Effectively Monitoring the Health of the Anti-phishing Ecosystem

Dan Geer and Adam Oest

Shepherd: Rik Farrow

Despite receiving extensive attention from the cybersecurity community, phishing remains a major, and agile, threat to Internet users. In early 2020, attackers quickly pivoted in an attempt to take advantage of the increased reliance on the Internet (especially by new and unsuspecting users) amid the worldwide outbreak of COVID-19. Google Safe Browsing (GSB) detected a record number of phishing websites in the first half of 2020, alongside more than 100 million daily phishing emails and hundreds of millions of other scam emails.[i] Any level of risk multiplied by numbers like 10^8 can only be described as substantial. The security community must rise to the occasion and do more to protect users targeted by large-scale phishing and disrupt the criminals engaged in it.[ii] The trendline in Figure 1 is perhaps the best summation of the situation.

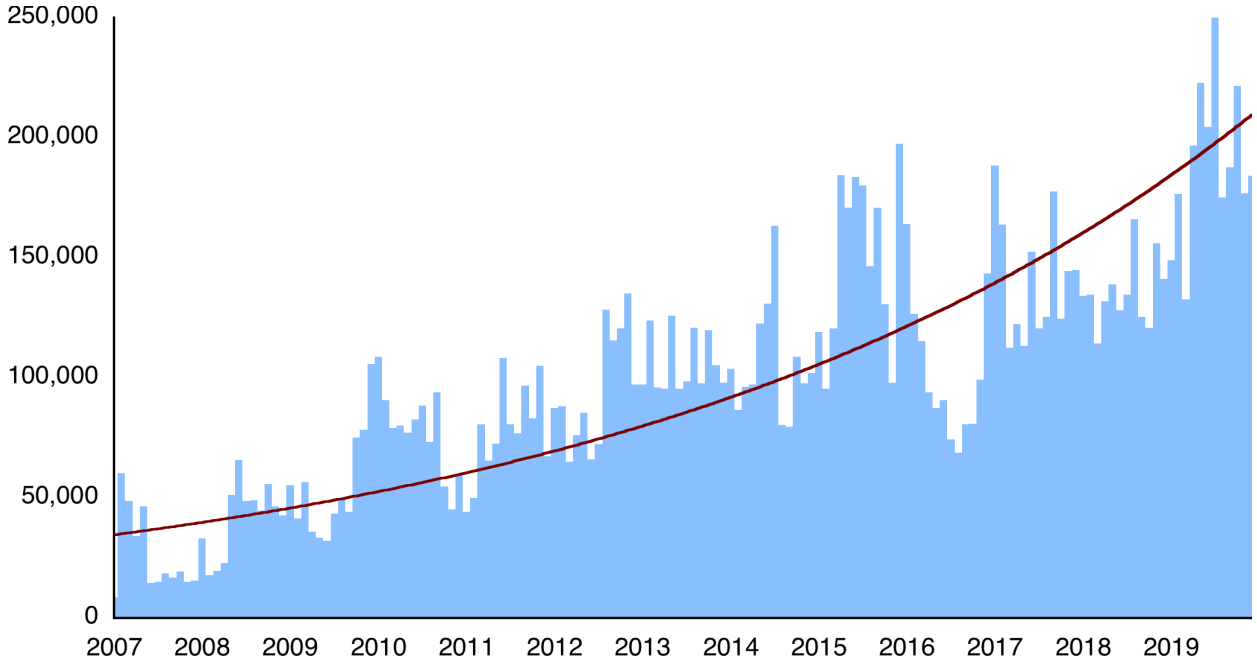

Figure 1: Phishing websites detected by Google Safe Browsing, by month

Most cyberattacks exploit technical vulnerabilities in computer systems; however, phishing attacks instead target the users of those systems with deceit, tricking the human user sitting in front of the screen into taking actions they would not normally take, such as sharing sensitive personal information. In turn, these actions put the victim users and others at risk. Even though phishing may appear trivial on the surface, it is not – modern attacks are marked by a high degree of sophistication, especially measures to evade timely detection by anti-phishing systems. That growing sophistication makes phishing particularly dangerous.

