Characterizing and Mitigating Phishing Attacks at ccTLD Scale (extended)

Moura, Giovane C. M.; Daniels, Thomas; Bosteels, Maarten; Castro, Sebastian; Müller, Moritz; Wabeke, Thymen; van den Hout, Thijs; Korczyński, Maclej; Smaragdakis, G.

Publication date
2024

Citation (APA)
Moura, G. C. M., Daniels, T., Bosteels, M., Castro, S., Müller, M., Wabeke, T., van den Hout, T., Korczyński, M., & Smaragdakis, G. (2024). *Characterizing and Mitigating Phishing Attacks at ccTLD Scale (extended)*. Delft University of Technology, Faculteit Elektrotechniek, Wiskunde en Informatica.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Characterizing and Mitigating Phishing Attacks at ccTLD Scale (extended)

TU Delft Technical Report EWI-TR-2024-1
Aug. 2024, Updated Sept. 2024

Giovane C. M. Moura
SIDN Labs
Arnhem, The Netherlands
Delft University of Technology
Delft, The Netherlands

Thomas Daniels
DNS Belgium
Leuven, Belgium
KU Leuven
Department of Computer Science
Leuven, Belgium

Maarten Bosteels
DNS Belgium
Leuven, Belgium

Sebastian Castro
.IE Registry
Dublin, Ireland

Moritz Müller
SIDN Labs
Arnhem, The Netherlands
University of Twente
Enschede, The Netherlands

Thymen Wabeke
SIDN Labs
Arnhem, The Netherlands

Thijs van den Hout
SIDN Labs
Arnhem, The Netherlands

Maciej Korczyński
University of Grenoble Alps
Grenoble, France

Georgios Smaragdakis
Delft University of Technology
Delft, The Netherlands

Abstract

NOTE: this paper is an extended version with appendices from the original ACM CCS 2024 paper [65].

Phishing on the web is a model of social engineering and an attack vector for getting access to sensitive and financial data of individuals and corporations. Phishing has been identified as one of the prime cyber threats in recent years. With the goal to effectively identify and mitigate phishing as early as possible, we present in this paper a longitudinal analysis of phishing attacks from the vantage point of three country-code top-level domain (ccTLD) registries that manage more than 8 million active domains – namely the Netherlands' `.nl`, Ireland's `.ie`, and Belgium's `.be`. We perform a longitudinal analysis on phishing attacks spanning up to 10 years, based on more than 28 thousand phishing domains. Our results show two major attack strategies: national companies and organizations are far more often impersonated using malicious registered domains under their country's own ccTLD, which enables better mimicry of the impersonated company. In stark contrast, international companies are impersonated using any domains that can be compromised, reducing overall mimicry but bearing no registration and financial costs. Although most research works focus on detecting new domain names, we show that 80% of phishing attacks in the studied ccTLDs employ compromised domain names. We find banks, financial institutions, and high-tech giant companies at the top of the most impersonated targets. We also show the impact of ccTLDs' registration and abuse handling policies on preventing and mitigating phishing attacks, and that mitigation is complex and performed at both web and DNS level at different intermediaries. Last, our results provide a unique opportunity for ccTLDs to compare and revisit their policies and impacts, with the goal of improving mitigation procedures.

Keywords

Phishing, DNS, ccTLD, Registry, Mitigation, Policy

1 Introduction

Phishing is an online scam enticing users to share private or financial information using social engineering and other fraudulent tactics. Web phishing is among the most popular forms of phishing, where a user is maliciously redirected to a website under a criminal group's control, *e.g.,* by clicking a link in a text or e-mail message.

Various security and investigative organizations have identified phishing as one of the most important cyber threats. Phishing is the top digital crime type identified by the U.S. Federal Bureau of Investigation (FBI) in its 2023 annual report [92]. They report more than 300 thousand complaints and more than US$160 million in losses that were directly attributed to phishing, and hundreds of millions of losses to Personal Data Breaches that utilized phishing as a social engineering and attack vector tactic in 2022 alone. The U.S. Cybersecurity and Infrastructure Security Agency (CISA) has also identified phishing as one of the most important cyber threats to national security [23]. The European Union Agency for Cybersecurity (ENISA) also ranks phishing as one of the top three cyber threats in its Threat Landscape for 2023 highlighting that "phishing is once again the most common vector for initial access" [22]

for individuals, corporations, and governments in Europe. Reports from other organizations around the world conclude that phishing continues to be one of the most harmful cyber threats today [44–46].

Previous works characterizing phishing attacks have been restricted to specific impersonated companies and relatively short time frames. For instance, by analyzing web server logs, Oest et al. [71] analyzed a year's worth of phishing data targeting a large, undisclosed payment provider. In another study, Bijmans et al. [12] focused on phishing attacks against ten banks in the Netherlands over five months, identifying 1,363 domains used in these attacks by analyzing certificate transparency logs [15].

In this paper, we address data and timeframe limitations by collaborating directly with country-code TLD (ccTLD) registries, which possess the necessary historical and longitudinal datasets. We partner with three European ccTLDs operators: the Netherlands' `.nl` (managed by SIDN [80]), Ireland's `.ie` (managed by .IE Registry [42]), and Belgium's `.be` (managed by DNS Belgium [11]).

This collaboration allows us to access longitudinal datasets spanning up to 10 years, including phishing blocklists, domain name registration records, and active and passive DNS and HTTP measurements. In total, we characterize *28,754 phishing domains used in attacks against 1,233 companies from 86 countries.*

Our study also allows us to perform "natural experiments," *i.e.,* non-controlled experiments for variables determined by ccTLDs in the past, to understand the impact of ccTLD registration and phishing mitigation policies. For the first time, it is also possible to study the influence of ccTLD policies on phishing mitigation in the long chain of intermediaries that includes DNS providers, hosting providers, registries, and registrars.

Our contributions can be summarized as follows:

- We perform a longitudinal analysis spanning up to 10 years for the three ccTLD registries, covering more than 28 thousand domains that were used as attack vectors for web phishing attacks.
- Our analysis shows that there are two phishing strategies (§4). National companies are often impersonated using newly registered domains. This class of attacks seems to leverage users' brand trust and fluency in the local official language. In stark contrast, international companies are typically impersonated by old domains, likely compromised websites. These companies are typically global brands that are attacked by criminals that exploit legitimate domains and without the cost or burden of registering new domains.
- We evaluate the market segments and show that banks, financial institutions, and technology companies are most impersonated in both phishing strategies (§4). Our analysis also shows that 11% of the impersonated companies account for more than 58% of all second-level domains (SLDs) used for phishing attacks. We show that compromised domains dominate the phishing landscape, contributing to more than 80% of SLDs that are weaponized in phishing attacks.
- We demonstrate how the ccTLD registration policy has an impact on reducing phishing activity (§4) . Restricted registration virtually eliminates phishing attacks that utilize new maliciously registered domains, but it does not prevent attacks using compromised domain names, which form the bulk of the attacks.
- Our analysis shows that mitigation is complex and varies significantly depending on the ccTLD registry and its policies (§5). We see that DNS and web mitigation are often used independently and simultaneously, and the registry can only cover the DNS mitigation part.
- We present a call for action (§6) for the research and operations communities. First, we need more research on detecting compromised domains, given they form the bulk of attacks and most studies focus on new domain names. Secondly, we demonstrate that ccTLDs should share threat intelligence given companies from one country are often attacked using domains from another country's ccTLD. Lastly, our results present an unique opportunity for registries to compare each other's practices. These results are under discussion within the three ccTLDs, spanning legal, management, and abuse handling departments. Consequently, policy adjustments may be implemented to enhance abuse mitigation.

2 Background

2.1 ccTLD registry operations

Our work leverages data from three ccTLD registries: `.nl`, `.ie`, and `.be`. A DNS registry's main task is to run the domain registration process, curating a registration database that lists details about registered domains under their ccTLD. This includes the registrant's personal data (name, address, e-mail), registration and expiration dates and the domain's associated authoritative DNS servers.

