLs, all else equal. Styx exploit kit URLs are 2.1 times more likely to be blacklisted and Blackhole landing page URLs are 1.2 times more likely to be blacklisted compared to uncategorized URLs. We attribute the higher blacklisting rate to the relative ease with which these kits can be identified from their URL structure.

	Coefficient	Odds Ratio	p value
Intercept	-2.304	**0.100**	0.000
URL features			
IP address	-0.004	0.996	0.945
Has a Path?	-0.196	**0.822**	0.031
Executable	1.017	**2.765**	0.000
Fake AV	-0.838	**0.433**	0.000
Styx	0.746	**2.109**	0.000
Blackhole Landing Page	0.196	**1.217**	0.000
Malware Blacklist Features			
Blocks Users?	0.611	**1.843**	0.000
Costs Money	0.298	**1.347**	0.000
$\chi^2 = 1144.501$, p value = 0.000			

TABLE III: Table of coefficients for logistic regression

	Coefficient	Odds Ratio	p value
URL features			
IP address	0.056	1.058	0.210
Has a Path?	-0.207	**0.811**	0.012
Executable	0.896	**2.449**	0.000
Fake AV	-0.814	**0.443**	0.000
Styx	0.750	**2.118**	0.000
Blackhole Landing Page	0.179	**1.196**	0.000
Malware Blacklist Features			
Blocks Users?	0.538	**1.713**	0.000
Costs Money	0.300	**1.351**	0.000
$R^2 = 0.055$			

TABLE IV: Table of coefficients for survival regression.

By contrast, Fake AV URLs are 0.4 times as likely to be blacklisted compared to "other" URLs. One explanation for this could be that the particular variants we identified are especially effective at hiding their behavior to investigators (e.g., through cloaking and only infecting the first page-view to an IP address).

Both malware blacklist features are positive and statistically significant. This means that blacklists that block users are 1.8 times as likely to blacklist a URL compared to blacklists that do not. This result runs counter to our expectations. We had thought that blacklists blocking URLs would take a more cautious approach, leading to a less comprehensive blacklist. That they in fact caught more URLs suggests that the blacklists could be more confident in their assessments.

We also found blacklists that cost money are 1.3 times as likely to blacklist a URL compared to blacklists that are free. This result is more expected, given that more expensive goods should provide better service.

B. Survival Analysis

In our logistic model from Section IV-A we only consider whether a URL is blacklisted or not by a particular blacklist. However, timing also matters. Many of the URLs in the MDL are part of spam campaigns, so blacklisting a URL a week after the campaign is less useful than blacklisting that same URL at the time of the campaign. We used a proportional hazards model using the time in hours to blacklist a URL as the response variable [9].

We look at the hazard rate $h_{ij}(t)$ of URL i using service j, where β_1, β_2, β_3, β_4, β_5,β_6 are the constant coefficients for our URL variables and β_7, β_8 the constant coefficients for our malware blacklist variables.

$$\begin{aligned} h_{ij}(t) =& h_0(t)\exp(\beta_1 \text{IP address}_i + \beta_2 \text{Has Path}_i \\ &+ \beta_3 \text{Executable}_i + \beta_4 \text{Fake AV}_i \\ &+ \beta_5 \text{Styx}_i + \beta_6 \text{Blackhole}_i \\ &+ \beta_7 \text{Blocks?}_j + \beta_8 \text{Costs Money}_j) \end{aligned}$$

The results are similar to the logistic regression. We find that the exploit kits are blacklisted more quickly, while fake antivirus takes longer to block. Once again, executables are the easiest to fend off. Meanwhile, blacklists that block URLs and those that cost money are also detecting faster.

Figure 5 visualizes the survival probabilities predicted using the proportional hazards model. For each graph, we plot the survival probability varying only one of the factors while holding all the others at their mean value. For example, in the top left figure we see that fake antivirus URLs are hardest to blacklist, with over 95% remaining unblacklisted for two weeks. Executables are blacklisted most quickly, with around 20% blacklisted within 24 hours holding all other factors constant, rising steadily to around 27% by the end of the monitoring period. The Styx exploit kit is not far behind, while the survival probability for the Blackhole exploit kit over time is only slightly worse than for uncategorized malware.

