Examining Newly Registered Phishing Domains at Scale

Sharad Agarwal*[1] and Marie Vasek†[1]

[1]Department of Computer Science, University College London (UCL)

Abstract

Phishing has been prevalent for over two decades, evolving recently into new forms, such as smishing or SMS phishing. Despite its long reign, it continues to be significant, deceiving victims globally. Cybercriminals compromise benign websites or register new domains to host phishing web pages. These pages impersonate brands and lure victims into providing their personal or financial details. The infrastructure criminals exploit differs for compromised benign websites and newly registered domains, requiring different mitigation approaches. In this paper, we investigate new domains that cybercriminals register with the sole intent to host phishing websites. We analyze 15 126 unique newly registered domains over a period of 11 months. These domains have an average lifetime of 8.6 days, and miscreants re-register old domains to use them for phishing. The third-party infrastructure these domains exploit differs greatly from infrastructure where hacked websites or the general population of websites concentrate. This allows us insight into where efforts can be taken to take down these maliciously registered domains quicker, from a relatively new ASN to the `.COM` registry to more anonymous registrars. From our findings, we derive a set of recommendations for stakeholders towards reducing the effects of this scourge.

1 Introduction

Despite the extensive amount of work done by academics and industry alike, phishing remains one of the most prominent cybercrime activities as of 2023 [39]. Defender efforts have shifted over the years, matching the shifting behavior of attackers. For instance, the Rock Phish gang in the mid-2000s led defenders to work to secure wild card DNS. Similarly, Google rolled out their safe browsing initiative in 2007, which bundled their anti-phishing blocklist with browsers, simplifying the previously complicated consumer security landscape and allowing browser and email intelligence to inform each other. While we can leverage existing knowledge, this repeated game of whack-a-mole makes it tricky to rely exclusively on the large body of work covering this ever-changing landscape.

One critical longstanding issue for miscreants in this space is acquiring new phishing domains to use. Cybercriminals can compromise benign websites either by uploading additional content to an existing website or altering the existing domain's DNS to add a path to the criminal website. This then allows them to host their phishing page, bootstrapping the reputation of the phishing page to the reputation of the hacked website. Furthermore, these domains can be tricky to remediate because the best practice is to save the good content (and fix the hole that the attacker came in from) and remove the phishing content. That can be expensive and time-consuming.

Otherwise, miscreants need to register new domain names (or re-register existing ones) whose sole purpose is to purport cybercrime. These domains can impersonate the relevant brand, be general-purpose for multiple types of brand impersonation, or be randomly generated. Each of these has its own purpose – the first two can trick users who read domain names, but do not recognize the deception. Previous research indicates that users often fail to recognize domains with unusual top-level domains (TLDs), additional terms, or those using subdomains [107]. Criminals can take advantage of this and set up domains using various squatting methods to deceive users [61, 64, 89, 144] or simply use an incorrect TLD [91]. Well-crafted adversarial web pages can trick many, including IT experts [35]. Randomly generated domains are useful for circumventing restrictions on domain registrations that some registrars use to avoid accidentally registering domains for malicious users.

Our paper concentrates on malicious domain registration towards carrying out phishing attacks. These domains only carry out malicious ends, so they are easier to take down by contacting the affected third parties (registrar, hosting provider, even domain registry) or adding the domain to a blocklist indefinitely. However, increasing competition among DNS providers and the availability of top-level domains (TLDs) make it easier for criminals to register domains. Furthermore, protective DNS services are readily available through proxy

*sharad.agarwal@ucl.ac.uk
†m.vasek@ucl.ac.uk

services like Cloudflare, and websites can be easily set up via free online services like Google Firebase. This complicates efforts to block domains based on IP address and hinders takedown companies from more effectively removing these sites.

While previous work concentrates on compromised, re-registered domains or phishing email/webpage detection, our work focuses on a comprehensive investigation of newly registered phishing domains. Other recent work that covers this area focuses on registration of phishing domains on particular TLDs [46, 140], re-registration of malicious domains [66, 71, 81] or differentiating between phishing domains which were hacked vs. maliciously registered [78, 120]. Infrastructure abused for malicious registrations differs greatly from infrastructure abused to host phishing web pages on compromised websites; this comparison leads us to offer up actionable information for third parties and governments trying to clean up third parties. This narrow focus also allows us to offer insights into the current state of abused infrastructure used to support maliciously registered phishing pages.

This research investigates newly registered domains to answer the following research questions:

RQ1 *What infrastructure do criminals abuse to host newly registered phishing domains?*

RQ2 *What brands do cybercriminals impersonate?*

RQ3 *How long are these domains active?*

RQ4 *How effective are antivirus vendors in detecting newly registered phishing domains?*

This paper provides the following contributions:

- We investigate 15 126 newly registered phishing domains over 11 months. We demonstrate how the distribution of various abused third parties (certificate authorities (CAs), registrars, top-level domains (TLDs), and autonomous systems (ASes)) differs compared to work investigating hacked web pages or the population of the Internet writ large.
- Particularly, we discover in Section 5.4 that more than 79.3% of newly registered phishing domains are issued TLS certificates compared to 66.7% of all phishing websites [60]. Furthermore, 190 phishing domains were issued over 50 TLS certificates since their first registration.
- In Section 5.3, we identify a relatively new (less than 2 years old) AS targeted to host a large number of newly registered phishing domains, despite only hosting 17 754 total domains.
- A considerable amount (26%) of newly registered domains abuse Cloudflare to evade detection and takedown compared to one-third of all phishing websites [54], 39% of smishing websites [92] and 19% of the Internet [27].
- There are discrepancies in Google Safe Browsing detection results for the identified newly registered domains abused for phishing over different platforms, largely in how the API vs. web platform handles phishing web pages only on subdomains.
- Newly registered phishing domains are active for an average of 8.6 days compared to 21 hours for an average phishing campaign [99]. We additionally find that 25.7% of our sample were domains that were previously registered; these overwhelmingly were originally registered for cybercriminal purposes.

2 Background

In this section, we provide the required background for our paper. We start with a quick overview of the major stakeholders involved in the domain name ecosystem. We then explain how criminals conduct phishing and smishing by compromising already-hosted websites and newly registered domains. Lastly, we discuss the attacker's preferences and what it means for the abused third parties.

2.1 Domain Name Ecosystem

Criminals compromise already-hosted websites or register new domains to conduct phishing and smishing. In both cases, third-party services are abused to host phishing web pages. However, depending on the scenario, they may or may not interact with the criminals and take required actions accordingly. Therefore, it is necessary to understand all the major stakeholders involved and their roles in a domain name registration ecosystem.

ICANN. The Internet Corporation for Assigned Names and Numbers (ICANN) is a non-profit organization that coordinates the global Internet's system of unique identifiers, i.e., numerical addresses and domain names. ICANN is mainly responsible for generic top-level domains (gTLDs).

Registry. Registries operate all individual top-level domains (TLDs) and maintain zone files that list all domain names and their authoritative name servers using Rapid Zone Update. IANA classifies TLDs into six groups: the most common are generic (gTLD) and country-code (ccTLD) [56]. Registries need to be approved and follow ICANN policies to operate a gTLD, whereas the influence that ICANN has on ccTLDs like .uk or .ca is significantly lower. Certain registries, like the national ones, proactively detect abuse on their TLDs and work with the relevant registrars for domain takedowns [18].

Registrar. A registrar sells domain names to a registrant. They need to be accredited with ICANN or a ccTLD registry to sell a domain name for any specific TLD. For example, IONOS is registered with Nominet, the registry managing `.uk` ccTLD and can sell domain names with a `.uk` TLD to users. A registrar's primary responsibility is maintaining registration records, including the registrant's contact information and the authoritative name servers associated with the domain. ICANN provides a list of over 2 000 accredited registrars that manage generic top-level domain registration services.[1]

The 2013 Registrar Accreditation Agreement requires ICANN-Accredited registrars to provide abuse contact information and take steps to investigate reports of abuse. This provides the registrars with the power to take reasonable and prompt steps to investigate and respond appropriately to any reports of abuse, including de-registering domains in line with their investigations [52].

Autonomous System (AS). An autonomous system (AS) is a large group of networks with a unified routing policy that serves a set of IP prefixes [26]. Many ASes are also *hosting providers* who provide servers which host and serve content. ASes are supposed to perform due diligence before providing hosting services and proactively deal with abuse on their networks. They have the power to remove harmful content hosted on their networks to prevent users from accessing it. Similar to the registrars, they should have contacts where defenders can report phishing and promptly investigate those reports to take necessary actions.

Certificate Authority (CA). Certificate authorities store, sign, and issue digital certificates to websites validating their domains. These digital certificates are used by TLS (transport layer security) protocols which allow systems to verify identify and provide encrypted communication. In the case of maliciously registered domains, CAs can detect abuse when criminals request that they issue a TLS certificate for the newly registered domain to host a phishing web page. They can utilize the historical data of abuse for that domain before issuing a certificate and even have the power to revoke certificates for confirmed abuse.

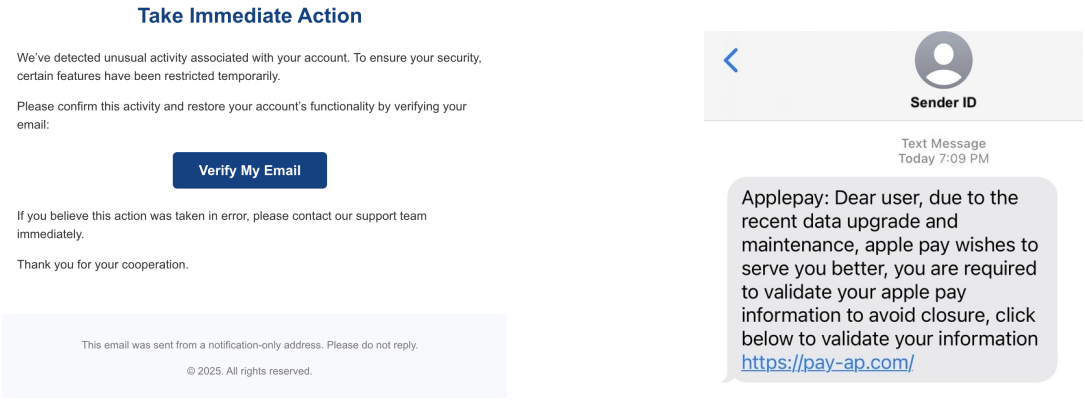
2.2 Compromised Websites

Criminals set up phishing web pages by identifying vulnerabilities on existing websites and exploiting these vulnerabilities to gain access and host web pages that impersonate known brands [82, 86, 136]. These phishing websites are typically uploaded to a *compromised website* into a separate directory from the rest of the webpage in order to evade detection and not accidentally interfere with the website operation. Criminals then link to the uploaded phishing webpage using the entire long URL. For example, `hxxps://talofdrag[.]tempurl[.]host/wp-admin/net/auth/autorisatioss/f9a015b70637e871832a0b3f585d9274/` hosts a phishing web page on a compromised website, impersonating a popular entertainment brand - *Netflix*.[2]

Medium of Abuse. Email allows criminals to hide the complete path of the phishing web page hosted on a compromised website using text or button hyperlinks. Fig. 1a shows how criminals hide the full path of the phishing webpage using a hyperlink button in the email.

Infrastructure. Criminals hack innocent websites to host phishing web pages. Phishing pages take advantage of the domain's reputation and existing infrastructure to perpetuate phishing attacks. Note that they do not have to communicate directly with any registrar, AS, or CA to host their content. While threat actors do not have to pay any monetary infrastructure cost when compromising a legitimate website, legitimate domain owners sometimes bear higher costs after being abused for this content like clean up costs and reputational costs. However, cybercriminals do require the technical skills to identify and exploit vulnerabilities to host a phishing web page; sometimes miscreants here hire others to exploit webpages on their behalf making this method not cost-free for the attacker. We can see cybercriminals interacting with a legitimate website to host a phishing web page in Fig. 2. In most cases, the service providers are also unaware that a phishing web page is abusing their services on a legitimate domain unless reported or the website is proactively monitored.

[1] ICANN accredited registrars - `https://www.icann.org/en/accredited-registrars`
[2] Phishing website reported to PhishTank - `https://phishtank.com/phish_detail.php?phish_id=8959437`



(a) Example of a phishing email that hides the complete path of a phishing webpage using a hyperlink button.

(b) Example of a smishing text message that uses newly registered domain.

Figure 1: Example of phishing email embedding the malicious URL and a smishing text message displaying the malicious URL in plain sight.

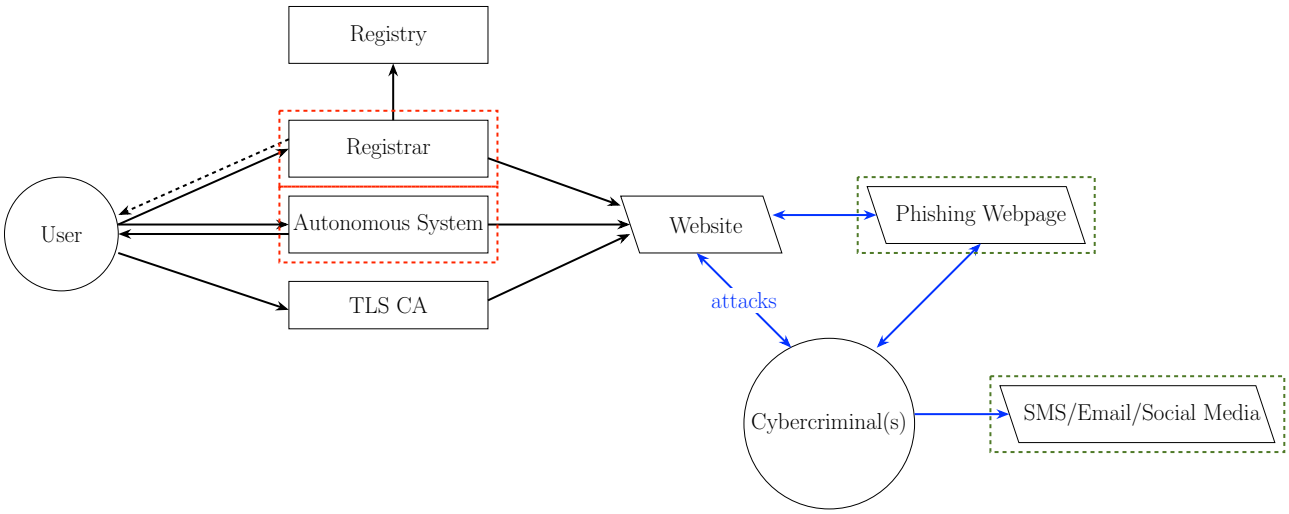


Figure 2: Cybercriminals host a phishing webpage by compromising a legitimate website. Black lines: user activities and stakeholder interactions. Blue lines: cybercriminal activities. Green boxes: intelligence signals to detect phishing. Red boxes: stakeholders who get reports from threat intelligence organizations.