Domain registration: It starts with a *registrant* (*e.g.,* a person) requesting a *registrar* (*e.g.,* GoDaddy) to register a domain name with a registry (*e.g.,* Verisign for `.com`). The registrar then executes this request on behalf of the registrant, once payment and other checks are cleared. Domain registrations can also be requested by resellers, which are intermediaries between a registrant and registrar. Thus, the registrar will register on behalf of resellers. To provide a perspective on the size of this industry, the `.nl` zone encompasses over 1k+ accredited registrars and 46k+ resellers.

Registration policy: In our study, both `.nl` and `.be` have open registration policies, while `.ie` has a restricted registration policy [40] – which means that only individuals and businesses related to Ireland can register `.ie` domain names (having an ID, passport, or business relation), whereas `.nl` and `.be` are open to anyone.

Domains are typically registered for a period of time, often one year, although `.nl` allows for registration periods of three and six months as well, with the option for renewal thereafter. Once registered, domain names are inserted into the DNS *zone file* [62], which contains the authoritative name servers for all delegated domains under the TLD. These zone files serve as input for *authoritative DNS servers*, which are a type of DNS server which knows the "contents of the zone from local knowledge" [31]. With this information, a domain name can be resolved by *recursive resolvers*, which, on behalf of users, map domain names to IP addresses.

2.2 Evaluated ccTLDs and other TLDs

As can be seen in Table 1, the demographics of the countries as well as the registration policies and pricing for the evaluated ccTLDs differ. In total, the ccTLDs we evaluate in this study have more than 8M active domain names.

Table 1: ccTLDs overview.

ccTLD	.nl	.ie	.be
# Domains	6.1M	330.1k	1.7M
Registration Policy	Open	Restricted	Open
Country Population	17.5M	4.9M	11.5M
Domains/1k Inhabitants	350	67	147
Domain cost price (€)	3.55	1.25	4.00
Official Languages	1	2	3
GDP per capita (€)	61k	105k	53k

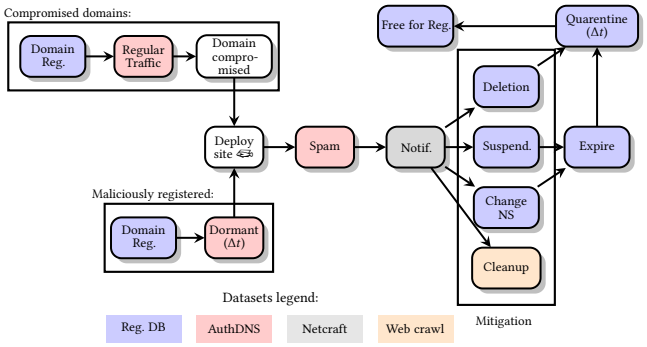


Figure 1: Phishing domains life cycle.

Other TLDs: The largest *generic* TLD (gTLD) is .com, which has 157M domain names, followed by .net, with 12M (as of July 18, 2024) [94] – however, both of those are *generic* TLDs used globally. Our ccTLDs are primarily used by businesses and individuals in their respective countries. Germany's .de is the largest ccTLD (17.4M in 2022 [19]), followed by The United Kingdom's .uk (9.1M in 2024 [69]). .nl is the third-largest ccTLD (6.1M). While many TLDs do not disclose their size publicly, some industry reports exist [14, 21], although often incomplete. Both .nl and .be are popular in their own countries, whereas in Ireland many companies often opt to use .com domain names instead of .ie.

2.3 Phishing domain types

Phishing attacks resemble aggressive mimicry in the biological world, where the mimics *deceive* their prey [66], embodying the proverbial "wolf in sheep's clothing" scenario.

We observe two primary categories of phishing domain names: *maliciously registered* and *compromised* domains [16, 55, 58, 61, 84]. Maliciously registered domains are those in which attackers register it themselves and take care of the associated steps – configuring its DNS and hosting the phishing website (Figure 1). These domains may be immediately used after registration to carry out attacks, or the attacker may wait for some time before carrying out the attacks – which is referred to as "aged domains".

Compromised domains, on the other hand, occur when attackers exploit someone else's website to host a phishing site. They resemble parasitic relationships in the natural world, where parasites live off (or in) another organism [73]. Typically, this is achieved by exploiting vulnerabilities in websites, particularly those using susceptible content management systems (CMSes) such as WordPress and Joomla [90, 93]. While it is the website that is exploited, for consistency, we refer to them in the paper as "compromised domain names". We include examples of of three real-world phishing sites in Appendix A .

Phishing Kits: Phishing sites can be built with the help of phishing kits and phishing as a service, with templates tailored to specific organizations [12, 27]. LabHost, a phishing-as-a-service provider busted in 2024 [24, 70, 89], offered convincing phishing sites mimicking 170 organizations covering "financial institutions, postal delivery services and telecommunication services providers, among others" [24], with subscription fees averaging US$ 249 per month.

After that, the phisher must actively promote the phishing URLs, typically resorting to spam methods (*e.g.,* e-mail, social networking). This initiates a race against time, as the detection of phishing activities may become imminent, which may lead to mitigation.

2.4 Domain name price influence in abuse

Previous research and industry reports highlight that pricing is an important factor for attackers when choosing the TLD and registrar for domain name registrations. Free SLDs were previously offered by Freenom, which caused many crooks to use them for phishing [1, 2, 72]. After being sued by Meta [72], Freenom stopped offering free SLDs, and at the moment of this writing, no registry offers free SLDs.

When China's .cn registry changed its registration requirements (from open to requiring IDs to register domains) and increased the price roughly tenfold, the volume of blacklisted .cn domain names reduced while the number of Russia's .ru spamming domains increased [59]. These results suggest migration from TLD responding to policy and price change (waterbed effect).

Another study [49] pointed a shift from abuse from legacy gTLDs (*e.g.,*.com) to new gTLDs (*e.g.,*.xyz) – some of the new gTLDs offer domains for less than US$1. For comparsion, a .com SLD costs US$10.26 [20]. The results suggest that low domain price correlates with abuse, although they could not fully confirm this due to a lack of comprehensive pricing data. It is important to note that pricing is only relevant to maliciously registered domains. Compromised domain do not incur any registration costs on the attacker side.

2.5 Phishing mitigation

Phishing mitigation can occur at two application levels: DNS and Web, or both. At the DNS level, the domain name used in phishing can be *deleted* from the namespace and from the zone file. It can also be *suspended*, where it is deleted (delisted) from the DNS zone but not from the namespace. Lastly, it can remain in the zone and namespace, but have its authoritative DNS servers (NS records) changed to a safe server.

In all three methods, the domain name will not resolve any longer to the phishing website. These mitigation options can be performed at the reseller (if there is one), registrar, or at the registry. These entities have the technical and legal capabilities of doing so.

At the Web application level, the phishing website can be cleaned up. This action can be done by the hosting providers and webmasters. It is also important to patch vulnerable web software for compromised domains, otherwise they can be exploited again.

Mitigation per domain type. Compromised domains and maliciously registered ones require different mitigation strategies [84]. Compromised websites should not be mitigated at the DNS level, given it will make the website unreachable, impairing the operations of legitimate websites. Web mitigation is recommended first (Cleanup in Figure 1). For malicious domain names, it is not enough to mitigate at one application level: if a domain is suspended, the attacker can register a new domain pointing to the phishing website. If the website is mitigated, then the attacker can redirect the domain name to another website. That said, one mitigation type partially solves the problems during an attack. Mitigating at both DNS and Web levels increases the economic cost for the attacker, reducing the incentives to use the chosen DNS zone. If deleted, the domain name subsequently expires, eventually reentering the pool of available domains for registration, after a quarantine period. This period lasts 40 days for `.nl` and `.be`, and 45 days for `.ie`.