We see much less variation between when an IP address is used in a URL compared to a domain. This is not surprising given that the difference is not statistically significant. Similarly, URLs without a path are blacklisted slightly more quickly than those with one (though this difference is statistically significant, as reflected in Table IV).

Finally, the bottom right graph examines the effect on survival probability of the blacklist-level characteristics. Regardless of the malware characteristics, blacklists that cost money and actively block URLs blacklist URLs much quicker than those that are free and do not block URLs. Blocking has the bigger effect, though, since free services that block blacklist quicker than paid services that do not also block.

V. Related Work

Sheng et. al. analyzed phishing blacklists using fresh phishing URLs [10]. They compared time to blacklist against time for the phishing campaign to finish, finding that 72% of the phishing URLs were blacklisted within 48 hours of their initial check by an average phishing blacklist service, whereas only 2% were blacklisted before their initial check. They also consider the role of using heuristics in blacklisting phishing URLs and found them to be effective in protecting users and resilient to false positives.

A number of previous works have used multiple antivirus products to more effectively detect malware binaries. CloudAV [11] uses 12 antivirus products to catch more malware binaries than a single AV product. This work also analyzes the detection rate as a function of time and, like our work, shows

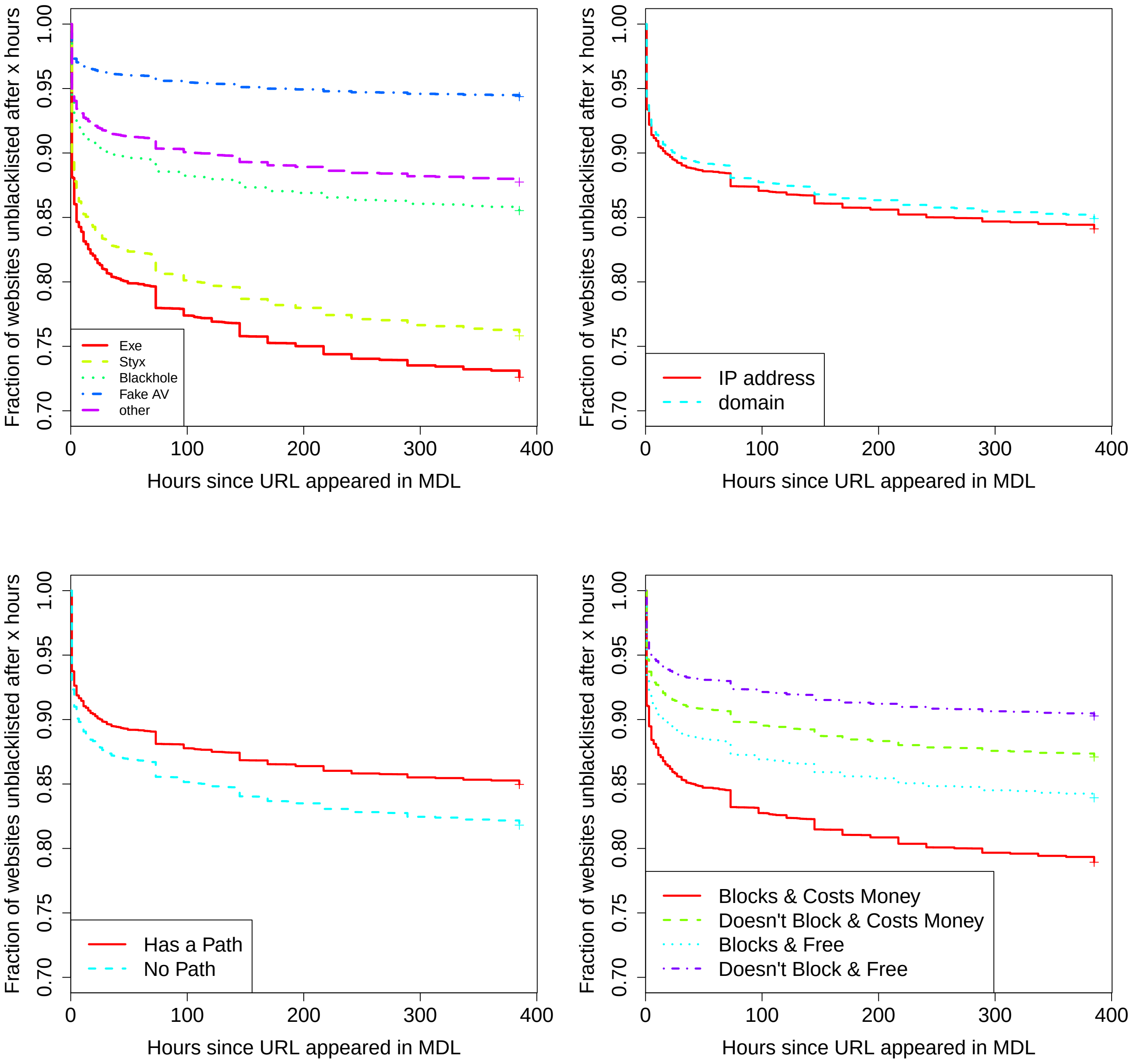

Fig. 5: Survival probability functions using the best-fit proportional-hazards model. Each graph shows the effect on survival probability when varying only one of the factors.

that more recent samples are less likely to be detected, though this relationship is not strictly monotonic. Other research builds upon this looking at the diversity in detection of malware binaries [12] [13]. This work shows similar trends as our work, namely the log-log detection [13] and the shape in the cumulative failure rate [12].

Provos et. al. described drive-by downloads and how to identify malicious URLs to create an efficient, malware blacklist [3]. Another type of previous literature that this research builds upon is the classification of malware type based on structure. John et. al. analyzes suspicious URLs based on URL properties to detect malicious search-redirection attack URLs [2]. Prakash et. al. manipulated features of current phishing URLs to better detect future phishing URLs [14].

VI. Concluding Remarks

We have presented an empirical analysis of malware URL blacklisting. Our goal has not been to judge which service is “better” than the others, but rather to seek explanations for the wide variation in detection rates and speed across

URLs and services. Our analysis has focused on submissions to the public Malware Domain List (MDL), which were repeatedly evaluated against the meta-scanner VirusTotal at regular intervals.