Identification. Legitimate websites are usually not monitored for abuse unless the owner opts into a paid service. Defenders can identify a phishing web page hosted on a compromised website once a user flags it (Fig. 2). After confirming a phishing web page, the URL gets added to a blocklist. However, as these websites are known to serve a legitimate purpose, the URL blocklists are dynamic, so a website can be removed once the harmful content hosted on it is removed. Adding the domain name or the hosting IP address to a permanent blocklist would significantly impact the website's traffic, business, and hosting provider's reputation. While defenders would send a take-down request, registrars cannot take down a legitimate domain. The hosting provider can also not stop hosting the website without trying to help fix it, as it would be overly punitive. In such cases, the registrar and/or hosting provider contacts the site administrator, asking to fix the vulnerability that lets the criminal compromise the website and remove the harmful content from their website.

2.3 Maliciously Registered Domains

While some criminals prefer to compromise websites, others tend to register new domains to host phishing. Depending on threat actors, they have different strategies based on target victims. They often use techniques like typosquatting, i.e., purposefully registering a domain name that is a typo of a known brand's domain name, which can then be used to host a phishing page [2, 118, 119]. These domain names can easily be misread or have the correct name with a different top-level domain (TLD). Not all maliciously registered domains fall into this pattern. Others contain words associated with types of brands/sectors like delivery or banking companies. For example, criminals register domains such as `local-courier-id[.]com` and `support-myevri-locator[.]`

com to impersonate parcel delivery companies and send them to lure victims into providing their personal information.

Medium of Abuse. Unlike email, SMS does not allow obfuscating URLs in text. Sending the complete path for a phishing webpage hosted on a compromised website plainly would raise users' suspicion and not deceive them into clicking the link. Threat actors tend to register new domains to steal users' personal details over SMS (SMS phishing) [92, 118]. Fig. 1b shows how criminals use a newly registered domain impersonating *ApplePay* for smishing.

Infrastructure. Unlike compromised websites, criminals register domains, purchase hosting services, and arrange TLS certificates with the sole intent of perpetuating cybercrime: hosting a phishing website or spreading malware. We can see the interactions between cybercriminals and other stakeholders in Fig. 3.

Malicious registration provides some advantages to the criminals. First, they can register a domain based on the target campaign. While some domains are tied to brands such as `support-myevri-locator[.]com`, others are related to the smishing campaign and are abused for multiple brands. For example, `local-courier-id[.]com` can be abused for any delivery/parcel-related smishing campaign. Second, cybercriminals can select an AS that does not proactively remove abuse even if contacted by a defender or knowingly supports abuse and protects customers from abuse reports (aka Bullet-Proof Hosting providers which are often based in countries where they can evade law-enforcement [5, 76]). Legitimate websites generally use TLS certificates for encrypted communication between the server and the website. To gain users' trust, criminals also use TLS certificates of their choice with newly registered domains.

Free third party services can be abused to host phishing on new domains. Companies like Google provide free web hosting services along with a `web.app` domain.[3] For example, `login-auth-device[.]web[.]app`. In such cases, a cybercriminal does not need to buy a domain name from a registrar, hosting services from an AS, or even a TLS certificate. Criminals abuse these services to set up and host phishing pages without cost.

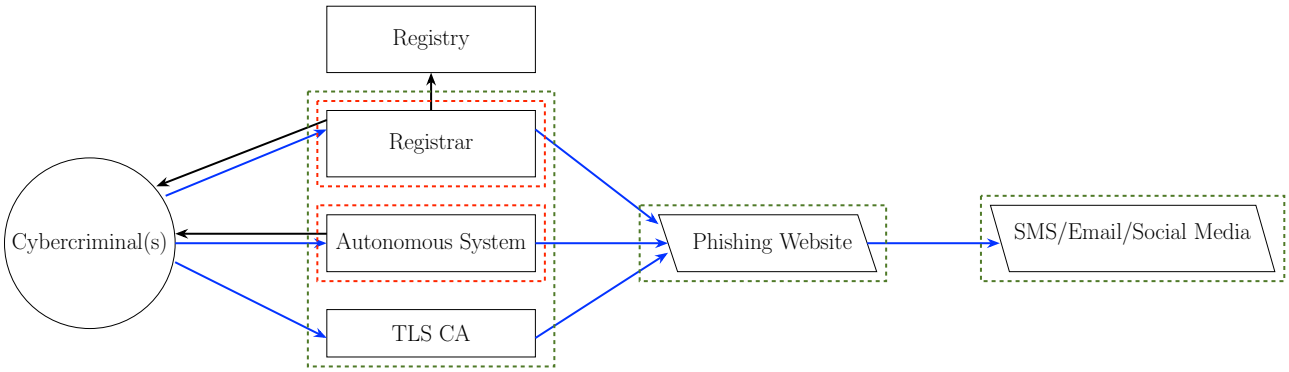


Figure 3: Cybercriminals register a new domain solely intending to host phishing. Black lines: user activities and stakeholder interactions. Blue lines: cybercriminal activities. Green boxes: intelligence signals to detect phishing. Red boxes: stakeholders who get reports from threat intelligence organizations.

Identification. Defenders can target stakeholders to takedown newly registered domains abused for smishing (Fig. 3). Unlike compromised websites, the surface exposure with newly registered domains is higher. Criminals register domains impersonating brands, provide details to the registrars, and are exposed from day zero. Defenders use zone files, IP addresses, and freshly acquired TLS certificates to identify newly registered malicious domains. Additionally, some registries monitor the domains on their TLDs to check for abuse and work with the respective registrars for takedowns. Many registrars provide a service to automatically register a wide range of possible typo-squatted domains when a trademark owner wants to register a domain [2]. More recently, the top registrars have started to offer a paid service to block certain kinds of typo-squatted domains for brand protection [45].

Takedown strategies here differ from removing content from hacked webpages. The maliciously registered domains can be added to a domain blocklist, preventing access from users. Additionally, mobile network operators use these blocklists to block text messages containing the domain. Complying registrars can suspend the domain as per ICANN policy and apply a transfer lock to prevent the registrant from attempting to evade the mitigation action and resume using the domain name [53]. Vigilant hosting providers often remove malicious content within hours of its discovery (either by report or self-monitoring) [83]. In the case of bulletproof hosting,

[3]Google provides firebase hosting service for free - `https://firebase.google.com/pricing`

the IP addresses or a range of addresses can be added to blocklists [5]. The hosting providers can also be reported or blocked for abuse by Internet service providers (ISPs) further limiting their reach [126].

2.4 Attacker's Preferences

While certain threat actors have preferred to compromise websites for phishing over emails for almost two decades [82, 99], there has been a rise in registration of malicious domains [96] to conduct phishing over SMS, smishing (SMS phishing) [119]. Here, we discuss the various reasons that affects the methods attackers prefer over the others.

Cost. Compromising websites saves criminals the cost of setting up the infrastructure. As the websites criminals compromise are already hosted, they do not have to bear the infrastructure cost to host phishing web pages. Additionally, it saves criminals the interaction with the stakeholders and the data (such as stolen credit cards) they would have to share with them to procure services such as buying a domain name from a registrar or hosting services. However, often cybercriminals mete out different responsibilities including compromising websites. This process comes at a cost.

Setting up infrastructure on new domains costs money. However, criminals can choose free third-party services such as the ones provided by Google Firebase. This saves them the infrastructure cost and also benefit from the AS's reputation. Note, typo-squatted domains with an unpopular TLD is not expensive (cf. Fig. 4).

Website Reputation. Hosting a phishing web page by compromising a legitimate website helps criminals use the trusted site's reputation, which helps bypass email security filters and makes the phishing email seem more credible to the recipient [28]. New domains start with low/no reputation and reregistered domains often carry forth old reputation. This can hamper the attractiveness of maliciously registered domains.

Deceiving Users. With compromised websites, criminals obfuscate the complete path in an email using hyperlinked text or a button, making it difficult for the user to identify the phish. On the other hand, newly registered domains allow criminals to select a more convincing domain name to lure victims. Based on the threat actor groups and their target victims, they (broadly speaking) prefer different strategies to host phishing web pages. For example, criminals that conduct smishing tend to use smaller [92] and more convincing domain names such as `pay-ap[.]com` to impersonate *Apple Pay*.

Hosting Persistence. Criminals have preferred to compromise websites [86]. While certain defender strategies focus on domain registration using zone files, TLS certificates issued in real time to identify malicious domains, others monitor websites using known exploits, hosted on rogue ASes making them susceptible to compromise [121,136]. While regular monitoring of susceptible websites could help in getting them fixed before being compromised [145], most ASes are unable to detect malicious activities on websites [21]. This helps criminals as not all small brands pay for regular monitoring and not all ASes proactively detect compromised websites. However, a compromised website hosting a phishing page can be confirmed once an anti-phishing defender manually evaluates a phishing report. Once a legitimately registered domain is compromised, the registrars or registries would not immediately proceed with a takedown. Instead, the stakeholders would contact the hosting provider or the site administrator directly asking them to remove harmful content. Depending on the service provider, these reports could take time, providing the criminals additional time to deceive victims.

In case of newly registered domains, criminals can select registrars and ASes that are less proactive in takedowns or support them in conducting abuse. Some threat actors even prefer using bullet proof hosting services which support abuse and are based in countries where it's difficult for law enforcement to proactively take action. The automatable, public contact information for the webpage given by criminals for newly registered domains is mostly incorrect [24]. This helps them gain some time. A good hosting provider might try to contact them to investigate the defenders' complaints before proceeding for stricter actions such as takedowns.

Availability. Criminals search on the Internet to identify vulnerable websites and recompromise the same websites to host phishing pages [86, 90]. Based on the medium of abuse, such as in case of smishing, they purchase new domains or re-register the previously taken down maliciously registered domains [49, 68]. Others monitor and register legitimate domains as soon as they expire or close the entity [87].

Evade Detection. Smartphones lack the sophisticated security filters that protect email accounts, making it easier for fraudsters to reach victims [101]. Emails also contain additional metadata which SMSes lack; this metadata is often leveraged to identify phishing. To avoid detection, criminals often use evasive techniques, including geofencing [97,108]. Criminals target victims based on countries — this is straightforward to ascertain

for phone-based fraud. They set up websites so only the targeted country's users can access the URL. This reduces the likelihood of detection by defenders' automated systems based in other regions or countries. Similarly, criminals can restrict access based on device type. For example, in case of smishing, domains are often set up to deliver a phishing page on a smartphone but not on a desktop machine. Alternatively, criminals sometimes redirect to download illicit Android malware if opened on a Android phone but show the phishing page on a desktop machine. While can techniques can be used by compromised websites, they are more commonly found in maliciously registered domains.

2.5 Need to Investigate Maliciously Registered Domains

There has been a recent rise in SMS phishing, aka smishing [41, 42, 100, 101]. Phishing pages hosted by compromised websites have a long, non-deceptive path that cannot be disguised in SMS. Therefore, to lure victims into taking action, scammers need to register look-alike malicious domain names to impersonate brands [118]. As more entities, including government and private sectors, move towards sending SMS messages to interact with users and provide updates, criminals have also increased the abuse of this medium to impersonate brands using maliciously registered domains. For example, recently criminals have been taking advantage of the fact that services send messages to users regarding taxes for using a toll road. Reports show that criminals abused multiple domains in various countries including US, Australia and Ireland to conduct toll tax smishing campaigns [37, 72, 119]. This indicates an increase in maliciously registered domains, making it a crucial research problem.

Culpability for abused third-party services. While criminals abuse third parties to maliciously register domains and host phishing pages, continuing to provide services after being aware of abuse indicates support for the criminals. A service that invests in maliciously registered domain detection, expedites responses to abuse notifications, and collaborates with threat intelligence organizations can be viewed as fulfilling its obligations – even if some malicious domains slip through. Conversely, providers that fail to implement basic safeguards or disregard abuse reports could be seen as facilitating the abuse. The infrastructure criminals exploit when compromising a legitimate website differs from that in case of maliciously registered domains and so does the takedown strategies. While the ASes hosting the legitimate website that gets compromised indirectly becomes a victim, the ones hosting maliciously registered domains could be colluding with criminals to support their illicit activities.

Summary. As SMS scams increase, attackers' reliance on maliciously registered domains indicates that criminals invest more in the infrastructure to set up smishing campaigns. This points towards the evolving landscape and how attackers refine their methods based on the characteristics of the medium they abuse. Attackers choose methods best suited to exploit the weakness and user behavior associated with the medium they target to abuse. It also suggests that there will likely be an increase in specialized infrastructure and tactics targeted to exploit smartphone users, highlighting the need for better defense strategies.

While most prior work investigating the infrastructure criminals abuse utilizes aggregators that contains unlabeled mixed websites [13, 60, 82, 99], it is necessary to understand what that means for tackling abuse.For example, the TLD or registrar distribution of hacked websites mostly shows how domains are targeted, rather than reflecting any security practices of the underlying infrastructure. Whereas the registrar distribution of maliciously registered domains delineates infrastructure providers who are either being abused by cybercriminals or harboring cybercriminals. Their attractiveness can reflect low prices, insufficient screening processes, and/or ease of use by cybercriminals in other ways. Such insights into the infrastructure could help industry bodies to produce best practice guidelines to efficiently tackle maliciously registered domains. It also motivates aggregators/defenders to properly label the identified phishing URLs for expedited actions. It will also aid government to create the required policies to stop criminals from registering new domains.

In the next section, we discuss the related work (Section 3) that will lay out the novelty of our work, which focuses on maliciously registered domains. Section 4 describes our methodology for collecting and enriching newly registered domains criminals abuse for phishing. Analyzing the enriched data, we present the insights into the infrastructure that criminals exploit to register new domain and host phishing pages in Section 5. Lastly, in Section 6, we discuss our findings and suggest mitigations.

3 Related Work

Here, we consider the large body of phishing research divided by the data source of the phishing URLs. While some only consider compromised websites (§3.1) and others only consider maliciously registered domains (§3.2); the vast majority of work considers mixed feeds of the two (§3.3), which weakens intervention measures suggested by the authors.

3.1 Compromised Websites

Criminals exploit vulnerabilities in legitimate websites to host phishing pages. For over a decade, this has been the most common method for criminals to conduct phishing. Criminals compromise and re-compromise machines to host phishing sites [86]. Web servers running content management systems like WordPress are more likely to be exploited to host phishing sites [98, 136]. This homogeneity creates a ripe target environment for attackers. Cybercriminals benefit from a website's good reputation when conducting phishing attacks on compromised benign websites. Compared to setting up a new domain, compromised websites can save time and money.