Other mitigation methods include blocklists [56], such as Google SafeBrowsing [26], which warn users in their browsers about the threat. Their efficacy is predicated in the ability and speed in which phishing websites are detected by the blocklist. Upstream providers, which are not technically responsible for content, may pressure clients to mitigate phishing attacks or cut service [32, 50, 57, 87].

2.6 TLDs and abuse handling policy

The Internet Corporation for Assigned Names and Numbers (ICANN) oversees the delegation of TLDs to registries [43]. ICANN's policies differ for country code TLDs (ccTLDs) and generic TLDs (gTLDs, e.g., .com, .org) due to their unique characteristics.

ccTLDs are generally regarded as national resources. As such, ccTLDs lack a contract with ICANN, preventing ICANN from enforcing abuse mitigation procedures. Instead, ccTLDs and ICANN have *formal agreements*, known as "Exchange of Letters" [37], where registries pledge to properly operate their country's ccTLD zone. These agreements do not encompass abuse handling, allowing ccTLDs to establish their own abuse handling policies and procedure in consultation with the local Internet community.

Generic TLD (gTLD) registries, however, have a *contractual relationship* with ICANN [36]. Proposed amendments in Dec. 2023, if ratified, will mandate gTLD registries to proactively mitigate abuse ([38], and §4 in [35]). ICANN also requires that its accredited registrars – which are able to register gTLD domains such as `.com` must adhere to specific abuse handling guidelines (§3.18 in [34]).

2.7 ccTLDs' phishing handling procedures

Table 2 summarizes the phishing mitigation procedures employed by the ccTLDs we evaluate in this paper. Their detailed abuse mitigation policy is covered in Appendix B .

Phishing detection: all three ccTLDs utilize abuse reports provided by Netcraft and accept external abuse complaints. In addition, `.ie` manually verifies each registration due to its restricted registration policy, before the domain registration is finalized.

`.nl` and `.be`, in turn, in the evaluated period, employed both manual and automated (ML-based) verification. During the evaluated period, `.be` used a rule-based system (now replaced by an ML-based one [10]). Such system flags roughly 30% of new domain registrations to be followed-up by a registrant identity verification. These

Table 2: ccTLDs phishing mitigation procedure.

	`.nl`	`.ie`	`.be`
Detection			
Netcraft	✓	✓	✓
Complaints	✓	✓	✓
Manual verification	✓	✓ (strict)	✓
ML-based verification	✓	–	✓
Notification	3rd-party (Netcraft) and Registrar	3rd-party (Netcraft)	Registrant and Registrar
Phishing Mitigation			
Suspend	✓ After 66h	✓ After 30 days	✓ASAP
Delete	✓	✓After two weeks	✓
Change NS	–	–	✓

`.be` domains are not delegated in the zone until this procedure is concluded. As such, this inhibits many potential phishing attacks – but we cannot measure it given our datasets only show confirmed attacks. `.nl`, in turn, has manual and automated processes that flag domains similarly, but differently from `.be`, it does not prevent delegation: the registration verification process happens *after* the domain is delegated.

Phishing notifications: Regarding notifications, all ccTLDs directly notify the respective registrars. `.nl` and `.ie` also hire the notification service from Netcraft's, which informs all parties known to be associated with the phishing attack, including DNS and upstream providers, and registrars. `.be` is the only one to also notify the registrants.

Moreover, `.nl` also notifies Netcraft of phishing sites its analysts detected independently – and therefore uses Netcraft's notification systems to notify the responsible parties. In this period, 1,552 `.nl` from the 25,389 SLDs (6.1%) were detected by `.nl` analysts and added to the Netcraft list.

Phishing mitigation: all three ccTLDs have policies granting them the right to remove domain names from their zone and delete them from their namespace, as discussed in §2.5 and shown in Table 2.

Each registry, however, has different moments to suspend domains. `.be` is the fastest, suspending as soon as possible after the Netcraft notification. `.nl`, in turn, will wait for 66h before suspending the domain [78]. `.ie`, in turn, will suspend a domain after 30 days if the registrant doesn't respond to contact attempts, or exceptionally within two business days if deemed harmful.

`.be` has a policy to first suspend the NS records and delete a domain name after two weeks if the registrant is unable to prove its legitimacy – so the initial mitigation (suspension) is followed by a deletion. In its transparency report, `.nl` states that it took down 9,305 domain names in 2024 – not only related to phishing, but multiple types of abuse [81].

3 Datasets

We leverage two main types of datasets in this study: a phishing blocklist (§3.1) and the registries' own datasets (§3.2).

Table 3: Netcraft phishing blocklist dataset.

	.nl	.ie	.be
Starting date	2013-09-16	2019-07-30	2019-08-29
Ending date	2023-06-05	2023-08-25	2023-06-05
Period	~10 years	~4 years	~4 years
Domains (SLDs)	25,389	555	2,810
URLs	137,880	4,542	27,346

3.1 Netcraft phishing blocklist

We begin our analysis with Netcraft [68], a commercial phishing blocklist and a leading industry threat intelligence provider. It catalogs URLs associated with phishing attacks, along with pertinent metadata. Netcraft is well-regarded among registries for its low false positive rates. We use only a *subset* of Netcraft's phishing blocklist, focusing on phishing attacks occurring within the ccTLDs relevant to this study — each registry has a paid subscription to only events in their zone.

Netcraft detection methods: Netcraft's precise detection methods are proprietary (their "secret sauce"). According to their website [67], they utilize various data sources including DNS queries, web crawls, app store searches, multiple search engines and social media searches, Domain-based Message Authentication, Reporting, and Conformance (DMARC) [51], and X.509 certificate analysis used in TLS connections.

We contacted their engineers and they said they employ multiple vantage points across the globe, "including residential and mobile networks as well as other techniques like user agent cloaking". Their tools also interact with forms, and according to them, their "automation to appear as if it were a real visitor is sufficient to avoid CAPTCHA interstitial pages." After detection, "a mixture of automation and human expertise is used to validate phishing URLs and automation is responsible for the large majority of these". Given we do not have access to their detection systems, we cannot verify these claims, we only use their final output: the blocklist itself.

Table 3 shows our phishing datasets – the subsets of the Netcraft blocklist for phishing, for each ccTLD. We see roughly 25k SLDs over a 10-year period for .nl, 555 for .ie, and 2,810 for .be, both over a 4-year period. Given a single SLD may have multiple phishing URLs, we see a larger number of URLs.

Figure 2 shows the timeseries of SLDs used for phishing attacks for the ccTLDs we evaluate. We also include a linear regression of the number of monthly SLDs in Figure 2, showing a decrease in the number of SLDs used over the year for all ccTLDs.

Alternative phishing feeds. Besides Netcraft, we also evaluated another feed, namely the Anti Phishing Working Group (APWG) phishing feed [5], for the same period. We found that APWG listed far fewer domains than Netcraft (roughly 10% for .nl, Table 4), and that most of them were already included in Netcraft. The extra domains provided by APWG would account for 2.2% of the total for .nl. Given that most of them were included and did not have manual validation of the remaining entries, we decide to not use them in this study.

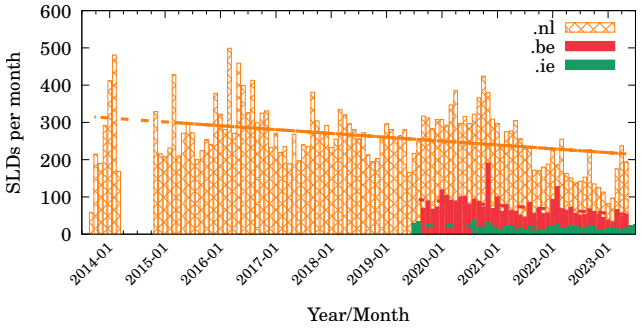


Figure 2: Phishing dataset timeseries (SLDs)

Table 4: Netcraft and APWG SLDs listed.