We have presented statistical models relying on two broad classes of explanatory factors that we believed to affect whether and when a suspicious URL gets blacklisted. The first category includes characteristics of the URL itself. We found that, for instance, executables and common exploit kits such as Blackhole and Styx are blacklisted very quickly, whereas malware designed to evade detection such as fake antivirus is in fact less likely to end up on a blacklist (and take longer to be put there). This suggests that the tactics criminals employ play a large role in affecting the success of our defenses.

But there can also be variations in detection rates not caused by the attacker's actions. Hence, the second class of explanatory variables we studied had to do with the blacklists themselves. Access to some blacklists cost money (e.g., Kaspersky URL Advisor), while others are given away for free (e.g., Google Safe Browsing). We found that on the whole, the free services did not fare as well as the paid ones in flagging malicious URLs. A more surprising finding, perhaps, is that blacklists that actively prevent subscribers from visiting suspicious URLs block more URLs faster than others.

Of course, there are many additional factors at play beyond those we have identified in this paper. This can be attributed to some of the paper's limitations, many of which we hope to address in future work.

A few limitations arise from the source of malware URLs we relied upon. The MDL is a low-volume list; our collection gathered about 120 malware URLs a month. Because of this, the URL features we could collect sufficient data on was limited. There are many more exploit kits than Styx or Blackhole, for example, but there were too few URLs in our collection to draw meaningful conclusions regarding their impact on detection. For example, we did observe several traffic distribution service (TDS) URLs in the MDL, but not enough to yield conclusive results.

Another promising area for further research if higher-volume data could be obtained is to compare blacklisting of legitimate websites that have been hacked to those registered by criminals (e.g., `http://hillaryklinton.ru:8080/forum/links/column.php`). We hypothesize that malicious websites could be blacklisted more quickly since they contain no legitimate content, which minimizes the negative consequences of a false positive.

The MDL is also a well known, publicly available list. Because of this, it is perhaps naïve to think that no services beyond the three we omitted use the MDL as input. Provided that the services do not take the MDL as input, then our study does in fact compare the coverage across services. If, on the other hand, they do take the MDL as input, then our study tests the effectiveness of their checker (i.e., sites not flagged as malicious have been deemed clean). We could get to the bottom of this only if we obtained a private stream of malware URLs to check.

A final limitation tied to the use of the MDL as input is that we cannot be certain whether a URL is in fact delivering malware at the time the service blacklisted them. Many malware URLs are shut down shortly after they are reported, and MDL reports all the URLs on its list. Consequently, we expect that at least some of the blacklisting is not necessarily beneficial to users.