Phishing evolves through a cat-and-mouse game between attackers and defenders, between criminals and organizations they impersonate – well-known brands, email providers, web browser developers, security software vendors, and more. Impersonated organizations form an anti-phishing ecosystem to protect potential victims of phishing through technical controls and educational measures. Phishing attack mechanisms specifically seek to avoid detection by the ecosystem, as has been confirmed by research showing that attacks specifically crafted to be evasive are responsible for the majority of real-world damage caused by phishing.[iii],[iv]

Anti-phishing Defenses

To understand why phishing attacks benefit from avoiding or delaying detection, we must first discuss a key anti-phishing mitigation used across the ecosystem: browser-based warnings. Today's major web browsers, including Google Chrome, Mozilla Firefox, Apple Safari, and Microsoft Edge, display a prominent alert whenever a user tries to visit a known phishing website. A key advantage of these alerts is that users can be protected far more quickly compared to traditional anti-phishing measures such as website take-down.[v]

However, browser-based warnings rely on lists of known malicious website URLs which are only updated reactively: only after an initial incident or detection is registered by a corresponding backend system. This detection requires retrieving and scanning the content of a suspected URL to verify that it is indeed malicious. Thus, attackers can maximize the longevity of their websites (and the return-on-investment that motivates their attacks) by delaying the backend system's discovery of their phishing URL in the first place, and then making it difficult for that system to retrieve the content itself.

Techniques that delay browser-based warnings also help attackers slip past other anti-phishing mitigations, such as spam filters. Although spam filters will often proactively block phishing emails based on the message content alone, the phishing emails that fail to be detected initially can generally only be flagged retroactively if the link found in the email ends up on a list of malicious URLs. Attackers degrade this retroactive flagging by ensuring that links in emails differ from the phishing landing page, such as by leveraging link shortening and tracking services.

Challenges and Current Solutions

When defending against evasive phishing attacks, a key challenge the ecosystem faces in developing effective protections is overcoming the highly distributed nature of the phishing attacks themselves. For example, phishing emails that impersonate a given brand (a retailer like Walmart, say) will oftentimes be distributed to email inboxes provided by an entirely different organization and, naturally (if unfortunately), only a tiny subset of recipients will subsequently report them to the brand. In addition, the web browsers and security applications that implement anti-phishing mitigations are typically developed by yet another entity. The impersonated brand may, therefore, struggle to maintain an accurate and up-to-date understanding of the phishing attacks made against it while also being dependent on external vendors to protect its users.

To help victim organizations reduce gaps in attack visibility, clearinghouses track and aggregate phishing website URLs from across the ecosystem. These clearinghouses, such as PhishTank or the Anti-Phishing Working Group (APWG), allow different brands to view and share phishing detections through a centralized database. This data sharing can bolster the anti-phishing efforts of individual organizations by increasing the diversity and timeliness of URL detections available to them. Clearinghouses are also instrumental for research into phishing trends.

Despite the advantages of today's clearinghouses, the merge-aggregation of phishing URLs fails to paint a complete picture of the health of the anti-phishing ecosystem. Specifically, contextual metadata is missing about the detection and mitigation times of each URL. Without this information, it is difficult to pinpoint high-impact phishing URLs, and, consequently, it is difficult to understand the significance of increases or decreases in the raw count of phishing URLs being observed. In a sense, the ecosystem is thus

flying blind through the constant flak of phishing attacks.

Monitoring the Ecosystem's Health

In 2018, a group of researchers from Arizona State University, PayPal, and Samsung Research proposed a framework for measuring the latency of browser-based warnings using large samples of innocuous (but real-looking) phishing websites.[vi] The research revealed several critical detection gaps which were subsequently addressed by the browser vendors, such as inconsistencies between warnings shown on desktop and mobile browsers. More importantly, however, the study revealed that evasive phishing websites – those used by state-of-the-art criminals – could avoid being blocked for several hours (on average), thus giving attackers a prolonged window of opportunity to cause harm to their victims. This finding explains, in part, why attackers continue to be so persistent in launching phishing campaigns. It also shows a clear and measurable way in which the ecosystem can improve: reducing the detection latency and increasing the detection coverage of phishing websites.