However, it is essential to understand the difference since the reporting targets, as well as the actions required to take down the malicious content, vary significantly based on the intent of the website. Furthermore, if more websites are being hacked, it is vital to provide cleanup services and more secure hosting options for webmasters. Whereas if more domains are being maliciously registered, it is vital to provide more resources to infrastructure operators to stop the malicious content from being registered and hosted in the first place. Predictive analysis [121] could help reduce compromised sites.

For compromised phishing websites, phishing URL structure can vary. Phishing URLs hosted by compromised websites can look like IP addresses as hostnames, random domain names with a deceptive path, long deceptive subdomains, and random unintelligible domains [79, 98]. Oest et al. hypothesize that categories of URLs representing compromised infrastructure and ones deployed by phishers will not overlap much, but do not verify this [98]. Others have proposed detecting compromised phishing web pages through patterns in substrings of the URLs [145], visual differences between other pages on the same website [28] or using network logs [21].

3.2 Maliciously Registered Domains

Maliciously registered domains often have deceiving domain names to lure victims more easily than a compromised website or using short URLs [22, 79]. Typosquatting is a common way to deceive victims into clicking on phishing URLs where scammers register domain names that are typo variations of popular domains [2, 89, 144]. Unlike emails, phishing URLs cannot be obfuscated in SMS. To lure victims, criminals have to use deceptive domains impersonating the targeted brand [92, 98, 131].

Domain tasting is a five-day grace period to temporarily register a domain, after which the registration fee cannot be refunded [51]. In 2008, McGrath and Gupta found that the adversaries also abuse this feature, as 6.1% of domains were dropped during the grace registration period [79]. A few years later, Coull et al. identified 76% of distinct registrations from the complete VeriSign dataset as the result of domain tasting [29]. While miscreants have previously abused domain tasting, our research finds the average lifetime of newly registered domains is currently more than the five-day grace period.

Specific TLD Abuse. Certain threat actors abuse specific TLDs to register new domains and sometimes abuse newly announced TLDs more than the old ones. Halvorson et al. found that within the first month of registration, new gTLD domains appear on a domain blacklist more than twice as likely as legacy TLDs [46]. Focusing on a specific ccTLD, Vissers et al. observed 80% of malicious `.eu` TLD registrations are part of 20 long-running malicious campaigns, and 18% of them never appeared on any blacklist [140].

Other Infrastructure Abuse. Depending on the threat actors, cybercriminals may purchase new and inexpensive web servers, rent from other actors, or host on available botnet infrastructure [80]. Bulletproof and anonymous hosting providers allow miscreants to host phishing websites and proactively help them by giving a few days of notice to move to another service when an abuse report is received [76]. These services are often registered in countries where law enforcement cannot easily reach them, making it difficult to take down these phishing websites. Since 2011, researchers have found different sets of TLDs [13], registrars [12, 13, 49], and hosting providers [13, 47, 49] abused to register and host malicious domains. However, the ecosystem has changed since then. Furthermore, each project covers different sets of infrastructure operators towards other aims, rather than comprehensively analyzing as many infrastructure operators as practicable towards proposing malicious domain mitigation strategies.

Discovering Malicious Domains. Previous work has leveraged machine learning towards a hope of discovering malicious domains early in their lifetime, at registration time or even beforehand [11, 40, 47, 48, 105, 116, 130, 149, 150]. Our work could help improve these algorithms with insights into their changing concentrations on different infrastructure operators.

3.2.1 Re-registered domains

Previous studies have uncovered that cybercriminals often reuse/re-register domains previously registered by themselves or others [3,49,68,71]. Moore and Clayton uncovered the reregistration of bank domains after banks close; cybercriminals take advantage of the dead bank's reputation for illicit activities [87]. Miramirkhani et al. study how domains are selected to be re-registered [81]. Lauinger et al. study 'drop-catch' services that compete to re-register domains [67]. In a follow-up study, they found that most re-registrations happened at the earliest possible time and observed re-registration of deleted domains in large batches hours later [66]. Malicious actors often delete their domains before expiration, enabling their re-registration, supporting continuous abuse [12].

Attacks supporting domain re-registration. Popular websites can contain remote JavaScript inclusions that point to expired domains. These then can be re-registered to perform code injection attacks, as uncovered by Nikiforakis et al. [94]. Reregistration of expired domain names could hijack thousands of domain names via outdated name servers, as Vissers et al. demonstrate [139]. Similarly, Schlamp et al. demonstrated how attackers could hijack an entire autonomous system by re-registering the expired domains [112,113].

While previous research on malicious domain registration primarily focuses on the re-registration of domains, we investigate newly registered domains identified by our collaborating organization that are abused to host phishing websites.

3.3 Mixed websites

Overwhelmingly, most previous phishing research leverages phishing data feeds from industry, aggregators, and other online forums such as Twitter [15, 36, 79, 82, 98, 109–111, 115, 136]. These aggregator data feeds primarily consist of compromised websites and do not separate compromised websites from maliciously registered phishing domains. This makes it difficult to differentiate the infrastructure abused between compromised websites and maliciously registered phishing domains, making takedowns and insights necessary to facilitate takedowns harder. Data collected by aggregator websites are open to many; we note that the overlap between the aggregators and take-down companies has historically been significant [84].

TLS certificates abused for phishing. Bijmans et al. found scammers use multiple TLS certificates for phishing websites [13]. CAs (Certificate Authorities) provide TLS certificates for secure website communication. The CAs do not revoke most TLS certificates issued to phishing websites [13]. Borgolte et al. propose a new authentication method for trust-based domain validation by incorporating the existing trust of a name into the validation process [17]. Kim et al. study the abuse of TLS certificates in aggregator data that mainly consists of hacked websites [60]. Our research is the first to provide insights into TLS certificates abused by new maliciously registered phishing domains.

Phishing removal. Phishing removal has mainly focused on taking down IP addresses or domain names hosting malicious content. Previous research found that non-Rock-Phish sites are taken down sooner, and IP address removal is slower than domain removal [82]. Scammers use this to their advantage and deploy multiple IP addresses and domains to keep their phishing sites running. Strategies used by Rock-Phish and other fast-flux attacks include deploying multiple IPs and replacing them automatically to keep the website up and running [79, 82], lengthening the lifetime of the phishing site. Fast-flux attacks pose the greatest phishing threat and continue to send spam for many longer-lived websites until they are finally removed [88]. While some service providers are faster than others at removing phishing sites, particular brands can get fraudulent sites removed more quickly [83]. Compared to law enforcement agencies, specialized brand protection firms are more effective at taking down phishing websites [50].

Phishing detection and categorization. Previous studies provide different methods to detect phishing or otherwise malicious domains [14, 59, 77, 116]. Others developed classification models to differentiate between compromised and newly registered domains [78,120]. We empirically analyze newly registered domains at scale, providing insights that could be used to improve these detection methods.

Our work investigates over $15k$ newly registered domains in real time, provided by our collaborators. We show how infrastructure criminals abuse new domain registrations to conduct phishing and how this has changed compared to previous studies. We support the need to differentiate compromised and newly registered phishing domains as the mitigation strategies vary significantly.

4 Methodology

This section describes the datasets used in our research to investigate the newly registered domains abused for phishing and how we enrich the collected data to answer our research questions (Section 1).

4.1 Datasets

Newly Registered Domains. We collaborate with Cyber Defence Alliance (CDA), a not-for-profit member-based international threat intelligence organization that provides a daily data feed of identified malicious fully qualified domain names between July 5, 2023, and May 23, 2024. Table 1 presents an overview of our overall data collection with a summary of our data enrichments.

FQDNs	Domains	Data Collection	TLS CAs	Top-level Domains (TLDs)	Registrars	Autonomous System Names (ASNs)	Brands
15 975	15 126	07/05/2023 - 05/23/2024	34	190	209	395	58

Table 1: Summary of the data we collect and enrich for maliciously registered domains.

Our collaborator is a coalition of multiple international banks with law enforcement support. The organization identifies newly registered domains that impersonate their clients and other generic domains that target specific verticals, such as delivery companies, or themes, such as 'new payee.' These domains do not include shortened URLs. However, they receive ad-hoc domain data from their members. They use various threat-hunting resources, such as investigating real-time TLS certificates issued to domains. This feed includes all domains they identify on a daily basis that are involved with current or future smishing or phishing campaigns. Financial institutions and mobile network operators ingest their daily data feed for takedowns and blocking SMS messages containing the identified malicious domains.

Mixed Websites (Mostly Compromised). We access the aggregate values of phishing URLs from the Anti-Phishing Working Group (APWG) eCrime Exchange (eCX) repository between July 2023 and May 2024 [9]. Table 2 shows the monthly number of maliciously registered domains we receive from our collaborator and the total number of domains collected by APWG eCX during the same time period.

Month	Domains	
	Maliciously Registered (Our Dataset)	Mixed Websites (APWG eCX)
07/2023	983	19 847
08/2023	1 601	23 325
09/2023	1 892	22 327
10/2023	1 633	24 478
11/2023	1 270	20 338
12/2023	1 293	16 085
01/2024	1 679	15 360
02/2024	1 464	8 548
03/2024	1 233	12 843
04/2024	1 632	12 229
05/2024	1 295	16 543

Table 2: Monthly distribution of maliciously registered domains ($n = 15\,975$) received from the threat intelligence organization between July 5, 2023, and May 23, 2024, and domains ($n = 191\,923$) discovered by APWG eCX [9] in the same period. (849 domains in our dataset and 32 053 domains in the APWG eCX dataset appeared more than once).

The AWPG is a member-based organization that runs a central repository to exchange Internet and machine-event data on cybercrimes, such as phishing. Their members consist of individual brands, industries, NGOs, and academic organizations that contribute phishing data to eCX. This repository contains new URLs reported daily and the targeted brands, where available, along with a confidence score. Although the APWG eCX includes domains on a larger scale, it does not distinguish between newly registered and compromised domains, making it unusable to answer our research questions directly.

4.2 Data Enrichment

We enrich the daily data feed of maliciously registered domains to help answer our identified research questions (**RQ1**-**RQ4**). Towards this end, we use multiple command-line tools and third-party services, which we describe

in this subsection.

Registrar. Our data provider collects registration information as soon as they identify a new phishing domain. We augment this data collection (and check it for accuracy) using the `whois` command locally on our system for the remaining domains (and a few of the known ones). The `whois` command provides details of a domain, including its registrar [31].

IP Address. We use the `dig` command[4] to find the corresponding IP address for every domain in the daily feed as soon as we receive it. IP addresses for domains are typically thought to uniquely correspond to a hosting provider who rents server space to the domain as well as routes incoming and outgoing traffic. However, not all IP addresses in our sample correspond to hosting providers. 31.7% are unroutable (nonexistent IP addresses or no answer), and at least 17.7% are proxy services. Proxy services are an intermediary between the hosting provider and their user attempting to access the website. They have a wide variety of legitimate uses, like allowing quick access to users around the world and preventing denial of service attacks. However, when they host cybercrime websites, they additionally serve to mask the downstream hosting provider.

Autonomous System (AS). We query `ipinfo.io` [57] and Cymru's ASN command-line lookup tool [129] for every IP address to get the details of the AS that the IP address belongs to. We perform these queries in real time when we receive the daily feed. Note that the range of IP addresses changes regularly, so timeliness here is vital.

TLS Certificate. Sectigo, a certificate management solution provider, provides a database of all TLS certificates issued to a domain, `crt.sh`. We query `crt.sh` through their API [102] for every domain as soon as we receive it. This provides us with all the TLS certificates ever issued to a domain in our dataset and their certificate authorities (CAs).

Passive DNS. The Passive Domain Name System (DNS) provides IP addresses that a domain has resolved to historically by logging DNS traffic on the Internet [147]. We use Spamhaus's passive DNS API [124] to find the last timestamp when an IP address was assigned to a domain and the first time Spamhaus saw it.

Antivirus Detection. VirusTotal [137] is an antivirus aggregator service containing over 70 antivirus scanners, allowing real-time lookup on domains to see all of the scanners that flag the domain. We use VirusTotal's public API [138] to collect information on our data feed in real-time. We augment this with Google Safe Browsing's API [44] (in real-time), which allow us insight into the service perhaps before they update their information on VirusTotal. In addition to Google Safe Browsing's API, Google also provides a website – Google Transparency Report, which shows a website's blocked status [43]. We programmatically check the status of domains on the Google Transparency Report website at the same time as querying the API to identify potential discrepancies between the API and the transparency report website.

4.3 Ethical Considerations

We collaborate with an international threat intelligence organization that identifies newly registered domains and provides us with a daily data feed. This does not contain any personally identifiable information (PII). §4.1 explains the organization's methods to collect the data responsibly and ethically. We also do not collect PII data when we query `whois` or other third-party services to enrich the data. Our department's research ethics committee reviewed and approved our research.

4.4 Limitations

We collaborate with only one threat intelligence organization. Due to business confidentiality reasons, we cannot share their methodology in depth for identifying newly registered domains impersonating entities. While we do not expect false positives in their data feed, as these are detected by their automated systems and then manually verified by their analysts before sharing the list of malicious domains with us, it is still prone to human error. The identified domains impersonate their clients and other specific verticals in the finance and delivery sectors. While this data might be biased toward their clients, it is similar to any other data source. For example, all PayPal-related phishing URLs are submitted to Phishtank, which is the most dominant entity in their dataset [25]. It is possible that they are not able to identify all domains impersonating brands in these sectors. This is a similar bias to other externally gathered phishing data sources, but is still, nonetheless, pertinent.

[4] `https://www.man.page/1/dig`

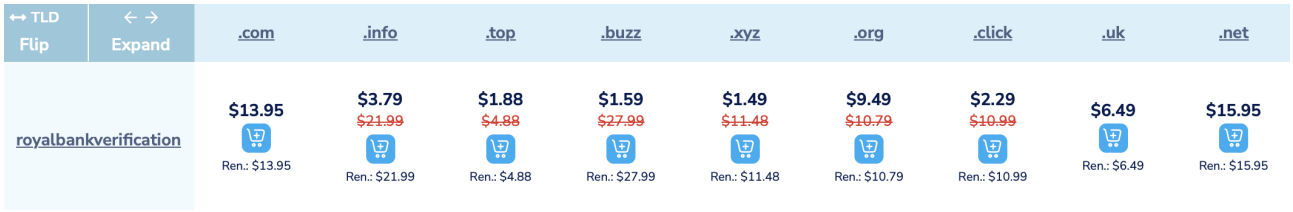


Figure 4: Price comparison for various TLDs for a domain on NameSilo. '.com' TLD is similarly priced as other available options.

5 Results

We present our insight into new maliciously registered domains abused in phishing (often SMS phishing) campaigns. Using our carefully constructed methodology, we answer our research questions towards understanding the use and abuse of maliciously registered domains. We divide this section by infrastructure operator before separately considering the domain lifetime.