	.nl	.ie	.be
Netcraft	25,389	555	2,810
APWG	2,770	280	297
∩	2,199	26	222
APWG only	571	254	75

Limitations. As any blocklist, the datasets we use are from a reactive nature – meaning that they report only known attacks [57]. This reactive approach leaves a window of vulnerability between the emergence of a new threat and its eventual addition to the blocklist.

3.2 Registries' datasets

We use datasets available at the registries, including the domain names registration database, authoritative DNS traffic and for .nl, we have web crawls of domain names. These web-crawl datasets are constructed using an in-house web crawler at the .nl registry (DMap[95]). This crawler scans the entire .nl zone monthly and all newly registered domains every 8 hours, for a period of 30 days. This crawler uses a headless Chromium browser to retrieve web data associated with domain names.

Compared to other studies, which typically relied on partial information about domain name registration (using whois [17] or commercial services), our datasets cover all details for every domain in the DNS zone, both current and historical. Access to these datasets is governed by the European and national laws of each ccTLD's country's jurisdiction. It is important to note that during the course of this research, while collaboration was ongoing, data was not shared among the ccTLDs. Instead, we shared code to analyze the datasets locally.

Dataset sharing: Unfortunately, we cannot share our raw datasets, only aggregated ones (see Netcraft's .nl phishing counts in [53]). The Netcraft blocklist is a commercial dataset and cannot be publicly released. Our registries' datasets contain personally identifiable information (PII), which prohibits public release due to the EU's GDPR and national laws.

4 Impersonated companies

ccTLDs, by definition, have a strong association with their countries: governments use them in their e-gov services [86], as well as local

Table 5: Target companies and market segments per ccTLD.

	.nl	.ie	.be
Targeted companies	1,057	206	546
Local	58	8	33
International	942	198	460
Unknown	57	0	53
Companies' countries	78	18	64
Market segments	114	52	108

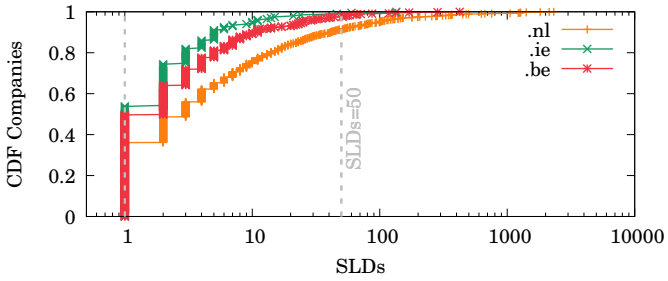


Figure 3: SLDs CDF for impersonated companies per ccTLD.

companies and services. As such, users are continuously exposed to domains under their respective ccTLD. To a certain extent, a ccTLD may be seen as a *brand*, and humans assign trust levels to brands [47, 88]. Brand trust is a critical aspect of consumer behavior and is a key factor in the success of a brand [74]. Next, we investigate if attackers exploit this association (or trust) in their attacks.

4.1 Companies' profile

We profile the impersonated companies we observed in our phishing datasets (§3.1). In the datasets, each URL includes a company (brand) which was impersonated in the attack. We then extract all listed companies and manually identify both their country of origin and the market segments they operate in. In the case of multinational companies, we use the country of their headquarters. For instance, we classify Netflix as a US-based company, even though it operates globally.

Table 5 shows the results. We see a large number of companies, most of them being *international, i.e.,* they are companies based in countries other than those of their ccTLDs (for some companies we could not find out where they operated from, such as some cryptocurrency companies, so we label them as "Unknown"). As such, we can clearly see that ccTLDs are used for phishing attacks *mostly against international companies outside the ccTLDs*, and not to exploit companies in the countries of the ccTLDs.

Companies phishing attack concentration: Next, we compute the distribution of the number of SLDs used against each company. Figure 3 shows the CDF for all ccTLDs. Across all ccTLDs, we see two patterns: some companies are targeted using one or a few SLDs, while others experience recurrent attacks. Most companies were targeted using a single SLD (36%, 54%, and 49% of their companies for `.nl`, `.ie`, and `.be`, respectively) or a few. However, few companies were attacked using at least 50 SLDs (9%, 0.1%, 3%).

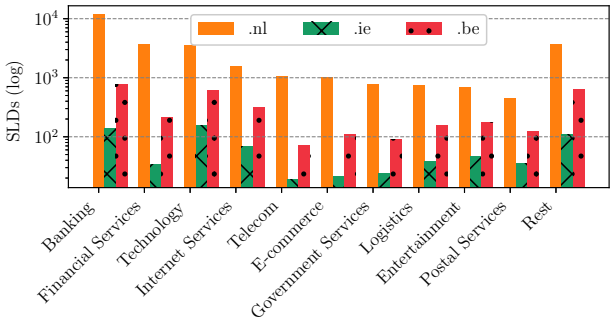


Figure 4: Targeted economic sectors.

Diverse market segments. We see many attacked market segments in Table 5. Figure 4 shows the counts of SLDs per segment. We see most phishing attacks are against banks and financial institutions, followed by technology companies. Banks also top the list of APWG phishing reports [6]. Examples of banks include Dutch banks (ING, ABN AMRO), and JP Morgan Chase (US).

The second most popular market segment is technology companies, which include US-based companies such as Microsoft and Apple. Then, we have Internet Services providers such as Yahoo (US), AoL (US) and NetEase (CN). In fourth place, we have Telecoms, such as Freebox (FR), Vodafone (UK), Orange (FR), and a Ziggo (NL). Coincidentally, these market segments were also present in the portfolio of LabHost (§2.3), the phishing-as-a-service provider.

4.2 Companies and country of origin

Next, we compute the number of companies, SLDs, and average and median age of the SLDs used in the attacks, for each country from the impersonated companies. We compute the SLD age by subtracting the phishing notification time from the domain registration time, which we obtain from the registries' databases.

American companies are the most popular. Table 6 shows that most impersonated companies are American, for all ccTLDs. Local companies rank #2 in `.nl` and `.be` zones, but not in `.ie`. Given that `.ie` has a restricted registration policy (§2.1), the results suggest it inhibits phishing against local companies.

Local phishing uses new domains, International uses compromised domains: Local companies, for `.nl` and `.be`, are attacked with new domain names – they have a median of fewer than 2.5 days – which suggests these are maliciously registered domains. That's a sharp contrast with the age of SLDs targeting American companies. The domains used in these attacks have, on average, more than 5 years of age. Given that these domains are old, we believe them to be *compromised domain names* (§2.3).

Most impersonated companies. Table 7 shows the top 5 impersonated companies for the ccTLDs, in terms of SLDs. We see that Microsoft tops the rank for all ccTLDs, with a median age of more than 5 years. Microsoft is an attractive target, given its large user base, and access to such accounts can provide them with e-mails and also a range of internal documents. Microsoft also tops the list of other industry phishing reports [85]. Similarly to Microsoft, we see PayPal among the top 5 for all ccTLDs. We see that two local business, a local bank (#3) and a financial service provider (#4),

Table 6: Impersonated companies and their countries.