A second set of limitations involve our method of malware verification. VirusTotal aggregates malware blacklists of only those services who agree to anonymously share data.[1] Thus, we are not able to evaluate the effectiveness of other blacklists whose operators choose not to share data with others. It would be interesting to study the effect of sharing on blacklist accuracy. We hypothesize that sharing would be positive effect, but we have no evidence to support this at present.

Another issue with using VirusTotal as we have in this paper is that the results for a given blacklist may differ from the experience of its customers. We threw out six blacklists that did not flag a single URL. Some of these services may in fact be dormant, but others simply give different answers to VirusTotal than they do to customers. For example, we found several examples of URLs submitted to MDL that the "urlQuery" service categorized as Blackhole landing pages but VirusTotal said that urlQuery missed. We note that VirusTotal has a disclaimer indicating that this may very well happen. Likewise, a service's confidence in the state of a questionable URL might display different things to VirusTotal than to users, which might skew our results.

Despite these limitations, we believe this work is a useful first step in an area ripe for further investigation. Identifying the factors that influence when a malware URL is blacklisted can help defenders improve the accuracy and speed of these services.

Acknowledgments

This work was partially funded by the Department of Homeland Security (DHS) Science and Technology Directorate, Cyber Security Division (DHS S&T/CSD) Broad Agency Announcement 11.02, the Government of Australia and SPAWAR Systems Center Pacific via contract number N66001-13-C-0131. This paper represents the position of the authors and not that of the aforementioned agencies.

References

[1] T. Moore, N. Leontiadis, and N. Christin, "Fashion crimes: trending-term exploitation on the web," in *ACM Conference on Computer and Communications Security*, Y. Chen, G. Danezis, and V. Shmatikov, Eds. ACM, 2011, pp. 455–466.

[2] J. P. John, F. Yu, Y. Xie, A. Krishnamurthy, and M. Abadi, "deSEO: Combating search-result poisoning." in *USENIX Security Symposium*, 2011.

[3] N. Provos, P. Mavrommatis, M. Rajab, and F. Monrose, "All your iFrames point to us," in *Proceedings of the 17th USENIX Security Symposium*, Aug. 2008.

[4] "Malware domain list," http://www.malwaredomainlist.com/.

[5] "VirusTotal," https://www.virustotal.com/.

[6] No Virus Thanks Blog, "Rogue security software XP Total Security spreads by email," http://blog.novirusthanks.org/2012/12/alert-unread-message-rogue-security-software-xp-total-security/.

[1] VirusTotal reports URLs/binaries to services if the URL/binary is flagged as malicious by at least one service; that flagging as malicious is the information that the services share.

[7] "Styx exploit kit," http://www.malwaresigs.com/2012/12/19/styx-exploit-kit/.

[8] M. Kührer and T. Holz, "An empirical analysis of malware blacklists," *Praxis der Informationsverarbeitung und Kommunikation*, vol. 35, no. 1, p. 11, 2012.

[9] D. R. Cox, "Regression models and life-tables," *Journal of the Royal Statistics Society, Series B*, vol. 34, pp. 187–220, 1972.

[10] S. Sheng, B. Wardman, G. Warner, L. Cranor, J. Hong, and C. Zhang, "An empirical analysis of phishing blacklists," in *Sixth Conference on Email and Anti-Spam*, 2009.

[11] J. Oberheide, E. Cooke, and F. Jahanian, "CloudAV: N-Version Antivirus in the Network Cloud," in *Proceedings of the 17th USENIX Security Symposium*, San Jose, CA, July 2008.

[12] M. Cukier, I. Gashi, B. Sobesto, and V. Stankovic, "Does malware detection improve with diverse antivirus products? An empirical study," in *32nd International Conference on Computer Safety, Reliability and Security*. IEEE, September 2013, to appear.

[13] I. Gashi, V. Stankovic, C. Leita, and O. Thonnard, "An experimental study of diversity with off-the-shelf antivirus engines," in *IEEE International Symposium on Network Computing and Applications*, 2009, pp. 4–11.

[14] P. Prakash, M. Kumar, R. R. Kompella, and M. Gupta, "Phishnet: predictive blacklisting to detect phishing attacks," in *INFOCOM*. IEEE, 2010, pp. 1–5.