5.1 Top-level Domains (TLDs)

As per the Internet Assigned Numbers Authority (IANA), more than 1 440 Top-level Domains (TLDs) are available as of May 12, 2024 [55]. We find 190 TLDs that criminals abuse to host newly registered domains identified to conduct phishing. Users cannot identify domains with unusual TLDs to conduct phishing [107]. This lends evidence to why criminals abuse various TLDs to register new domains for illicit activities.

We investigate the TLDs criminals prefer to abuse by aggregating the maliciously registered domains in Table 3. Unsurprisingly, our results reveal that `.com` is the most abused TLD (55.2%), followed by `.top` (7.3%). This is similar to the Spamhaus TLD reputation statistics as of June 2024 [122] and in line with most relatively recent academic research [62, 65, 71]. `.com` is also the most abused TLD in smishing campaigns [132]. We compare the TLD distribution of maliciously registered domains with the phishing sites in 2022 as per Kaspersky [58] and find that it somewhat follow the distribution (Table 4). However, criminals prefer to abuse the `.com` TLD significantly more for maliciously registered domains, which is not reflected from the Kaspersky phishing reports (two-proportion z-test comparing the proportion of the overall population which is registered on `.com`: $p - value < 0.0001$). This is likely due to the difference in the data analyzed by Kaspersky since they consider both maliciously registered and compromised domains. Our results will help stakeholders focus on most abused TLDs for maliciously registered domains, whereas any TLD can become a part of compromised website. This distinction is crucial for developing defense strategies and stakeholder resource allocation.

Although `.com` TLDs are perceived to be more expensive than other generic TLDs, their registration fees are often equivalent to alternate TLDs. For example, we find that the `.com` TLD registration for *royalbankverification* is similar in price compared to other TLDs, as shown in Fig. 4. We hypothesize that criminals abuse domains with `.com` TLDs to deceive users as they are easier to trust as the legitimate URLs of most global brands that criminals impersonate to conduct phishing use a `.com` TLD. The `.com` TLD is highly available, making it easy to find a wide array of potential registrars. Old TLDs like `.com` are also less likely to be auto-blocked by security vendors who sometimes block domains based on TLD [16, 32, 127].

TLDs	Domains
com	8 347
top	1 103
info	849
online	556
xyz	476
net	291
bond	237
shop	182
site	165
delivery	143
ru	143

Table 3: Top 10 top-level domains (TLDs) abused to register new domains that host phishing websites.

Similar to our results, previous research has found that old TLDs like `.com` host more bad content than new generic TLDs [62, 132]. We compare our results on TLD abuse with the usage of popular TLDs worldwide

as per W3Techs [141] to find whether this is disproportionate to the legitimate usage of popular TLDs. The distribution of `.com` TLD for phishing abuse seems to follow the distribution of `.com` TLD worldwide usage. (χ^2: $p = 0.2266$, two-proportion Z-test: $p = 0.2473$).

TLDs	Our Dataset (%)	Kaspersky Phishing (%)	TLDs	Our Dataset (%)	W3Techs (%)
com	55.18	17.69	com	55.18	46.0
xyz	3.15	13.71	org	0.56	4.5
fun	0.06	7.85	ru	0.95	3.4
org	0.56	3.89	de	0.60	2.9
top	7.29	1.80	net	1.92	2.7
ru	0.95	1.52	br	0.00	2.4
com.br	0.01	1.13	uk	0.09	2.2
co.uk	0.35	0.98	jp	0.00	1.8
de	0.60	0.98	fr	0.05	1.6
se	0.02	0.92	it	0.00	1.6
others	31.83	54.44	others	40.65	69.1
χ-squared Test	$\chi^2 = 99$, $df = 90$, $p = 0.2423$			$\chi^2 = 80.667$, $df = 72$, $p = 0.2266$	
Two Proportion z-Test	$p < 0.0001$			$p = 0.2473$	
95% CI:	(0.2420, 0.5077)			(-0.0562, 0.2398)	

Table 4: Top 10 top-level domains (TLDs) in our dataset compared to Kaspersky's phishing dataset from 2022 [58] and the most popular TLDs worldwide per W3Techs in 2023 [141]. The Two Proportion z-Test supports that the proportion of `.com` TLD distribution is same in maliciously registered domain and worldwide TLD usage.

Although increasing the prices for `.com` TLDs could deter such misuse [128], we do not recommend this as a feasible solution. The `.com` TLD is widely used. Increasing registration costs would disproportionately affect legitimate users, potentially excluding many from using this popular TLD (see Table 4). Addressing this issue requires more than just a whack-a-mole [25]. Instead, we propose that domain registrars take a more proactive role by suspending maliciously registered domains and conducting thorough due diligence before registration [73]. This strategy would help mitigate the abuse of reputable TLDs without negatively impacting legitimate users.

We classify the newly registered malicious domains as per the IANA TLD classification [56] in Table 5. A majority (93.8%) of the domains abuse gTLDs, followed by ccTLDs. Cybercriminals likely select gTLD based on the brands and sectors they target. For example, `.delivery` and `.online` host phishing websites targeting parcel delivery-related brands such as UPS (e.g., `tracking-ups.delivery`) and An Post (e.g., `anpost.delivery`) and online login portals (e.g., `myrbcprofile.online`), respectively.

TLD Group	Our Data TLDs	Our Data Domains	DomainTools Domains
generic (gTLD)	134	14 191	$217m$
country-code (ccTLD)	51	889	$122m$
generic-restricted (grTLD)	3	42	$2m$
sponsored (sTLD)	2	4	$583k$
infrastructure (.arpa)	0	0	0
test (tTLD)	0	0	0

Table 5: Domains across groups of TLDs in our dataset compared to DomainTools.

We identify 51 ccTLDs which, simply, are often cheaper TLD choices. 42 domains abuse generic-restricted TLDs, such as `.name`, `.biz`, and `.pro` while only 4 domains abuse sponsored TLDs, such as `.asia` and `.tel`. The distributions of these domains roughly follow the distribution of all domains, as per DomainTools who comprehensively count this using zone files and registry domain counts [34].

Criminals abuse free web-hosting services to host phishing websites [85, 111]. We identify only 35 (0.23%) domains that abuse Google's free Registry Domains – `.web.app` and `.firebaseapp.com`. This could be because it is easier for users to identify phishing domains with these TLDs. Google could also takedown or otherwise block these phishing websites faster since they Google's infrastructure.

5.2 Registrars

Our dataset identifies over 200 registrars that criminals abuse to register newly created domains to host phishing websites. We note that the `WHOIS` lookup [31] for 226 domains did not result in a registrar (138 of these are country-code TLDs and others are generic TLDs). We compare the top 10 registrars cybercriminals abuse to host newly registered phishing domains, against the top 5 registrars abused to host business email compromise domains as per the APWG [8, 10] in Table 6.

Our Dataset		**APWG Registrars**		**Top Registrars**	
Registrar	Domains	Q3 2023	Q1 2024	Registrar	Market Share (%)
NameSilo	3 331	Namecheap	Namecheap	GoDaddy	11.36
NiceNIC	2 182	Squarespace	Squarespace	Namecheap	2.85
Alibaba	876	1&1 IONOS	Hostinger	Tucows	1.67
GoDaddy	675	Hostinger	Tucows	Squarespace	1.22
PublicDomainRegistry	653	NameSilo	1&1 Ionos	GMO Internet Group	0.85
OwnRegistrar	561			Dynadot	0.85
Hosting Concepts B.V.	509			Ionos	0.84
Tucows	480			NameSilo	0.81
Jogjacamp	398			Gname	0.77
Reg.ru	396			Network Solutions	0.73

Table 6: Top 10 registrars abused to register newly created phishing domains ($n = 14\,900$), top 5 registrars as per APWG abused to host Business Email Compromise domains (BEC) [8, 10] and top 10 registrars by market share (as per domains registered) [33].

We identify the most abused registrar for newly registered phishing domains to be NameSilo[5], followed by NiceNIC[6]. Previous (mixed list) phishing research from 2021 [13] and a smishing research focused on US smish reports [132] found Namecheap as the most abused registrar. During the same data collection period, APWG reports Namecheap[7] as the most abused registrar, followed by Squarespace[8] to register business email compromise domains [8, 10]. Contrary to this, we only find 2.1% of malicious domains registered with Namecheap and Squarespace each. On the other hand, Spamhaus' reputation statistics are in line with our results, highlighting NameSilo as the most abused registrar associated with phishing [123].

We investigate NameSilo to understand why criminals prefer it over other registrars. We find free AI services on NameSilo that could help criminals quickly generate domain names. Fig. 5 shows a screenshot from the registrar's website where an individual can use AI to generate similar-sounding names for these options on their website. Next, we look into the second-most preferred registrar – NiceNIC. The option to buy domains through a cryptocurrency wallet such as Bitcoin likely attracts criminals to NiceNIC[9], which provides pseudo-anonymity for the transactions. NameSilo also accepts Bitcoin transactions in addition to the other available payment methods[10].

5.3 Autonomous Systems (ASes)

We look into IP addresses that the criminals abuse to host their domains. We queried every domain on the day we received the data feed. Only 10 326 (68.3%) resolved to an IP address; the remainder are either taken down by the time we receive the data or not resolving as a takedown mitigation strategy or cloaking attempt [69]. A sizeable chunk (2 681 or 26%) of resolved domains use the reverse proxy service Cloudflare. This contrasts to 19.3% of all websites which use the service [142]. This service allows users to hide their domains' IP addresses, proxying via Cloudflare's IP ranges.[11] Note that Cloudflare does not host content, per se; cybersecurity researchers would need to contact Cloudflare to identify the infrastructure hosting content for identified domains. Previous smishing research also identified Cloudflare being abused by criminals towards smishing [92]. We investigate the remaining IP addresses to identify the Autonomous Systems (ASes) that criminals abuse or collaborate with to host the newly registered domains for phishing. To this end, we find the top 10 ASes abused to host newly registered domains, as shown in Table 7. Prior work identified Amazon as the top hosting provider [92], which hosts 726 domains in our dataset.

[5] `https://www.namesilo.com/`
[6] `https://nicenic.net/`
[7] `https://www.namecheap.com/`
[8] `https://domains.squarespace.com/`
[9] `https://nicenic.net/news/messview.php?ID=21095`
[10] `https://www.namesilo.com/payment-options#alter-payment`
[11] `https://www.cloudflare.com/ips/`

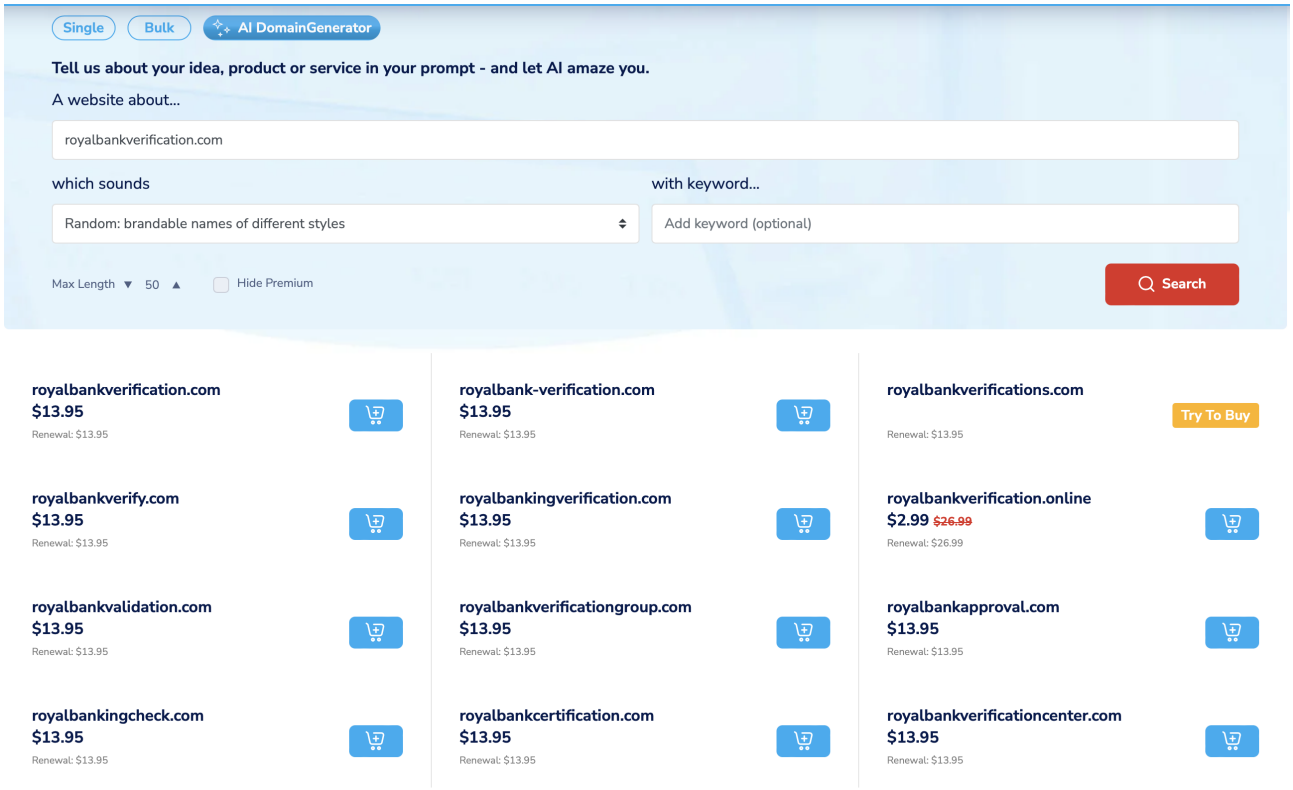


Figure 5: Generating multiple domain names using NameSilo's AI service. Most generated options have a '.com' TLD with an optimal price range.

AS Numbers	AS Names	Domains
AS200593	Prospero OOO	1 001
AS16509, AS14618	Amazon.com Inc	736
AS132203, AS45090	Tencent	421
AS45102	Alibaba (US) Technology Co.	351
AS47846	SEDO GmbH	321
AS54467, AS6134	XNNET LLC	238
AS133618	Trellian Pty. Limited	229
AS200019	ALEXHOST SRL	218
AS47583, AS204915	Hostinger International Limited	186
AS46606	Unified Layer	169

Table 7: Top 10 Autonomous Systems hosting newly registered phishing domains.

Most of the top 10 ASes are around 10 or more years old except two ASes that registered less than 2 years ago – AS200593[12] and AS54467[13]. AS200593 hosts over 1 000 domains, whereas AS54467 hosts only 236 domains. We discover that IP addresses from AS54467 are relatively evenly distributed with a mean of 1.2 domains per IP address, whereas AS200593 has particular IP addresses that are more abused than others.