Table 1 shows the blocking performance of Google Safe Browsing (used by Chrome, Firefox, and Safari) and Microsoft SmartScreen (used by Edge and Internet Explorer) in a controlled sample of 2,862 phishing websites reported to the ecosystem in 2020.[v] To illustrate, phishing websites with a landing page containing basic bot detection and a corresponding lure e-mail with a redirection link – a configuration now typical for attackers – would take an average of 2 hours and 43 minutes to be blocked by GSB, and in 15.7% (1 minus 84.3%) of cases, they would not be blocked at all. Moreover, all websites tested with newer, sophisticated evasion simply failed to be blocked. Such websites represent an emerging trend among phishing kit developers.

	Google Safe Browsing		Microsoft SmartScreen	
	latency	blocked	latency	blocked
Phishing website with no evasion	00:57	92.9%	02:48	93.2%
Phishing website with basic server-side evasion	00:59	94.2%	03:04	100%

Phishing website with basic server-side evasion and lure with redirection	02:43	84.3%	10:01	86.6%
Phishing website with advanced "behavior based" evasion (collectively 3 different types of JavaScript-based human verification)	N/A	0.0%	N/A	0.0%

Table 1: In-browser blocking performance

The ecosystem as a whole could benefit from the expansion of this measurement approach across a large dataset of phishing URLs, such as those aggregated by clearinghouses; in addition to collecting the phishing URLs themselves and timestamps of when they were reported, clearinghouses would also verify the presence of the URL on key blocklists from the time they are reported to the time they go offline. Figure 2 helps explain the importance of the latency measurements in the context of how attackers approach phishing campaigns.

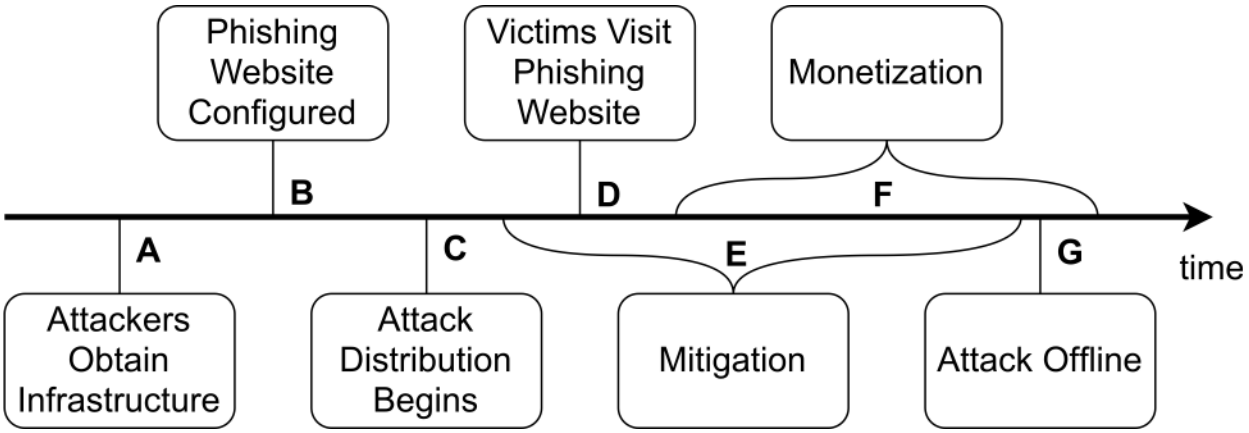

Figure 2: High-level stages of a typical phishing attack

Capturing these latencies would, in turn, allow for the following improvements to the security provided by the anti-phishing ecosystem:

Timely identification and mitigation of harmful threats

Discovering gaps in the blocking of URLs can improve the timeliness of identification of sophisticated attacks and targeted mitigation of the corresponding URLs. For example, if certain URLs are not being blocked due to evasion strategies implemented by

phishing websites, those strategies could be investigated, and the classification capabilities of detection systems improved. This is particularly important given the recent observation that a small proportion of (advanced) phishing websites is responsible for the majority of harm to victims.[iii]

Ensuring the propagation of phishing URL reports

In certain cases, phishing URLs may fail to be blocked by the ecosystem not because of a detection problem, but because those URLs were not reported to the appropriate anti-phishing entities in a timely or complete manner. Thus, with blocking data in hand, clearinghouses could identify shortcomings in reporting based on variables such as the targeted brand, the source of the original URL detection, or the submitted evidence showing the maliciousness of the URL.