The Netherlands ccTLD (.nl)

Country	Companies	SLDs	URLs	Avg. Age	Med. Age
All	1,054	26,740	125,304	1,829.6	1,641.0
US	268	13,205	69,851	1,600.0	1,751.5
NL	58	6,084	31,063	208.2	2.5
FR	55	1,532	8,434	1,762.0	1,766.2
GB	66	1,096	5,327	1,991.5	2,163.7
Rest	607	5,693	23,123	1,737.0	1,988.5

Ireland ccTLD (.ie)

Country	Companies	SLDs	URLs	Avg. Age	Med. Age
All	206	555	4,542	3,062.9	2,746.0
US	66	363	2,781	3,406.1	3,310.0
FR	15	47	574	2,475.6	2,102.0
DE	12	44	293	2,986.2	2,958.0
GB	9	16	48	2,849.0	2,834.0
Rest	92	135	782	2,932.1	2,445.5

Belgium ccTLD (.be)

Country	Companies	SLDs	URLs	Avg. Age	Med. Age
All	546	2,810	27,346	3,231.1	2,154.0
US	126	1,647	17,095	3,022.1	1,922.5
BE	33	254	1,480	1,522.1	2.0
FR	41	211	1,681	2,435.3	1,661.0
DE	23	205	1,816	3,151.7	2,877.0
Rest	323	684	5,286	3,595.1	2,888.0

Table 7: Top 5 impersonated companies for each ccTLD. Yellow rows are from local companies, blue rows are common companies in the three ccTLDs. Age is median in days.

The Netherlands ccTLD (.nl)

Company	SLDs	Age	CC	Sector
Microsoft	2,319	2,251.0	US	Technology
PayPal	2,134	1,751.0	US	Financial Serv.
ING Netherlands	1,815	1.0	NL	Banking
International Card Services	1,410	2.0	NL	Financial Serv.
Apple	1,276	1,775.0	US	Technology

Ireland ccTLD (.ie)

Microsoft	135	2,598.0	US	Technology
Webmail	60	1,921.0	US	Internet Services
Netflix	46	2,792.0	US	Entertainment
PayPal	24	2,279.0	US	Financial Serv.
DHL	21	2,301.0	DE	Logistics

Belgium ccTLD (.be)

Microsoft	424	2,016.0	US	Technology
Webmail	284	1,368.0	US	Internet Services
Netflix	170	3,039.0	US	Entertainment
PayPal	133	1,829.0	US	Financial Serv.
DHL	118	3,239.0	DE	Logistics

are also impersonated using many SLDs, which shows that local companies are also frequently impersonated, but with new domain names. We include the top 10 companies per ccTLD in Appendix C.

Not every local company is impersonated using new domain names. Figure 5 shows all local (Dutch) companies impersonated using .nl domains, and their number of SLDs and median age. The most popular impersonated local companies (left y axis) are prominently targeted using new domains (right y axis), but some are not. We see the same patterns in .be (Figures in Appendix D).

Consistency over time. We compute the median age per country, per year. We see that the results are consistent over time: local companies are impersonated often with maliciously registered domain names. We show the figures results in Appendix D. We also see that Belgian companies are also impersonated using .nl domains. We discuss this overlap in §4.4.

Table 8: Segments and countries of impersonated companies for malicious registered domains (< 7 days). Rows in yellow refer to segments from local companies.

.nl

Segment	CC	SLDs	Companies	Med. Age (days)
Banking	NL	4,165	15	1.1
Financial Services	NL	1,417	3	1.4
Banking	BE	610	14	2.9
Credit Union	US	335	34	4.7
E-commerce	NL	283	1	2.3
Government Services	NL	178	4	0.8
Telecom	NL	141	9	1.8
Postal Services	NL	109	1	0.9
Domain Registry	NL	93	1	4.5
Online Identity Management	NL	80	1	1.8

.be

Banking	BE	180	11	0.9
Credit Union	US	82	10	0.7
Banking	NL	64	7	1.0
Government Services	BE	29	2	5.7
Insurance	BE	29	1	1.2
Postal Services	BE	25	1	1.3
E-commerce	BE	11	1	0.8
Insurance	FR	9	1	2.1
Government	AR	5	1	0.7
Financial Services	BE	5	1	0.5

4.3 What are new domains used for?

We see that local companies are primarily targeted using newly registered domain names. Next, we turn to investigate what all new domains are used for. In the literature, many research works classify maliciously registered domains (new domains) if they are younger than 30 or 90 days by the time the phishing attack is deployed. We instead look into our datasets, and show what domain ages we see when the phishing attack is detected.

Figure 6 shows the CDF of the SLDs' age for each ccTLD, for all phishing domains. For .nl and .be, we see that there is not much difference between a week and 90 days in the number of SLDs, so instead we choose seven days – a stricter value.

Then, we proceed to group them by segment and country of origin from the impersonated brands. Table 8 shows the results. We see that most of the SLDs are used to impersonate local banks, but also e-commerce, government, and telecommunications (we omit .ie given it has very few SLDs in the first week). We see that Dutch banks rank #3 in phishing in the .be zone, and vice-versa.

4.4 ccTLDs and targeted companies compared

Our datasets revealed 1,233 targeted companies from 86 countries across all ccTLDs (distribution shown in Table 5). We compile the set of brands for each ccTLD and compute their intersections, and show the results as a Venn diagram in Figure 8. Next we explore the characteristics of each of these subsets in the figure.

4.4.1 Common companies among the ccTLDs. From the 1,233 companies we saw in our datasets, only 139 (11% of the total, highlighted text color matching Figure 7 and Figure 8) are found in all ccTLDs. Companies in this subset include Microsoft, PayPal, Apple, Google, and Netflix, which have a global presence and large user base (we list them in Appendix E.

Companies characteristics: What distinguishes them from the others? We find that they are mostly *international/global* (133 from

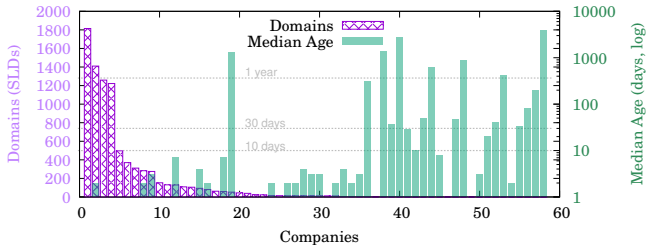


Figure 5: Dutch companies: number of .nl SLDs used in attacks, and median age, sorted by number of SLDs.

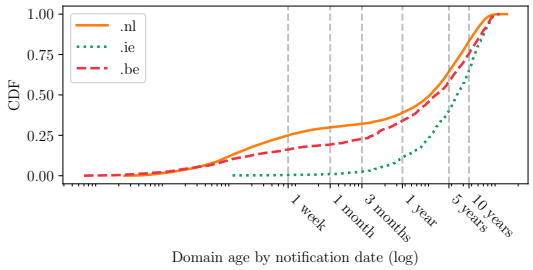


Figure 6: Domain age CDF by notification date.

24 countries, Figure 7), they are *popular* with attackers (they account for 58.6% of all SLDs used in phishing attacks in our datasets, or 21k SLDs out of 37k), and attacks against them are mostly done using *old domains* (having at least 5.4 years median age, or 1,967 days).

Attacker profile: As such, we can infer that attackers who targeted these companies used mostly compromised domain names (given they are old domains), regardless of the ccTLD. For these attackers, it does not matter if the phishing URL is a specific TLD; they leverage vulnerable, exploitable and thus free websites.

The exception to this are two Belgian organizations that have been phished using 376 .nl SLDs, with an median age of four days. These two organizations are the Government of Belgium and an International bank headquartered in Belgium. We speculate that this is also influenced by their sharing of one common language, which allows an attack group to use this language to attack companies in both countries.

4.4.2 ccTLDs exclusively targeted companies. Next we look in the other end of spectrum: what companies are only impersonated in one ccTLD? And why?

Most of the companies in our dataset are only found in the .nl zone – 639 companies. We attribute this partially to the fact that .nl is the largest zone of the three (Table 1), which creates the largest attack surface, and that the .nl dataset covers 10 years while the others cover 4 years.

When we compare the companies impersonated only in the .nl zone (or the 147 companies only in the .be zone), we see the same pattern: most are international, so it follows the same attacker profile as §4.4.1. However, 42 local Dutch companies are impersonated using only .nl domain names, from a total of 60 Dutch companies we saw in all ccTLDs. Local ones are impersonated with newly registered domain names (median age of 4 days). Among these local companies, a popular online marketplace tops the list (Marktplaats [60]), followed by the national mail company.