We investigate all the connections between IP addresses and domain names. Using this, we identify four IP addresses from AS200593 (Prospero OOO) which together host 859 domains from our dataset. We use the Spamhaus passive DNS API to perform reverse IP lookups to find the other domains and corresponding hostnames these IP addresses have hosted in the past year (summarized in Table 8). Manually investigating these new domains hosted on Prospero OOO, we find that they have been abused to host phishing domains, impersonating several brands of financial institutions, telecom and delivery companies, among other dubious aims. We note that this AS is listed on various abuse lists for hosting malicious content and engaging in other cybercriminal behaviors [135], such as attempting to hack into WordPress websites [148]. This AS has also featured in the recent security press, detailing how the AS was tied to phishing domains targeting victims, leading them to download malware on their machines [19, 74].

[12] `https://bgp.tools/as/200593`
[13] `https://bgp.tools/as/54467`

IP Address	Our Dataset	Spamhaus pDNS	
	Domains	Domains	Hostnames
IP 1	327	2 644	2 980
IP 2	296	1 473	2 077
IP 3	121	696	886
IP 4	115	1 686	2 072

Table 8: Four IP Addresses from 'Prospero OOO' (AS200593) that host newly registered domains abused for phishing.

5.4 TLS Certificates

Criminals abuse certificate authorities to procure TLS certificates for maliciously registered domains. We find 109 570 TLS certificates issued to 11 996 (79.3%) domains by 91 different TLS issuer IDs corresponding to 34 issuing organizations. The remaining 3 130 domains never had any TLS certificates issued. Similar to previous research [13], we find that criminals abuse multiple TLS certificates for phishing domains, ranging between 1 and 4 039 certificates per domain with a median of 3 certificates.

We find that criminals abuse three major TLS certificate issuing authorities – Let's Encrypt, Sectigo Ltd, and Google Trust Services – to conduct phishing on newly registered domains. While Let's Encrypt and Google Trust Services provide SSL certificates for free, Sectigo charges money as it provides features like a much longer validity period and multi-domain SSL [114]. While it is difficult to find the ground truth, we observe that other TLS issuing organizations like `ssl.com` are not so popular among criminals. This is likely because they charge a considerable amount for even the basic service.[14] Note that while we compare other results of ours to mixed phishing datasets, recent trend reports do not include CA distribution [8, 10].

Let's Encrypt is the most abused TLS certificate issuer [1], with over $77.5k$ certificates issued to more than $10.1k$ newly registered phishing domains (Table 9). The criminals' preference to abuse Let's Encrypt is unsurprising as it issues TLS certificates at no cost, and multiple hosting platforms/control panels collaborate with them to provide their TLS certificates. Previous work analyzing certificates for maliciously registered cybercrime domains also supports this finding [60]. However, we note that Let's Encrypt TLS certificates are only valid for 90 days which could artificially inflate their incidence in our dataset.

Let's Encrypt does not examine a domain before issuing a certificate, nor does it agree that CAs are the proper intermediaries to intervene; they argue that encrypted traffic is becoming a need, not a want, and all websites should be able to be encrypted [70]. As of June 2024, Let's Encrypt provides TLS certificates to almost 47% of the top 1 million websites on the internet [20, 143] and over 67% of the newly registered phishing domains, as shown in Table 9. The proportion of phishing domains abusing Let's Encrypt follows the general proportion of websites using Let's Encrypt certificates worldwide (χ^2 test considering the distribution of all certificate authorities used worldwide vs. for phishing domains $p = 0.2773$; two-proportion z-test considering only the proportion of all websites in the sample using Let's Encrypt with $p = 0.4102$). This indicates while there is a concentration, it's about what we would expect if cybercriminals made their decisions like normal web domain operators.

Our Dataset				W3Techs	
Issuing Organization	Certificates (#)	Domains (#)	Domains (%)	Issuing Organization	Market Share (%)
Let's Encrypt	77 594	10 140	67.0	Let's Encrypt	60.4
Sectigo Ltd / Comodo CA	9 038	2 427	16.0	GlobalSign	16.1
CPanel	9 032	676	4.5	Sectigo	6.2
Google Trust Services	6 784	3 623	23.9	DigiCert	4.5
ZeroSSL	4 090	245	1.6	GoDaddy	4.0
DigiCert	1 070	464	3.1	IdenTrust	2.1
Cloudflare	970	257	1.7	Certum	0.6
GoDaddy	402	244	1.6	Actails	0.5
Apple	187	4	0	Secom Trust	0.3
Globalsign	135	25	0.2	Entrust	0.1
χ-squared test	$\chi^2 = 66$, $df = 60$, $p - value = 0.2773$				
Two Proportion z-test	$p - value = 0.4102$			95% CI of the difference in proportions: (-0.209, 0.077)	

Table 9: Top 10 TLS certificate issuing organizations abused to register new domains that host phishing websites ($n = 11\,996$) and the most widely used certificate issuing organizations as of October 2024 per W3Techs (as per absolute usage) [143].

Unlike our other measures where each domain corresponds to at most one registrar or AS, we find multiple TLS certificates for some domains. Criminals abuse multiple certificates for both domains with and

[14] `https://www.ssl.com/certificates/basicssl/`

without brand names in them. For example, a generic domain abused to impersonate multiple brands – `myonlineportal[.]net` has 4 039 SSL certificates issued to it since it was first registered by six different certificate issuing organizations, 3 860 of which were issued by Let's Encrypt and 150 by CPanel. Using the Spamhaus passive DNS service, we query the domain name for associated hosts and identify 295 subdomains in the past year.[15] This reveals subdomains such as 'verifyamznprime' and 'servmicrosftmrorllnk' used to impersonate brands like Amazon Prime and Microsoft.

The other example where criminals impersonate Royal Bank of Canada is `rbc-i[.]com` that was issued 2 588 SSL certificates, of which 1 618 were by Let's Encrypt and the remaining by Sectigo Ltd. Spamhaus's passive DNS service shows that the domain was first seen in Jan 2017 and did not have any subdomains in the past year. The number of TLS certificates indicates that the domain has been re-registered or abused to impersonate the same brand several times since then.

5.5 Brands Targeted

We investigate 15 212 domains that target over 50 unique brands, the most popular of which we highlight in Table 10. We find that financial institutions are the most targeted industry, followed by the delivery/shipping sector. This contrasts with industry reports. APWG reports social media as the most targeted industry, followed by Webmail providers and financial institutions [8, 10]. Proofpoint, an international cybersecurity company's 2024 state of phish report, ranks one delivery/shipping company as the third most abused brand over email [106]. This difference is likely due to the different data sources (mixed-websites which could be hacked or maliciously registered) that APWG and Proofpoint ingest compared to our data feed (maliciously registered domains).

Brands	Sector	Domains
Royal Bank of Canada	Financial Institution/Banking	1 268
Royal Mail	Delivery/Shipping	1 015
Santander	Financial Institution/Banking	916
DHL	Delivery/Shipping	889
Apple	Tech	775
EVRI	Delivery/Shipping	702
An Post	Delivery/Shipping	497
Barclays	Financial Institution/Banking	277
Deutsche Bank	Financial Institution/Banking	240
AIB	Financial Institution/Banking	229

Table 10: Top 10 brands impersonated to lure victims.

Criminals impersonate financial institutions and parcel delivery/shipping companies to lure victims into providing their financial details. Anderson et al. found that long-lasting fraud, like online banking fraud, has increased from 2012 [6] to 2019 [7]. Parcel delivery/shipping scams have also existed for a few years [133, 134] but have significantly increased since the pandemic [103] and even ranked the top scam in certain countries [23]. Unsurprisingly, they are common during the festive season [146].

We identify over 1 200 newly registered domains impersonating the Royal Bank of Canada (RBC). One of the reasons for this could be that criminals can set up a phishing website for this brand easily compared to other brands. This is possible with the use of phishing kits. A phishing kit is a collection of tools that allows criminals to deploy a phishing website quickly [98]. Previous research analyzing phishing kits has demonstrated their use from criminals targeting financial institutions [13, 30].

We also identify popular global brands such as DHL and Apple being targeted by criminals. Impersonating widely used brands allows criminals to lure a much more extensive population without any geographical restrictions. This indicates that there exist different threat actors that follow various strategies for phishing. As with every other paper in this area, our identified brands are both a reflection of attacks in the wild as well as our underlying data's skew based on our data providers's data collection.

5.6 Antivirus Detection

We investigate the detection by other security companies for our domains by using APIs for one large security organization, Google Safe Browsing, and the prominent antivirus aggregator, VirusTotal. We query the endpoints for each of our 15 126 domains simultaneously as soon as we receive the data feed.

We use VirusTotal to identify the number of antivirus vendors that flag the identified newly registered domains abused for phishing. 503 (3.3%) domains did not return any results from VirusTotal (Table 11). The number of antivirus vendors reporting their results for different domains ranges between 51 and 95, with a

[15] Spamhaus passive DNS provides access to only one year of historical data.

median of 91. While 80.7% domains are marked malicious by at least one antivirus vendor, only 4.3% domains are identified by more than 15 (16.5%) vendors. This indicates the myriad ways in which different providers build up their block lists [38]. Our work focuses on newly registered domains; providers that wait until a significant number of their customers interact with the domains might list these domains after we consider their blocklisting status.

VirusTotal Results	Domains
Malicious = 0 and Suspicious = 0	909
Malicious $\geq$ 1	12 206
Malicious $\geq$ 3	8 281
Malicious $\geq$ 5	6 179
Malicious $\geq$ 10	2 604
Malicious $\geq$ 15	646
Suspicious $\geq$ 1	6 714
Suspicious $\geq$ 3	297
Suspicious $\geq$ 5	1
Suspicious $\geq$ 10	0

Table 11: VirusTotal antivirus detection results of newly registered phishing domains ($n = 14\,623$).

Google Safe Browsing	Unsafe Domains	Partially Unsafe Domains	Domains Undetected	No Available Data
API	3 012	-	12 114	-
Transparency Report	3 981	1 898	4 033	5 214
on VirusTotal	2 508	-	12 618	-

Table 12: Google Safe Browsing detection results for newly registered domains abused for phishing ($n = 15\,126$) through their API, transparency report website, and on VirusTotal.

We separately consider Google Safe Browsing, despite it's inclusion in VirusTotal. That is due to disparate results based on data collection method. The Google Safe Browsing API returns an empty string if the domain is not flagged as malicious; otherwise, it returns the threat type detected for the domain, such as 'social engineering.' We find that this API detects 3 012 (19.9%) domains (Table 12).

Surprisingly, the Google Safe Browsing transparency report website returns more detailed results than the API. This website relays information about their blocklist in a format digestable for, primarily, web users. We query domains from our dataset using the transparency report website and find different results, as shown in Fig. 8. Similar to the API, the transparency report website returns "no unsafe content found" when Google does not detect any malicious content on the website (Fig. 8a). Otherwise, it returns that the site is unsafe as shown in Fig. 8b. Note that there are two additional options which are displayed to users by this web frontend. The website displays "no available data" if Google has not seen the queried domain (Fig. 8c) or a caution for "some pages on the site to be unsafe" when only a subdomain or specific path is detected as hosting/delivering malicious content (Fig. 8d). The transparency report website cannot detect 4 033 (26.7%) domains while partially detecting 1 898 (12.5%) domains. However, it returned no available data for 5 214 (34.5%) domains (Table 12).

Results on the Google Safe Browsing transparency report website are different and much more detailed than its API. This is a reflection of their priority to avoid blacklisting entire domains when only a single page or subdomain contains malicious content. However, it can provide unintuitive results for security researchers using their automatic APIs on a domain.

We also find contrasting results when comparing the Google Safe Browsing API against VirusTotal's flag for Google Safebrowsing. This is inline with prior work [104]. Google only flags 2 508 (16.6%) domains as unsafe on VirusTotal compared to 19.9% on their API and 38.9% on their transparency report website (Table 12). This is likely due to the timeliness of updating VirusTotal. We encourage those that use Google Safe Browsing data for highly time sensitive purposes to use their API or transparency report directly.

5.7 Lifetime of Domains

The lifetime of the phishing website is important, as longer-running domains defraud more victims. Threat actors use different strategies to run phishing campaigns. We consider the start date of the domain to be the date our data provider identifies a domain. The end date is when Spamhaus passive DNS saw the last time the domain was associated with an IP address, i.e., the last date it resolved to a webpage. If this isn't available, we calculate the lifetime using the closest date between Spamhaus seeing the domain or the latest TLS certificate

being issued to the domain. Spamhaus could have seen the domain before our collaborator identify it and criminals procure the TLS certificate before or as soon as they start hosting the phishing page.

To this end, we find the lifetime of 14 112 maliciously registered domains ranging between 0 and 449 days, with a median of 1 day and a mean of 8.6 days. Previous smishing research observed an average domain lifetime of 12.241 days with a median of 0.53 hours [92]. In comparison, prior phishing research found non-rock phish domains have a mean lifetime of more than two days and an average of 18 days for fast-flux domains [82]. Others have also estimated the average lifetime of mixed phishing websites to be three days [79]. The strategies criminals use to host maliciously registered domains differ from a compromised website, affecting its lifetime.

We analyze the distribution of maliciously registered domains lifetime (Fig. 6). McGrath and Gupta found 1/3rd of mixed phishing websites last 55 mins while a quarter last 12 days [79]. In comparison, our research finds that while 89.45% of the domains were active for less than 2 days, only 6.59% were active for more than 15 days. This indicates that the majority of the maliciously registered domains are short-lived. One reason could be that the newly registered domains are taken down faster. Our collaborator shares identified domains with respective stakeholders, including mobile network operators, to filter SMS messages that contain these domains. Hence, it is also possible that criminals move to other domains once the medium of abuse blocks the domain.