Monitoring the health of anti-phishing efforts

The blocking latency of phishing URLs across the entire ecosystem is a key metric for tracking the collective health of anti-phishing efforts over time. In the long term, a decrease in blocking latency is desirable. Moreover, knowing how well phishing URLs are blocked can help put existing metrics – such as raw URL counts – into perspective. Increases in phishing URL counts without a corresponding decrease in detection latency could, for instance, suggest the need for additional mitigations.

Individual brands could similarly benefit from these metrics, as they would have a better understanding of how well, collectively, their customers are being protected. Such data would be particularly useful for brands commonly targeted by fraud, who might find it challenging to attribute losses directly to phishing amid similar threats such as data breaches or malware.

Phishing attacks will continue to evolve, and it is time for anti-phishing efforts to leverage the wealth of data available to the ecosystem. Using large-scale phishing data to monitor the effectiveness of URL blocking, and find the gaps being exploited by attackers can now be done. It only requires resolve.

Internet users and industry practitioners rely on controls that are core to the ecosystem.

Corporations and government entities rely on control testing to create risk assessments for their respective businesses. Methodical, continuous testing of a key ecosystem control by a non-biased entity is not only beneficial for phishing victims who are consumers, but also for the mitigation of phishing campaigns that target corporate accounts. As one effort, the Anti-Phishing Working Group (APWG) is building monitoring into its eCrime exchange (a phishing URL clearinghouse) to capture the latency between when a URL is reported to representative anti-phishing blocklists and when it gets blocked (if it actually is blocked).[vii]

The ecosystem must also look toward the future. Fundamentally, URL blocking is a reactive mitigation which will always leave some window of opportunity for attackers, even as the blocking latency declines. Proactive anti-phishing systems are comparatively less mature than reactive ones, but they will not share the same shortcomings.[viii] Thus, as the ecosystem works to eliminate gaps in reactive mitigations, it should also consider proactive approaches such as in-browser warnings that scan page content rather than relying solely on a list of known malicious websites.

[i] Google Safe Browsing Transparency Report, https://transparencyreport.google.com/safe-browsing/overview

[ii] FTC Received Nearly 1.7 Million Fraud Reports, and FTC Lawsuits Returned $232 Million to Consumers in 2019, https://www.ftc.gov/news-events/press-releases/2020/01/new-ftc-data-shows-ftc-received-nearly-17-million-fraud-reports

[iii] Adam Oest, Penghui Zhang, Brad Wardman, Eric Nunes, Jakub Burgis, Ali Zand, Kurt Thomas, Adam Doupé, Gail-Joon Ahn: “Sunrise to Sunset: Analyzing the End-to-end Life Cycle and Effectiveness of Phishing Attacks at Scale”, Proceedings of the 2020 USENIX Security Symposium

[iv] Camelia Simoiu, Ali Zand, Kurt Thomas, Elie Bursztein: “Who is targeted by email-based phishing and malware? Measuring factors that differentiate risk”, Proceedings of the 2020 ACM Internet Measurement Conference

[v] Adam Oest, Yeganeh Safaei, Penghui Zhang, Brad Wardman, Kevin Tyers, Yan Shoshitaishvili, Adam Doupé, Gail-Joon Ahn: “PhishTime: Continuous longitudinal measurement of the effectiveness of anti-phishing blacklists”, Proceedings of the 2020 USENIX Security Symposium

[vi] Adam Oest, Yeganeh Safaei, Adam Doupé, Gail-Joon Ahn, Brad Wardman, Kevin Tyers,:“PhishFarm: A Scalable Framework for Measuring the Effectiveness of Evasion Techniques against Browser Phishing Blacklists”, Proceedings of the 2019 IEEE Symposium on Security and Privacy

[vii] PhishFarm Block List Latency Monitoring, https://ecrimeresearch.org/phishfarm

[viii] Bin Liang, Su Miaoqiang, You Wei, Shi Wenchang, Gang Yang: "Cracking classifiers for evasion: a case study on the Google's phishing pages filter", Proceedings of the 2016 International Conference on World Wide Web