Attacker profile: we see two attacker profiles. The first one consists of attackers who exploit any domain name they can find – just like in §4.4.1. However, we see an attacker's profile that exploits the trust Internet users have in their ccTLD: they register new domains under the ccTLD of their targeted companies, and get to choose domain names in the process. We analyze these domain names' strings and see they are in Dutch words, which, in turn, raises the bar for the attacker, suggesting these attacks are likely to come from individuals or groups operating locally or in neighboring countries speaking the same language.

.ie *exclusive companies:* We see only 20 companies, 17 being international and 3 local among the .ie exclusive companies. The attacks use only 22 SLDs. Given these small numbers, we think these companies only impersonated with .ie domain names are incidental, using old, compromised domain names. Restricted registration inhibits malicious registered domain names. However, there have been cases of malicious registrations within the .ie zone. In a particular case, 120 domains were maliciously registered using falsified identification documents. The registration occurred on a Friday, and over the weekend, .ie received reports of abuse associated with these domains. Subsequently, these domains were deactivated. This incident indicates that attackers attempt to circumvent restricted registration policies by employing counterfeit documentation.

4.4.3 .nl and .be only companies. We find 247 companies impersonated using .nl and .be exclusively. Out of these, 32 are "local", being either Dutch or Belgian. However, they account for 70% of the SLDs in this subset, with a median age of two days, indicating they are maliciously registered. In fact, most of SLDs used against Dutch and Belgian companies fall into this category, *i.e.,* they are impersonated using *both* ccTLDs.

This finding demonstrates the need for registries to collaborate to mitigate phishing in their own countries. By mitigating .be phishing SLDs, it protects Dutch users, and vice-versa. In this way, the mitigation methods employed in one zone affect the users in the other ccTLD country. These local companies are impersonated by the attacker profile described in §4.4.2, in which they use new domain names crafted to mimic the legitimate website.

We wonder if this is a consequence of having phishing kits tailored for banks in both countries. Given that many banks operate in both countries, and they share the same language, it is relatively easy to phish for clients in both countries using the same domain. We find 82 SLDs in the .nl zone that were used exclusively to phish Dutch and Belgian companies, with a median age of 1.29 days, while we found 2 .be SLDs with a median age of 5.56 days doing the same.

The remaining subsets have mostly compromised domain names, and we include its characteristics in Appendix F. We also investigate time correlation between impersonated brands attacks between the ccTLDs and found a weak correlation (Appendix G).

Subset	Companies	Countries	SLDs	SLDs .nl	SLDs .ie	SLDs .be	Median age (days) .nl	Median age (days) .ie	Median age (days) .be
∪ all	1,233	86	37,215	32,762	728	3,725	1,641	2,746	2,154
∩ all	139	27	21,175	18,040	651	2,484	1,938	2,739	1,921
Intl. companies	133	24	20,090	17,084	636	2,370	1,971	2,748	1,967
Dutch companies	2	1	573	549	6	18	659	3,460	614
Irish companies	2	1	61	31	6	24	3,570	3,535	3,599
Belgian companies	2	1	451	376	3	72	4	1,695	812
.nl only	639	70	3,888	3,888	–	–	1,602	–	–
Intl. companies	597	69	2,808	2,808	–	–	1,756	–	–
Dutch companies	42	1	1,080	1,080	–	–	4	–	–
.ie only	20	12	22	–	22	–	–	2,398	–
Intl. companies	17	11	19	–	19	–	–	2,255	–
Irish companies	3	1	3	–	3	–	–	2,540	–
.be only	147	38	213	–	–	213	–	–	5,671
Intl. companies	135	37	164	–	–	164	–	–	6,340
Belgian companies	12	1	49	–	–	49	–	–	108
(.nl ∩ .be) - ∩ all	247	51	11,611	10,603	–	1,008	1,394	–	1,969
Intl. companies	215	49	3,673	3,064	–	609	1,585	–	2,490
Dutch companies	13	1	7,151	7,061	–	90	1	–	2
Belgian companies	19	1	787	478	–	309	2	–	2

Figure 7: Subsets characteristics. SLDs refers to non-unique SLDs, given a same SLD can be used for multiple companies.

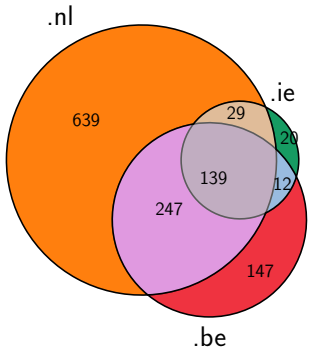


Figure 8: Venn diagram of impersonated companies.

5 Mitigation

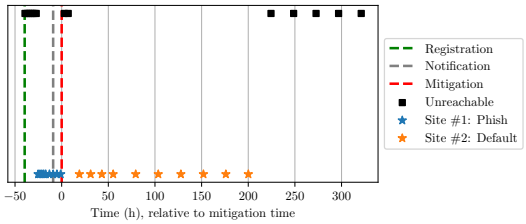
5.1 Mitigation in practice

Phishing mitigation is not a single event carried out by a single actor. In practice, we see actions taken by registrars, registries, hosting providers, webmasters, all happening independently to mitigate phishing attacks. We demonstrate this in two examples (Figure 9), drawn from two phishing attacks observed in the .nl zone. We include their screenshot captured independently by our crawler in Appendix H.

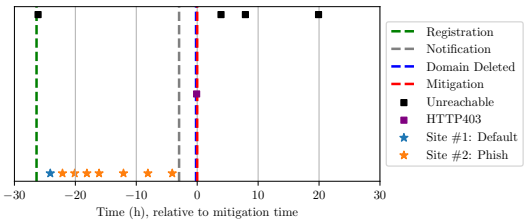
French Bank Phishing: In Figure 9a, we show the timeline of a phishing domain used to impersonate a French bank, which was mitigated only at the web level and left to expire at the DNS level. This figure is made with a combination of various datasets. From Netcraft, we extract the phishing notification and mitigation time, and plot them as dashed lines, relative to the mitigation time. From the .nl registration database, we retrieve the domain registration time, and plot it as a green dashed line (t=-39h).

The remaining data points are web measurements obtained from our in-house crawler, which crawls every domain name multiple times after registration (§3.2). Up to t=-27h, our crawler cannot find a website, so we mark it as unreachable. Then, between -25h and -1h, our crawler fetches the phishing site: the .nl domain redirects the crawler to a .click domain, where the actual phishing is hosted. We know this is the phishing site given it matches the data reported by Netcraft. This redirect takes place at web level, using HTTP redirects [25].

Netcraft detects the phishing site and notifies at t=-7h (so it took seven hours to mitigate it). After the mitigation time reported by Netcraft, our crawler data confirms the mitigation (it cannot retrieve the phishing website any longer). Then, around t=10h, the website becomes reachable again, but instead of the phishing site, we now see a default landing page of the hosting provider. Lastly, after t=200h, the website becomes unreachable again, and our crawler does not crawl it anymore. This domain was then left by the registrar to expire, meaning it was not deleted before the end of its 1-year lease.



(a) Web level phishing mitigation of a French bank

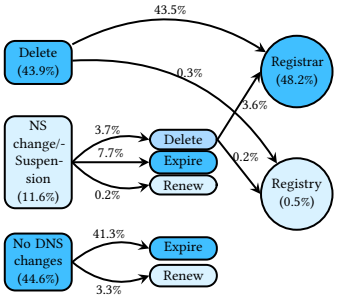


(b) DNS level mitigation of Tinder phishing

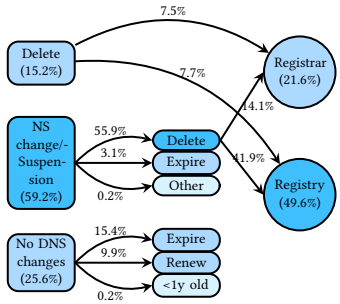
Figure 9: Phishing mitigation examples.