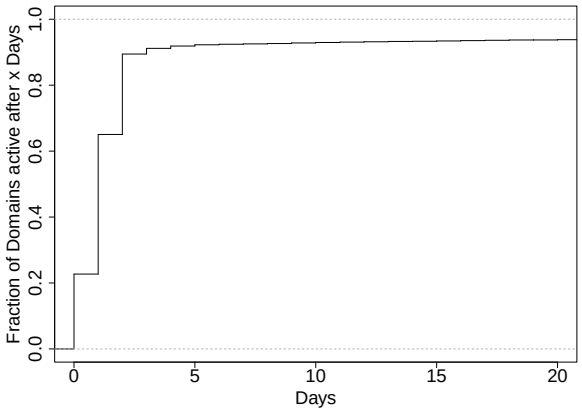
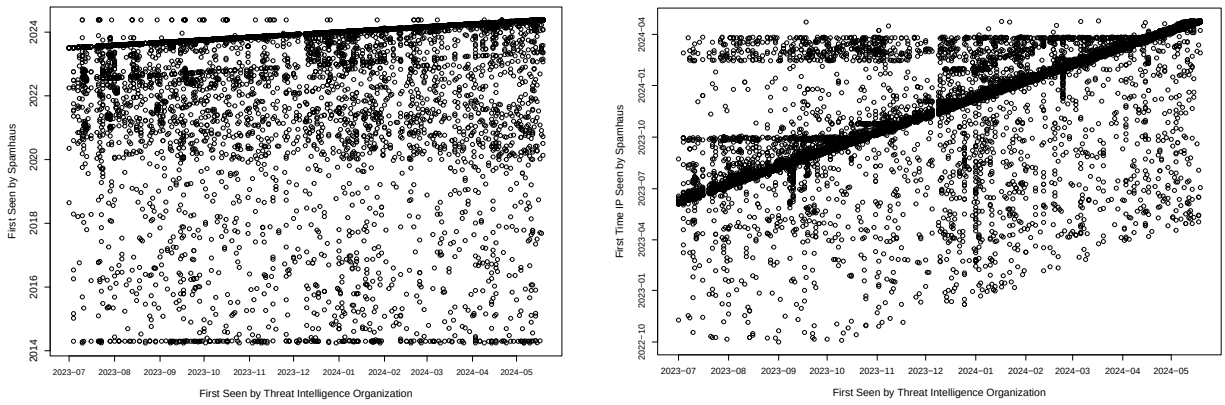


Figure 6: Cumulative distribution for the number of days newly registered phishing domains were active (n = 14 112) with a median of 1 day and mean of 8.6 days.

In section 5.4, we identify multiple TLS certificates were issued to maliciously registered domains over the years, which lends to the evidence of domains being reused for phishing. To investigate this further, we plot the first date when we receive the domains from our collaborator and the first time Spamhaus saw these domains in Fig. 7a. We find that some domains are re-registered and reused by criminals for phishing. This follows the results from previous work investigating only the `.com` TLDs [49].



(a) Comparison of the start date vs the date Spamhaus first saw the domains (n = 15 126).

(b) Comparison of the start date vs the first date an IP address was associated with the domains (n = 13 728).

Figure 7: Comparison of the start date vs the date Spamhaus saw the domains. (a) indicates criminals re-register domains and (b) visualizes the time criminals wait after registering the domain to conduct phishing.

We also plot the first date when we receive the domains from our collaborator against the timestamp collected

from passive DNS in Fig. 7b. The plot shows that the domains are abused by criminals to deploy a phishing website as soon as they are registered. However, sometimes, they wait for the antivirus vendors' detection algorithms before deploying the phishing website. This indicates that threat actors use different strategies to conduct phishing.

6 Discussion

Fighting phishing abuse from maliciously registered domains is a collaborative effort. In this section, we discuss some of the key insights we make and suggest potential mitigations for relevant stakeholders.

Cybercriminals register domains to conduct smishing. Over the years, threat actors have preferred to compromise legitimate websites to host phishing web pages [79, 86, 98, 136]. Noticing the shift of users and organizations towards SMS, criminals have started luring users into phishing via SMS. Unlike email, the complete path of the phishing page on a compromised website cannot be obfuscated in a text message. Sending the complete URL of a compromised website in plain sight with a text message that impersonates a completely different entity cannot work to lure a victim. Hence, criminals prefer to register new domains, host phishing pages that impersonates a brand and send them in SMS texts to deceive users into taking actions [118, 119]. SMS, RCS or iMessage abuse detection systems are not as sophisticated as email filters to detect and flag smishing texts containing maliciously registered domains. We suggest that stakeholders, including mobile network operators, should collaborate with threat intelligence organizations to detect and take down maliciously registered domains. Defenders should prioritize the monitoring of newly registered domains and look for signs of mass registrations or patterns that can correlate with SMS scam campaigns. Additionally, aggregators should more broadly differentiate between maliciously registered phishing domains and compromised websites, helping relevant stakeholders take appropriate actions.

Cybercriminal preferences on infrastructure to abuse changes over time. Compromised or mixed websites are what most existing studies into the phishing ecosystem focus on [79, 83, 86, 98, 115, 136]. While this provides rich insights into the victim websites hacked to serve phishing pages, takeaways are limited to this type of abuse. Takedown processes are different for webpages hosted by maliciously registered versus compromised domains. Abuse desks working at infrastructure operators need to treat these types of abuse differently – without differentiation, next steps are muddled.

The infrastructure abused by criminals has evolved over the years, depending on stakeholders' changing prices and new policies implemented to reduce domain abuse. While recent work has looked into specific TLDs for maliciously registered domains [63, 140], our work examines a broader dataset of maliciously registered domains across many TLDs. In 2020, the UK National Cyber Security Center (NCSC) found Namecheap to be the most popular host of UK-government themed SMS phishing [93]. However, we find that criminals preferred registrar is NameSilo (Section 5.2) and `.com` TLD is the most abused for registering new phishing domains (Section 5.1). This indicates that criminals preferences change over time.

Criminals maximize profit by rotating through new domains, procuring cheap infrastructure and running short-lived campaigns before getting identified. Phishing kits [98] allow them to quickly launch large number of campaigns targeting more victims in a short time. Registrars should proactively monitor the most abused TLDs to prevent threat actors registering new domains. They should ingest intelligence feeds from defenders to perform timely take downs of known maliciously registered domains. We suggest registries to closely work with their affiliated registrars that are more abused than others and share best practice guidelines to detect and deal with abuse.

Some ASes collude with criminals, allowing them to conduct illicit activities such as phishing [95, 117]. These are often referred to as bulletproof hosting providers [75, 76]. We highlight (Section 5.3) one relatively newly registered AS which is significantly more abused than others. We find over $6k$ suspicious domains in the past year hosted by four identified IP addresses within this AS's newly registered IP range. This indicates that the AS is either not proactive in detecting illicit content or knowingly hosts illicit websites. It is also possible that there exist rogue resellers who are abusing IP ranges of particular ASes for phishing. Regardless of the exact scenario, hosting providers have a significant role in performing proper due diligence and stopping bad actors from polluting the internet. We suggest that stakeholders allocate more resources toward identifying maliciously registered domains to take down relevant infrastructure efficiently.

Stakeholders overlook prior domain abuse. Once a domain is identified for abuse, defenders report it to the stakeholders and request for it to be blocked. Counterintuitively, criminals re-register domains that have been previously abused. While most of the domains in our dataset were registered for the first time, we find that criminals do frequently re-register domains to host phishing websites (Fig. 7a). These domains have been known to conduct illicit activities in the past and are being re-registered to conduct phishing. Others found that

over 14% domains taken down are released back to the market in less than a year [4]. While expired domains can be resold, there is no valid reason to procure a domain impersonating a legitimate entity. In the past, criminals re-registered expired domains owned by reputed entities to evade detection [49, 87]. However, our work indicates that they also re-register maliciously registered domains (e.g., `rbc-i[.]com` first seen in 2017). We suggest that newly registered typo-squatted domain identified for phishing abuse should be permanently banned [49]. Although some registrars already perform due diligence, a more collective effort from the industry is required to curb domain abuse.

Similarly, newly registered domains were issued TLS certificates with a median of 3 certificates per domain (Section 5.4). Given that CAs are frequently abused by criminals [13], we recommend that CAs should incorporate antivirus vendor detection results or threat intelligence feeds to assess for potential abuse before issuing certificates. We note that Let's Encrypt used to use Google Safe Browsing results before granting TLS certificates. This approach might need to be tweaked with a more relevant, mostly maliciously registered list for better effects. While compromised websites would likely result in false positives, in the case of maliciously registered domains, the CAs could leverage antivirus detection results, significantly mitigating phishing.

7 Conclusion

We investigate 15 126 newly registered domains identified by our collaborating international threat intelligence organization. Our work highlights the continuing need to work with registrars and their associated registries to prevent abuse on their platforms. There still remain inadequate levels of abuse happening on newly registered domains, which our work focuses on. While others highlight `.com` and the associated abuses, we note that `.com` does not have to be disproportionately abused in newly registered phishing domains, contrasting with other abuse vectors. We are particularly concerned with the number of malicious domains which are then re-registered year upon year, as other researchers have highlighted in years past [125]. Our work yields evidence towards the need to protect every registry from abuse. We note a similar concern with TLS certificates issued and then reissued multiple times (up to 4 039) from TLS certificate providers, an issue not limited to just free certificate providers.

While the prevalence of newly registered domains involved in online fraud writ large is not a new finding [105, 130], our work highlights how it is still a threat even today. Despite this common finding, many security vendors do not proactively handle this scourge. Given the mixed intent in the domains of other works, focusing on maliciously, newly registered domains allows us insights into behavior exhibited by cybercriminals over behavior correlated with hacked websites.

While phishing and other online fraud is illegal in many jurisdictions, it is often not pursued by law enforcement, rather by communities of people interested in combating phishing through financial or professional interests. This sets it apart from many other sorts of crimes which are primarily handled by law enforcement. While we are not suggesting to overload police with every single phishing email or domain, we call for governments to address this growing issue by proper enforcement of laws against platforms enabling abuse, as we point out when discussing AS abuse and/or proper incentives for platforms to adequately resource their abuse desks towards proactively as well as reactively preventing abuse.

Acknowledgments

We would like to thank the Cyber Defence Alliance (CDA) for collaborating with us on this research by sharing their data feed. We also thank DNS Research Federation (DNSRF) and Spamhaus for providing access to the APWG eCX and passive DNS data, respectively.

References

[1] Josh Aas, Richard Barnes, Benton Case, Zakir Durumeric, Peter Eckersley, Alan Flores-López, J. Alex Halderman, Jacob Hoffman-Andrews, James Kasten, Eric Rescorla, Seth Schoen, and Brad Warren. Let's encrypt: An automated certificate authority to encrypt the entire web. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, CCS '19, page 2473–2487, 2019.

[2] Pieter Agten, Wouter Joosen, Frank Piessens, and Nick Nikiforakis. Seven months' worth of mistakes: A longitudinal study of typosquatting abuse. In *Network and Distributed System Security*, 2015.

[3] Gautam Akiwate, Stefan Savage, Geoffrey M. Voelker, and K C Claffy. Risky bizness: risks derived from registrar name management. In *Proceedings of the 21st ACM Internet Measurement Conference*, IMC '21, page 673–686, 2021.

[4] Eihal Alowaisheq. Cracking wall of confinement: Understanding and analyzing malicious domain takedowns. In *The Network and Distributed System Security Symposium (NDSS)*, 2019.

[5] Sumayah Alrwais, Xiaojing Liao, Xianghang Mi, Peng Wang, XiaoFeng Wang, Feng Qian, Raheem Beyah, and Damon McCoy. Under the shadow of sunshine: Understanding and detecting bulletproof hosting on legitimate service provider networks. In *2017 IEEE Symposium on Security and Privacy (SP)*, pages 805–823, 2017.

[6] Ross Anderson, Chris Barton, Rainer Böhme, Richard Clayton, Michel JG Van Eeten, Michael Levi, Tyler Moore, and Stefan Savage. Measuring the cost of cybercrime. *Workshop on the Economics of Information Security*, pages 265–300, 2013.

[7] Ross Anderson, Chris Barton, Rainer Bölme, Richard Clayton, Carlos Ganán, Tom Grasso, Michael Levi, Tyler Moore, and Marie Vasek. Measuring the changing cost of cybercrime. *Workshop on the Economics of Information Security*, 2019.

[8] Anti-Phishing Working Group (APWG). Phishing activity trends report 4th quarter 2023. `https://docs.apwg.org/reports/apwg_trends_report_q4_2023.pdf`, 2023.

[9] Anti-Phishing Working Group (APWG). The APWG eCrime exchange (ECX). `https://apwg.org/ecx/`, 2024.

[10] Anti-Phishing Working Group (APWG). Phishing activity trends report 1st quarter 2024. `https://docs.apwg.org/reports/apwg_trends_report_q1_2024.pdf`, 2024.

[11] Manos Antonakakis, Roberto Perdisci, David Dagon, Wenke Lee, and Nick Feamster. Building a dynamic reputation system for DNS. In *19th USENIX Security Symposium (USENIX Security 10)*, August 2010.

[12] Timothy Barron, Najmeh Miramirkhani, and Nick Nikiforakis. Now you see it, now you Don't: A large-scale analysis of early domain deletions. In *22nd International Symposium on Research in Attacks, Intrusions and Defenses (RAID 2019)*, pages 383–397, sep 2019.

[13] Hugo Bijmans, Tim Booij, Anneke Schwedersky, Aria Nedgabat, and Rolf van Wegberg. Catching phishers by their bait: Investigating the dutch phishing landscape through phishing kit detection. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 3757–3774, August 2021.

[14] Leyla Bilge, Engin Kirda, Christopher Kruegel, and Marco Balduzzi. Exposure: Finding malicious domains using passive dns analysis. In *Network and Distributed System Security*, pages 1–17, 2011.

[15] Marzieh Bitaab, Haehyun Cho, Adam Oest, Penghui Zhang, Zhibo Sun, Rana Pourmohamad, Doowon Kim, Tiffany Bao, Ruoyu Wang, Yan Shoshitaishvili, Adam Doupé, and Gail-Joon Ahn. Scam Pandemic: How Attackers Exploit Public Fear through Phishing. In *2020 APWG Symposium on Electronic Crime Research (eCrime)*, pages 1–10, November 2020.

[16] BlueCat. Which top-level domains to block and how to do it right. `https://bluecatnetworks.com/blog/which-top-level-domains-to-block-and-how-to-do-it-right/`, 2024.

[17] Kevin Borgolte, Tobias Fiebig, Shuang Hao, Christopher Kruegel, and Giovanni Vigna. Cloud strife: Mitigating the security risks of domain-validated certificates. In *Proceedings of the 2018 Applied Networking Research Workshop*, ANRW '18, page 4, 2018.

[18] Eleanor Bradley. Domain watch-ing for phishers. `https://www.nominet.uk/domain-watch-ing-for-phishers/`, 2019.

[19] Brilliance Security Magazine. Thwarting a russian-based cyberattack on a global bank. `https://brilliancesecuritymagazine.com/cybersecurity/thwarting-a-russian-based-cyberattack-on-a-global-bank/`, 2023.

[20] SSL providers Web Usage Distribution. `https://trends.builtwith.com/ssl`, 2024.

[21] Davide Canali, Davide Balzarotti, and Aurélien Francillon. The role of web hosting providers in detecting compromised websites. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, page 177–188, 2013.

[22] Sidharth Chhabra, Anupama Aggarwal, Fabricio Benevenuto, and Ponnurangam Kumaraguru. Phi.sh/$ocial: the phishing landscape through short urls. In *Proceedings of the 8th Annual Collaboration, Electronic Messaging, Anti-Abuse and Spam Conference*, CEAS '11, page 92–101, 2011.

[23] Citizens Advice. Scams linked to parcel deliveries come top in 2023. https://www.citizensadvice.org.uk/about-us/media-centre/press-releases/scams-linked-to-parcel-deliveries-come-top-in-2023/, 2023.