Tinder Phishing: Figure 9b shows the timeline of a domain used to impersonate the dating site Tinder, which was mitigated at both web and DNS levels. We make the graphs following the same steps as in Figure 9a. This website was detected by .nl analysts (and not Netcraft) and inserted into Netcraft to use their notification services. The website was then cleaned up (see HTTP 403 error) and then deleted from the .nl zone by the registrar.

Comparing both cases, we see that French bank phishing was mitigated by the hosting provider or webmaster and the domain name was left to expire, whereas the Tinder phishing was first mitigated at the web level and then at the DNS level by the registrar. These two examples demonstrate the different ways and actors that can mitigate phishing sites.



(a) .nl (6,218 domains)



(b) .be (454 domains)

Figure 10: DNS mitigation: new domains.

5.2 Phishing mitigation at the DNS level

Next, we focus on DNS-level mitigation, by analyzing the registries' registration databases. Before we analyze the results, we divide the datasets into two: new domains and old domains. The rationale is that new domains tend to be maliciously registered, while old domains tend to be compromised websites. As discussed in §2.5, they require different mitigation methods. We choose a threshold of 7 days to classify a domain as maliciously registered, following the analysis shown in §4.3. Given our datasets span over many years, some SLDs can be registered several times in the period, so we consider each new registration independently.

5.2.1 New domains mitigation. Figure 10 summarizes the DNS mitigation results, for new domain names (< 7 days), for .nl and .be (we omit .ie due to its low number of maliciously registered domains). We show the chain of events that lead to mitigation at the DNS level.

Policy impact: who acts first? Recall the ccTLDs abuse handling policies (§2.6); .nl first allows other parties to mitigate, while .be often takes action first. We see clearly in Figure 10 how these policies reflect in the results: roughly 50% of the .nl malicious registered domains are deleted by the *registrars*, and less than 1% by the registry. For .be, we see 71% of deletes – the registry being responsible for 50% and the registrars for another 21%. We see that most of the malicious .be domains are first suspended (removed from the zone) or have their DNS records changed, which conforms to the policy of .be.

Disposable domains: we see that the majority of the new domains are either deleted or simply not renewed (expired). These account for 96.5% for .nl, and 89.9% for .be, thus we can conclude that they are disposable – attackers likely make enough profits within the period that the phishing website is active, and have no interest in renewing it.

Mitigation speed: once a phishing website is active, speed is of the essence to prevent successful attacks. We show in Figure 11 the time it takes to mitigate a phishing attack after its notification by Netcraft. For .be, we see that 60% of the .be domains are suspended or have their NS records changed in the first day, compared to 40% for .nl (Figure 11a). Out of the 269 .be domains, 239 (88.8%) were suspended by the .be registry itself. For .nl, however, 720 domains (out of 6218, 11.5%) had NS changed, all by their registrars.

Domain deletions take longer. We see in Figure 11b how fast domains are deleted for each ccTLD. For .be, we see a significant increase at 15 days – this conforms to their policy of waiting two weeks before proceeding with a delete. For .nl, we see spikes around 30, 60 days, and one year. Given that virtually all deletes are performed by the registrar, these are a consequence of their policies of deleting domains after certain intervals. As such, for both registries, we see that deleting is often resorted to as a second step in the mitigation strategy.

5.2.2 Old domain names mitigation. The majority of old SLDs are renewed post mitigation, for the three ccTLDs (59% for .nl, 94% for .ie, and 79.3% for .be). We believe this is because old domains tend to be *compromised* and then cleaned up at the web level, while we do not see changes at the DNS level.

This matches our intuition that domain names of compromised websites are valuable and thus not disposable. Their registrants have an interest in keeping these domain names active, so they are renewed, after removing the harmful content from the website.

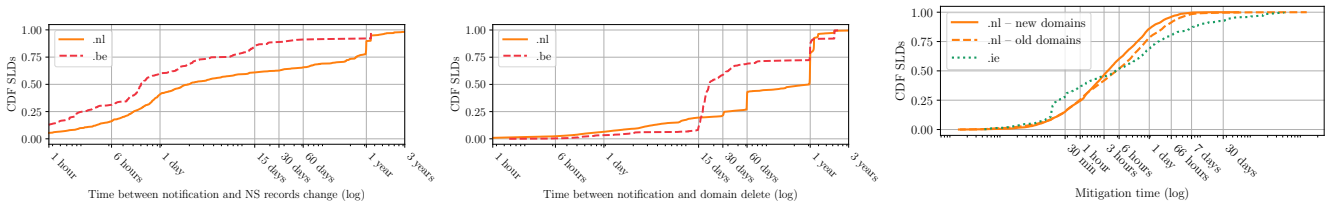
Interestingly, however, we see that 27% of old domains from .nl are deleted. We investigated this and we found it was due to one registrar deleting domains on the 12th day after notification – due to its own policy. So these domains could also be labeled as maliciously registered, which would increase the renew ratio for .nl. We include the CDFs of domain name delete and NS record changes in Appendix D, as well as the mitigation figures.

5.3 Phishing mitigation seen from crawlers

Next we analyze data obtained from web crawlers that visit phishing websites. These crawlers mimic user behavior when attempting to access a phishing URL, though some sites may use cloaking techniques to mislead crawlers [97].

We first start with the results reported by Netcraft. For each SLD, we compute its mitigation time, which we define as the time between the Netcraft notification to resolution *i.e.*, when the phishing attack was mitigated – either at DNS and/or web level – so it cannot be reached by a browser. We have this data only for .nl and .ie, given these ccTLDs subscribed to their notification services (§2.7), and, as such, have the phishing URLs tracked by Netcraft.

Figure 11c shows the results. We see that for .nl, both new and old domains are mitigated at comparable speeds, even if old domains account for 75% of all .nl compromised domains (Table 2). .ie, in turn, has faster mitigation times – 40% mitigation within the first hour, while .nl reaches only 20%. After 24h, 80% of the .nl domains are mitigated while roughly 70% of the .ie are mitigated.



(a) DNS mitigation: Domain suspension **(b) DNS mitigation: Domain delete** **(c) Mitigation reported by Netcraft**

Figure 11: Phishing mitigation times.

We speculate that this is due to the different environments both operate in – different registrars, hosting providers, and so forth – and the number of domain names (`.nl` sees roughly 10× more phishing domains than `.ie`).

When compared to DNS level mitigations seen at the registry (Figure 11a and Figure 11b), we see that the web mitigation times for `.nl` are significantly shorter than DNS mitigation. Specifically, 40% of the maliciously registered `.nl` SLDs have their NS records changed on the first day, while web mitigation data shows 80% mitigation on the first day post-notification.

Why such difference? Results from Figure 11c are a compound of *all* sorts of mitigations at all levels (§2.5) by all intermediaries: DNS and Web mitigations by registries, registrars, hosting providers, DNS providers, and webmasters. As such, it reflects a more realistic view of mitigation.

Is this mitigation fast enough? The mitigation shown in Figure 11c shows time intervals after notification – however, as shown in Figure 9, phishing sites can go unnoticed for hours. This interval – detection time – combined with mitigation time – may be enough time for phishing attacks to be profitable. Therefore, reducing both intervals should be the goal for registries wishing to reduce phishing attacks in their zones.

5.3.1 Can we trust mitigation times from Netcraft? The mitigation times reported by Netcraft are used by decision makers within the registry to guide policies and strategies. Given that they are self-reported using vantage-points inaccessible to us, we cannot validate them, especially considering they span over 10 years.

We can, however, validate parts of the reports, using the same methodology we described to build the timelines of phishing attacks in §5.1. Not all cases adhere to the patterns shown. Appendix I shows cases involving undetectable phishing web cloaking, domain deletes, and undetected instances.