[24] Richard Clayton and Tony Mansfield. A study of whois privacy and proxy service abuse. In *13th Workshop on the Economics of Information Security*, 2014.

[25] Richard Clayton, Tyler Moore, and Nicolas Christin. Concentrating correctly on cybercrime concentration. In *WEIS*, 2015.

[26] Cloudflare. What is an autonomous system? | what are asns? https://www.cloudflare.com/en-gb/learning/network-layer/what-is-an-autonomous-system/.

[27] Cloudflare. Browser market share report for 2024 q2. https://radar.cloudflare.com/reports/browser-market-share-2024-q2, 2024.

[28] Igino Corona, Battista Biggio, Matteo Contini, Luca Piras, Roberto Corda, Mauro Mereu, Guido Mureddu, Davide Ariu, and Fabio Roli. Deltaphish: Detecting phishing webpages in compromised websites. In *ESORICS 2017: 22nd European Symposium on Research in Computer Security*, pages 370–388, 2017.

[29] Scott E Coull, Andrew M White, Ting-Fang Yen, Fabian Monrose, and Michael K Reiter. Understanding domain registration abuses. *Computers & Security*, 31(7):806–815, 2012.

[30] Marco Cova, Christopher Kruegel, and Giovanni Vigna. There is no free phish: An analysis of "free" and live phishing kits. In *Proceedings of the 2nd Conference on USENIX Workshop on Offensive Technologies*, WOOT'08, pages 1–8, 2008.

[31] Leslie Daigle. Whois protocol specification, 2004. https://www.rfc-editor.org/rfc/rfc3912.

[32] Elena Deola. How to block top-level domains (tld's). https://flashstart.com/how-to-block-top-level-domains-tlds/, 2024.

[33] Domain Name Stat. Domain name registrars list - domain name stat. https://domainnamestat.com/statistics/registrar/others, 2024.

[34] DomainTools. Domain count statistics for TLDs. https://research.domaintools.com/statistics/tld-counts/.

[35] Ajka Draganovic, Savino Dambra, Javier Aldana Iuit, Kevin Roundy, and Giovanni Apruzzese. Do Users Fall for Real Adversarial Phishing?" Investigating the Human Response to Evasive Webpages. In *2023 APWG Symposium on Electronic Crime Research (eCrime)*, pages 1–14, 2023.

[36] Christine E Drake, Jonathan J Oliver, and Eugene J Koontz. Anatomy of a phishing email. In *CEAS*, 2004.

[37] eFlow. eflow fraudulent text messages. https://www.eflow.ie/news/eflow-fraudulent-text-messages/, 2024.

[38] Álvaro Feal, Pelayo Vallina, Julien Gamba, Sergio Pastrana, Antonio Nappa, Oliver Hohlfeld, Narseo Vallina-Rodriguez, and Juan Tapiador. Blocklist babel: On the transparency and dynamics of open source blocklisting. *IEEE Transactions on Network and Service Management*, 18(2):1334–1349, 2021.

[39] Federal Bureau of Investigation (FBI). Federal Bureau of Investigation Internet Crime Report 2023. https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf, 2023.

[40] Mark Felegyhazi, Christian Kreibich, and Vern Paxson. On the potential of proactive domain blacklisting. In *Proceedings of the 3rd USENIX Conference on Large-Scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More*, LEET'10, page 6, 2010.

[41] Bree Fowler. Check your messages: Scam texts on the rise. https://www.cnet.com/tech/services-and-software/check-your-messages-scam-texts-on-the-rise/, 2022.

[42] Global Anti-Scam Alliance (GASA). Scammers steal £11.4 billion from britons in 1 year as 71% fail to report scams - state of scams in the united kingdom 2024. https://www.gasa.org/post/state-of-scams-in-the-united-kingdom-2024-11-billion-stolen, 2024.

[43] Google. Google Safe Browsing - Google Transparency Report. https://transparencyreport.google.com/safe-browsing/search.

[44] Google. Safe Browsing APIs (v4). https://developers.google.com/safe-browsing/v4/, 2024.

[45] Craig Hale. Domain registrars can now block users from registering typo-laden domains for nefarious purposes. https://www.techradar.com/pro/security/domain-registrars-can-now-block-users-from-registering-typo-laden-domains-for-nefarious-purposes, 2024.

[46] Tristan Halvorson, Matthew F. Der, Ian Foster, Stefan Savage, Lawrence K. Saul, and Geoffrey M. Voelker. From .academy to .zone: An analysis of the new tld land rush. In *Proceedings of the 2015 Internet Measurement Conference*, IMC '15, page 381–394, 2015.

[47] Shuang Hao, Nick Feamster, and Ramakant Pandrangi. Monitoring the initial dns behavior of malicious domains. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, page 269–278, 2011.

[48] Shuang Hao, Alex Kantchelian, Brad Miller, Vern Paxson, and Nick Feamster. Predator: proactive recognition and elimination of domain abuse at time-of-registration. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 1568–1579, 2016.

[49] Shuang Hao, Matthew Thomas, Vern Paxson, Nick Feamster, Christian Kreibich, Chris Grier, and Scott Hollenbeck. Understanding the domain registration behavior of spammers. In *Proceedings of the 2013 Conference on Internet Measurement Conference*, IMC '13, page 63–76, 2013.

[50] Alice Hutchings, Richard Clayton, and Ross Anderson. Taking down websites to prevent crime. In *2016 APWG Symposium on Electronic Crime Research (eCrime)*, pages 1–10, 2016.

[51] ICANN. Agp (add grace period) limits policy. https://www.icann.org/resources/pages/agp-policy-2008-12-17-en, 2008.

[52] ICANN. Advisory: Compliance with dns abuse obligations in the registrar accreditation agreement and the registry agreement. https://www.icann.org/resources/pages/advisory-compliance-dns-abuse-obligations-raa-ra-2024-02-05-en, 2024.

[53] ICANN. Advisory: Compliance with dns abuse obligations in the registrar accreditation agreement and the registry agreement. https://www.icann.org/resources/pages/advisory-compliance-dns-abuse-obligations-raa-ra-2024-02-05-en, 2024.

[54] Interisle Cconsulting Group. Phishing landscape 2024. https://static1.squarespace.com/static/63dbf2b9075aa2535887e365/t/66cde404c8345e766972319c/1724769286084/PhishingLandscape2024.pdf, 2024.

[55] Internet Assigned Numbers Authority (IANA). List of top-level domains. https://data.iana.org/TLD/tlds-alpha-by-domain.txt, 2024.

[56] Internet Assigned Numbers Authority (IANA). Root zone database. https://www.iana.org/domains/root/db, 2024.

[57] IPinfo. The trusted source for ip address data. https://ipinfo.io/.

[58] Kaspersky Lab. Distribution of top-level domains used by phishing sites in 2022. https://www.statista.com/statistics/1256788/phishing-sites-tlds/, 2023.

[59] Issa Khalil, Ting Yu, and Bei Guan. Discovering malicious domains through passive dns data graph analysis. In *Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security*, ASIA CCS '16, page 663–674, 2016.

[60] Doowon Kim, Haehyun Cho, Yonghwi Kwon, Adam Doupé, Sooel Son, Gail-Joon Ahn, and Tudor Dumitras. Security analysis on practices of certificate authorities in the https phishing ecosystem. In *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security*, ASIA CCS '21, page 407–420, 2021.

[61] Panagiotis Kintis, Najmeh Miramirkhani, Charles Lever, Yizheng Chen, Rosa Romero-Gómez, Nikolaos Pitropakis, Nick Nikiforakis, and Manos Antonakakis. Hiding in plain sight: A longitudinal study of combosquatting abuse. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, page 569–586, 2017.

[62] Maciej Korczynski, Samaneh Tajalizadehkhoob, Arman Noroozian, Maarten Wullink, Cristian Hesselman, and Michel Van Eeten. Reputation metrics design to improve intermediary incentives for security of tlds. In *2017 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 579–594, 2017.

[63] Maciej Korczynski, Maarten Wullink, Samaneh Tajalizadehkhoob, Giovane C. M. Moura, Arman Noroozian, Drew Bagley, and Cristian Hesselman. Cybercrime after the sunrise: A statistical analysis of dns abuse in new gtlds. In *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, ASIACCS '18, page 609–623, 2018.

[64] Neeraj Kumar, Sukhada Ghewari, Harshal Tupsamudre, Manish Shukla, and Sachin Lodha. When diversity meets hostility: A study of domain squatting abuse in online banking. In *2021 APWG Symposium on Electronic Crime Research (eCrime)*, pages 1–15, 2021.

[65] Neeraj Kumar, Sukhada Ghewari, Harshal Tupsamudre, Manish Shukla, and Sachin Lodha. When diversity meets hostility: A study of domain squatting abuse in online banking. In *2021 APWG Symposium on Electronic Crime Research (eCrime)*, pages 1–15, 2021.

[66] Tobias Lauinger, Ahmet S. Buyukkayhan, Abdelberi Chaabane, William Robertson, and Engin Kirda. From deletion to re-registration in zero seconds: Domain registrar behaviour during the drop. In *Proceedings of the Internet Measurement Conference 2018*, IMC '18, page 322–328, 2018.

[67] Tobias Lauinger, Abdelberi Chaabane, Ahmet Salih Buyukkayhan, Kaan Onarlioglu, and William Robertson. Game of registrars: An empirical analysis of Post-Expiration domain name takeovers. In *26th USENIX Security Symposium (USENIX Security 17)*, pages 865–880, August 2017.

[68] Tobias Lauinger, Kaan Onarlioglu, Abdelberi Chaabane, William Robertson, and Engin Kirda. Whois lost in translation: (mis)understanding domain name expiration and re-registration. In *Proceedings of the 2016 Internet Measurement Conference*, IMC '16, page 247–253, 2016.

[69] Nektarios Leontiadis, Tyler Moore, and Nicolas Christin. Measuring and analyzing {Search-Redirection} attacks in the illicit online prescription drug trade. In *20th USENIX Security Symposium (USENIX Security 11)*, 2011.

[70] Let's Encrypt. The ca's role in fighting phishing and malware. `https://letsencrypt.org/2015/10/29/phishing-and-malware/`.

[71] Chaz Lever, Robert Walls, Yacin Nadji, David Dagon, Patrick McDaniel, and Manos Antonakakis. Domain-z: 28 registrations later measuring the exploitation of residual trust in domains. In *2016 IEEE Symposium on Security and Privacy*, pages 691–706, 2016.

[72] Linkt. Latest scams and security alerts. `https://www.linkt.com.au/help/security/latest-scams`, 2024.

[73] He (Lonnie) Liu, Kirill Levchenko, Mark Felegyhazi, Christian Kreibich, Gregor Maier, and Geoffrey M. Voelker. On the effects of registrar-level intervention. In *Proceedings of the 4th USENIX Conference on Large-Scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More*, LEET'11, March 2011.

[74] Adam Lopez. Hyas insight uncovers and mitigates russian-based cyberattack. `https://www.hyas.com/blog/hyas-insight-uncovers-and-mitigates-a-russian-based-cyberattack`, 2023.

[75] Dhia Mahjoub. Sweeping the ip space: the hunt for evil on the internet. In *Virus Bulletin Conference*, 2014.

[76] Dhia Mahjoub. Behaviors and patterns of bulletproof and anonymous hosting providers. *Usenix Enigma Conference*, 2017.

[77] Pratyusa Manadhata, Sandeep Yadav, Prasad Rao, and William Horne. Detecting malicious domains via graph inference. In *Proceedings of the 2014 Workshop on Artificial Intelligent and Security Workshop*, pages 59–60, 2014.

[78] Sourena Maroofi, Maciej Korczyński, Cristian Hesselman, Benoît Ampeau, and Andrzej Duda. Comar: Classification of compromised versus maliciously registered domains. In *2020 IEEE European Symposium on Security and Privacy (EuroS&P)*, pages 607–623, 2020.

[79] D. Kevin McGrath and Minaxi Gupta. Behind phishing: An examination of phisher modi operandi. In *First USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET 08)*, San Francisco, CA, April 2008.

[80] Jason Milletary. Technical trends in phishing attacks. https://insights.sei.cmu.edu/library/technical-trends-in-phishing-attacks/, 2005.

[81] Najmeh Miramirkhani, Timothy Barron, Michael Ferdman, and Nick Nikiforakis. Panning for gold.com: Understanding the dynamics of domain dropcatching. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, page 257–266, 2018.

[82] Tyler Moore and Richard Clayton. An empirical analysis of the current state of phishing attack and defence. In *Workshop on the Economics of Information Security*, 2007.

[83] Tyler Moore and Richard Clayton. Examining the impact of website take-down on phishing. In *2007 APWG eCrime Researchers Summit*, pages 1–13, 2007.

[84] Tyler Moore and Richard Clayton. The consequence of non-cooperation in the fight against phishing. In *2008 APWG eCrime Researchers Summit*, pages 1–14, 2008.

[85] Tyler Moore and Richard Clayton. The impact of incentives on notice and take-down. In *Managing Information Risk and the Economics of Security*, pages 199–223. Springer, 2008.

[86] Tyler Moore and Richard Clayton. Evil searching: Compromise and recompromise of internet hosts for phishing. In *Financial Cryptography and Data Security*, pages 256–272, 2009.

[87] Tyler Moore and Richard Clayton. The ghosts of banking past: Empirical analysis of closed bank websites. In *Financial Cryptography and Data Security*, pages 33–48, 2014.

[88] Tyler Moore, Richard Clayton, and Henry Stern. Temporal correlations between spam and phishing websites. In *Proceedings of the 2nd USENIX Conference on Large-Scale Exploits and Emergent Threats: Botnets, Spyware, Worms, and More*, LEET'09, page 5, 2009.

[89] Tyler Moore and Benjamin Edelman. Measuring the perpetrators and funders of typosquatting. In *International Conference on Financial Cryptography and Data Security*, pages 175–191, 2010.

[90] Tyler W Moore and Richard Clayton. The impact of public information on phishing attack and defense. *Communications and Strategies*, (81):45–68, 2011.

[91] Paul Mutton. Lego vs cybersquatters: The burden of new gtlds. https://www.netcraft.com/blog/lego-vs-cybersquatters-the-burden-of-new-gtlds/, 2017.

[92] Aleksandr Nahapetyan, Sathvik Prasad, Kevin Childs, Adam Oest, Yeganeh Ladwig, Alexandros Kapravelos, and Bradley Reaves. On sms phishing tactics and infrastructure. In *2024 IEEE Symposium on Security and Privacy*, pages 1–16, 2024.

[93] NCSC. Active cyber defence: The fourth year. https://www.ncsc.gov.uk/files/Active-Cyber-Defence-ACD-The-Fourth-Year.pdf, 2020.

[94] Nick Nikiforakis, Luca Invernizzi, Alexandros Kapravelos, Steven Van Acker, Wouter Joosen, Christopher Kruegel, Frank Piessens, and Giovanni Vigna. You are what you include: large-scale evaluation of remote javascript inclusions. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, CCS '12, page 736–747, 2012.

[95] Arman Noroozian, Maciej Korczynski, Samaneh Tajalizadehkhoob, and Michel van Eeten. Developing security reputation metrics for hosting providers. In *8th Workshop on Cyber Security Experimentation and Test (CSET 15)*, 2015.

[96] Yevheniya Nosyk, Maciej Korczynski, Sourena Maroofi, Jan Bayer, Zul Odgerel, Andrzej Duda, Samaneh Tajalizadehkhoob, and Carlos Gañán. Infermal: Inferential analysis of maliciously registered domains. https://www.icann.org/en/system/files/files/inferential-analysis-maliciously-registered-domains-08nov24-en.pdf, 2024.

[97] Adam Oest, Yeganeh Safaei, Adam Doupé, Gail-Joon Ahn, Brad Wardman, and Kevin Tyers. Phishfarm: A scalable framework for measuring the effectiveness of evasion techniques against browser phishing blacklists. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1344–1361, 2019.

[98] Adam Oest, Yeganeh Safei, Adam Doupé, Gail-Joon Ahn, Brad Wardman, and Gary Warner. Inside a phisher's mind: Understanding the anti-phishing ecosystem through phishing kit analysis. In *2018 APWG Symposium on Electronic Crime Research (eCrime)*, pages 1–12, May 2018.

[99] Adam Oest, Penghui Zhang, Brad Wardman, Eric Nunes, Jakub Burgis, Ali Zand, Kurt Thomas, Adam Doupé, and Gail-Joon Ahn. Sunrise to sunset: Analyzing the end-to-end life cycle and effectiveness of phishing attacks at scale. In *29th USENIX Security Symposium (USENIX Security 20)*, pages 361–377, August 2020.

[100] Ofcom. 45 million people targeted by scam calls and texts this summer. `https://www.ofcom.org.uk/phones-and-broadband/scam-calls-and-messages/45-million-people-targeted-by-scams/`, 2021.

[101] Opsec Security. How can brands fight high-volume fraud? `https://www.opsecsecurity.com/insights/blog/how-can-brands-fight-high-volume-fraud/`, 2024.

[102] A. Paul. crtsh 0.3.1. `https://pypi.org/project/crtsh/`, 2021.

[103] Kevin Peachey. Parcel delivery texts now the most common con-trick. `https://www.bbc.co.uk/news/business-58233743`, 2021.

[104] Peng Peng, Limin Yang, Linhai Song, and Gang Wang. Opening the blackbox of virustotal: Analyzing online phishing scan engines. In *Proceedings of the Internet Measurement Conference*, IMC '19, page 478–485, New York, NY, USA, 2019. Association for Computing Machinery.

[105] Pawan Prakash, Manish Kumar, Ramana Rao Kompella, and Minaxi Gupta. Phishnet: Predictive blacklisting to detect phishing attacks. In *2010 Proceedings IEEE INFOCOM*, pages 1–5, 2010.

[106] Proofpoint. 2024 state of the phish. `https://www.proofpoint.com/sites/default/files/threat-reports/pfpt-us-tr-state-of-the-phish-2024.pdf`, 2024.

[107] Florian Quinkert, Martin Degeling, Jim Blythe, and Thorsten Holz. Be the phisher – understanding users' perception of malicious domains. In *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, ASIA CCS '20, page 263–276, 2020.

[108] Resecurity. Smishing triad impersonates emirates post to target uae citizens. `https://www.resecurity.com/blog/article/Smishing-Triad-Impersonates-Emirates-Post-Target-UAE-Citizens`, 2023.

[109] Sayak Saha Roy, Unique Karanjit, and Shirin Nilizadeh. Evaluating the effectiveness of phishing reports on twitter. In *2021 APWG Symposium on Electronic Crime Research (eCrime)*, pages 1–13, 2021.

[110] Sayak Saha Roy, Unique Karanjit, and Shirin Nilizadeh. What remains uncaught?: Characterizing sparsely detected malicious urls on twitter. In *Network and Distributed Systems Security (NDSS)*, pages 400–412, 2021.

[111] Sayak Saha Roy, Unique Karanjit, and Shirin Nilizadeh. Phishing in the free waters: A study of phishing attacks created using free website building services. In *Proceedings of the 2023 ACM on Internet Measurement Conference*, IMC '23, page 268–281, 2023.

[112] Johann Schlamp, Georg Carle, and Ernst W. Biersack. A forensic case study on as hijacking: the attacker's perspective. *SIGCOMM Comput. Commun. Rev.*, 43(2):5–12, April 2013.

[113] Johann Schlamp, Josef Gustafsson, Matthias Wählisch, Thomas C. Schmidt, and Georg Carle. The abandoned side of the internet: Hijacking internet resources when domain names expire. In *Traffic Monitoring and Analysis*, pages 188–201, 2015.

[114] Sectigo. Buy SSL / TLS Certificates. `https://www.sectigo.com/ssl-certificates-tls`.

[115] Steve Sheng, Brad Wardman, Gary Warner, Lorrie Cranor, Jason Hong, and Chengshan Zhang. An empirical analysis of phishing blacklists. *Conference on Email and Anti-Spam (CEAS)*, 2009.

[116] Hossein Shirazi, Bruhadeshwar Bezawada, and Indrakshi Ray. "kn0w thy doma1n name" unbiased phishing detection using domain name based features. In *Proceedings of the 23nd ACM on symposium on access control models and technologies*, pages 69–75, 2018.

[117] Craig A. Shue, Andrew J. Kalafut, and Minaxi Gupta. Abnormally malicious autonomous systems and their internet connectivity. *IEEE/ACM Transactions on Networking*, 20(1):220–230, 2012.

[118] SIDN. Typo domain names used for smishing. `https://www.sidn.nl/en/news-and-blogs/typo-domain-names-used-for-smishing`, 2020.

[119] Silent Push. Imp-1g: Silent push tracks new us gov smishing threat actor. https://www.silentpush.com/blog/imp-1g-smishing/, 2024.

[120] Ravindu De Silva, Mohamed Nabeel, Charith Elvitigala, Issa Khalil, Ting Yu, and Chamath Keppitiyagama. Compromised or Attacker-Owned: A large scale classification and study of hosting domains of malicious URLs. In *30th USENIX Security Symposium (USENIX Security 21)*, pages 3721–3738, August 2021.

[121] Kyle Soska and Nicolas Christin. Automatically detecting vulnerable websites before they turn malicious. In *23rd USENIX Security Symposium (USENIX Security 14)*, pages 625–640, 2014.

[122] Spamhaus. Reputation Statistics | General top-level domains (gTLDs). https://www.spamhaus.org/reputation-statistics/gtlds/phishing/, 2024.

[123] Spamhaus. Reputation Statistics | Registrars. https://www.spamhaus.org/reputation-statistics/registrars/phishing/, 2024.

[124] Spamhaus Technology. Passive dns. https://www.spamhaus.com/product/passive-dns/.

[125] Jonathan M. Spring. Modeling malicious domain name take-down dynamics: Why ecrime pays. In *2013 APWG eCrime Researchers Summit*, pages 1–9, 2013.

[126] Bruce Sterling. Tainted address blocks. https://www.wired.com/2009/11/tainted-address-blocks/, 2009.

[127] Janos Szurdi. A peek into top-level domains and cybercrime. https://unit42.paloaltonetworks.com/top-level-domains-cybercrime/, 2021.

[128] Janos Szurdi and Nicolas Christin. Domain registration policy strategies and the fight against online crime. *WEIS, June*, 2018.

[129] Team Cymru. IP to ASN Mapping Service. https://www.team-cymru.com/ip-asn-mapping.

[130] Ke Tian, Steve T. K. Jan, Hang Hu, Danfeng Yao, and Gang Wang. Needle in a haystack: Tracking down elite phishing domains in the wild. In *Proceedings of the Internet Measurement Conference 2018*, IMC '18, page 429–442, 2018.

[131] Daniel Timko and Muhammad Lutfor Rahman. Commercial anti-smishing tools and their comparative effectiveness against modern threats. In *Proceedings of the 16th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, WiSec '23, page 1–12, 2023.

[132] Daniel Timko and Muhammad Lutfor Rahman. Smishing dataset i: Phishing sms dataset from smishtank.com. In *Proceedings of the Fourteenth ACM Conference on Data and Application Security and Privacy*, CODASPY '24, page 289–294, New York, NY, USA, 2024. Association for Computing Machinery.

[133] UPS. Scam Text Examples 120822.xlsx. https://www.ups.com/assets/resources/webcontent/en_GB/scam-text-examples.pdf, 2022.

[134] UPS. UPS | Examples of Fraudulent Emails.xlsx. https://www.ups.com/assets/resources/webcontent/en_GB/fraud_email_examples.pdf, 2022.

[135] URLhaus. Urlhaus asn csv feed. https://urlhaus.abuse.ch/feeds/asn/200593/, 2024.

[136] Marie Vasek and Tyler Moore. Identifying risk factors for webserver compromise. In *Financial Cryptography and Data Security*, pages 326–345, 2014.

[137] VirusTotal. VirusTotal. https://docs.virustotal.com/docs/how-it-works.

[138] VirusTotal. What is the difference between the public API and the private API? https://docs.virustotal.com/docs/difference-public-private.

[139] Thomas Vissers, Timothy Barron, Tom Van Goethem, Wouter Joosen, and Nick Nikiforakis. The wolf of name street: Hijacking domains through their nameservers. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, page 957–970, 2017.

[140] Thomas Vissers, Jan Spooren, Pieter Agten, Dirk Jumpertz, Peter Janssen, Marc Van Wesemael, Frank Piessens, Wouter Joosen, and Lieven Desmet. Exploring the ecosystem of malicious domain registrations in the .eu tld. In *Research in Attacks, Intrusions, and Defenses*, pages 472–493, 2017.

[141] W3Techs. Most popular top-level domains worldwide as of december 2023 [graph]. https://www.statista.com/statistics/265677/number-of-internet-top-level-domains-worldwide/, 2023.

[142] W3Techs. Historical trends in the usage statistics of reverse proxy services for websites. https://w3techs.com/technologies/history_overview/proxy/all, 2024.

[143] W3Techs. Usage statistics and market shares of ssl certificate authorities for websites. https://w3techs.com/technologies/overview/ssl_certificate, 2024.

[144] Yi-Min Wang, Doug Beck, Jeffrey Wang, Chad Verbowski, and Brad Daniels. Strider Typo-Patrol: Discovery and analysis of systematic Typo-Squatting. In *2nd Workshop on Steps to Reducing Unwanted Traffic on the Internet (SRUTI 06)*, July 2006.

[145] Brad Wardman, Gaurang Shukla, and Gary Warner. Identifying vulnerable websites by analysis of common strings in phishing urls. In *2009 APWG eCrime Researchers Summit*, pages 1–13, 2009.

[146] Rebecca Wearn. Parcel delivery scam texts to spike this christmas. https://www.bbc.co.uk/news/business-59760326, 2021.

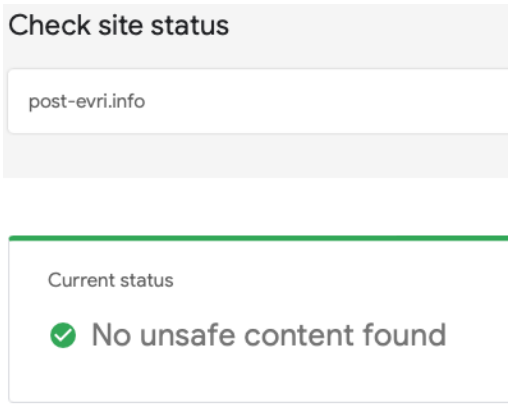
[147] Florian Weimer. Passive dns replication. In *FIRST conference on computer security incident*, volume 98, pages 1–14, 2005.

[148] WhoisXML API. Uncovering suspicious download pages linked to app installer abuse. https://circleid.com/posts/20240409-uncovering-suspicious-download-pages-linked-to-app-installer-abuse, 2024.

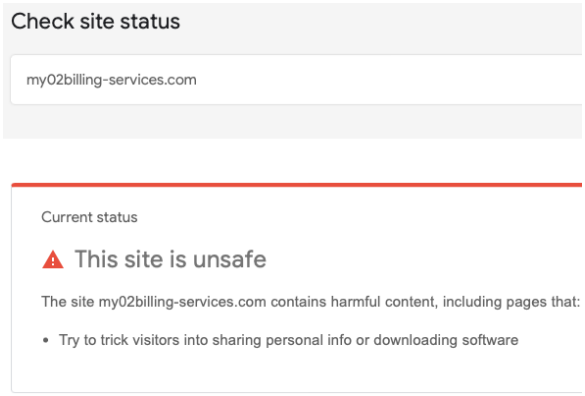
[149] Sandeep Yadav, Ashwath Kumar Krishna Reddy, A.L. Narasimha Reddy, and Supranamaya Ranjan. Detecting algorithmically generated malicious domain names. In *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement*, IMC '10, page 48–61, 2010.

[150] Yury Zhauniarovich, Issa Khalil, Ting Yu, and Marc Dacier. A survey on malicious domains detection through dns data analysis. *ACM Computing Surveys*, 51(4), July 2018.

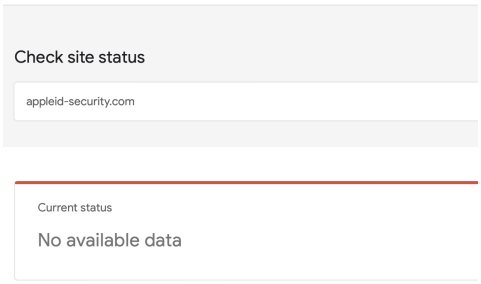
Appendix



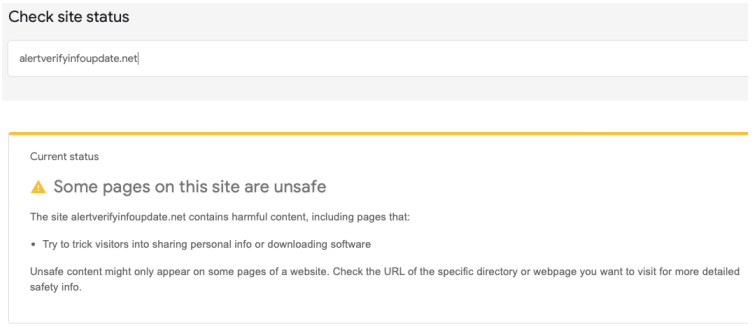
(a) No unsafe content.



(b) Domain is unsafe.



(c) Unavailable data.



(d) Warning for some pages on the domain.

Figure 8: Different possible results on Google Safe Browsing's transparency report website.