We analyze all phishing SLDs from 2022, totaling 1,925 reported by Netcraft to the `.nl` zone. We focus on domains no older than 7 days at notification time, yielding 322 domains. We have data points for 234 SLDs before and after Netcraft's reported mitigation time. The missing data is due to individual crawl runs that did not start.

We examine web mitigation for 234 phishing sites in the `.nl` ccTLD. Each web measurement point from our crawler is classified as `HTTP 200` (active site), Web mitigation (HTTP [3-4-5]*, indicating `HTTP 300`, `400`, and `500` errors [25]), or unreachable. If a domain has multiple statuses post-mitigation, we consider its last status.

Figure 12 presents a Sankey diagram of the 234 new SLDs pre and post-mitigation. Most domains (195) had a website, but the majority (119) became unreachable post-mitigation. In total, 140 SLDs are unreachable post-mitigation. Another 61 SLDs that had an active website post-mitigation returned another website, mostly landing or default pages.

We observe evidence of mitigation before Netcraft's timestamp (Unreachable and HTTP [3-4-5]* in Figure 12). We found no evidence of phishing websites being served post-mitigation. Figure 13 shows the age of each of the 234 SLDs at mitigation time and when it was deleted from the `.nl` zone. We see a concentration of domains being deleted on the mitigation day ($x = y$), and at one year – when they expire. The majority of these domains remain active even after their web mitigation.

This subset of SLDs thus corroborates Netcraft's reported mitigation times – our crawler, run from independent vantage points, did not find any phishing web sites *after* Netcraft's self-reported mitigation time for these 234 SLDs. As such, we have no evidence to not trust the mitigation times reported by Netcraft shown in Figure 11c.

5.3.2 Can we reduce detection times? Reducing phishing detection and mitigation time is key to disincentivize attacks in a TLD. Next we perform a post-mortem analysis of our crawling data to determine if phishing websites could be earlier identified, as shown in Figure 9, in which our crawler had found the phishing site *before* the notification. We cannot do this at scale retroactively, only for the domains we already had data on. We start with the 234 domains we discussed in §5.3.1 – which are all new domains. Out of those, only 36 had a phishing website at their home page – thus visible to our independent crawler (many phishing attacks hide their phishing sites in subdomains or subdirectories).

We then manually analyze their HTML title and HTML descriptions fields, and screenshots, and see if there are indications of the phishing site, and when. For example, if a phishing site had been reported impersonating Microsoft, we look at HTML title and description before and after Netcraft's notification, and screenshots, and see when the phishing site was first seen from our vantage point, if it was seen. We perform this analysis manually. From the 36 sites, we saw 18 with phishing sites either from the HTML title or description or from the screenshots.

Figure 14 shows the results. Even though it is a small set of domains, we see that 16 out of the 18 could have been possibly detected using data available at the registry, given they were able

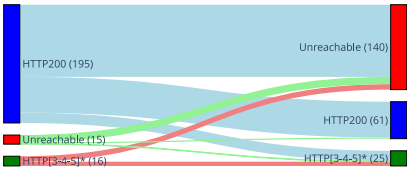


Figure 12: Sankey diagram of 234 SLDs using active web measurements from our crawler.

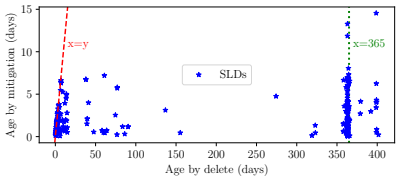


Figure 13: SLDs: age by mitigation and domain delete at .nl for new domains from 2022.

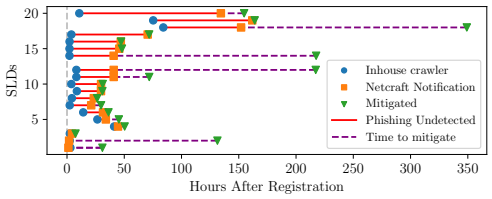


Figure 14: Phishing detection, notification, and mitigation.

Table 9: Phishing attack strategies

	Maliciously Registered	Compromised Domains
Bio. Analogy	Aggressive mimicry	Aggressive mimicry & parasitism
Share SLDs	20.00%	80.00%
Leverage ccTLD Trust	Yes	Likely no
Use	Local company or operating locally	Any, mostly international
Targeted Companies	< 5%	>95%
Restricted Reg.	Mitigated	Not Mitigated
Detection	Easier	More Difficult
Mitigation	DNS, Web	Mostly Web

to retrieve the phishing website before the Netcraft notification. For the other two, Netcraft notified before our crawl runs.

Although these few domains may not be representative for the entire dataset, it shows that we could reduce detection time for some newly registered domain names that host phishing in their home pages. We intend to investigate this as future work.

6 Discussion

Our analysis unveils two classes of attacks (Table 9), namely, those who prefer a do-it-yourself tactic, by registering and hosting their own domains and sites, and others that prefer to misuse someone else's hosting, by compromising their websites; we refer to the latter as "parasitism" tactic. Roughly 80% of phishing SLDs we observe are compromised domains.

We can only speculate why these strategies exist. Maliciously registered domain names enable better mimicry, as attackers can pick more convincing domains. For instance, instead of compromising a domain like `joesbbqhouston.net` for bank phishing, they can use a crafted domain such as `citibank-validate-card.com`. Additionally, choosing the same top-level domain (TLD) as the target company enhances overall mimicry effectiveness. However, it has financial costs. Compromised domains, in turn, are *cheaper* (as they require no domain registration, hosting, or DNS setup). They also make attack *attribution more difficult* since few traces are left beyond exploiting the compromised website. Additionally, leveraging an existing domain's benign reputation and achieving scalability without costs are advantages (previous research has demonstrated that websites are frequently scanned for vulnerabilities of CMSes [96]).

Generalization of our findings: Our findings are based on an analysis of three ccTLDs. Although we did not evaluate other TLDs, we believe that similar behaviors may be prevalent, particularly concerning compromised domains used for phishing. We would not be surprised if our observations also hold for registries with similar registration policies to those evaluated in this paper. As future work, we intend to collaborate with other TLDs to determine if the same behaviors can be found.

6.1 Call for action

We present the following action points based on our findings:

More research on compromised domains is needed. Most research works focus on maliciously registered domain names (*e.g.,* [29]). We need new solutions to detect compromised domains, given they form the bulk of SLDs. It is a challenge, given we cannot use registration or certificate features in the process, such as domain name or hosting infrastructure. Given they are parasites to legitimate websites, they leverage the reputation of their host. At the registry, we could use authoritative server traffic to detect suspicious domains (we demonstrate an example in Figure 15c in Appendix A)). However, it is challenging to validate the results: we cannot know from DNS queries the full URL of phishing attacks. Collaboration with hosting providers may be used in such cases in an attempt to validate results. In the meantime, blocklists, reactive by nature, will remain essential in detection of compromised domains.

Increase collaboration between ccTLDs to mitigate phishing: The results presented in §4.4 reveal that attackers frequently utilize many maliciously registered domains from one ccTLD to target companies in neighboring countries. In this context, mitigating domains within the `.be` zone not only protects users in Belgium but also benefits users in The Netherlands, and vice versa. Therefore, fostering collaboration between ccTLD registries is essential to reduce the uptime of phishing sites and improve user protection against scams. Note, however, that mitigating phishing at one TLD may drive attackers to another one (*e.g.,* [59]). Overall, the goal is to make it more expensive or difficult for them to succeed.

Revising registration and abuse handling policies in ccTLDs: This study represents the first comprehensive examination and comparison of abuse handling policies and registration policies across ccTLDs and their impact on phishing mitigation. Our research demonstrates that while a restricted registration policy effectively combats maliciously registered domain names, it falls short in addressing compromised domain names. Additionally, we show that

.be adopts a more active approach to mitigation compared to .nl in terms of suspending domain names, whereas .nl waits intermediaries to first mitigate, and takes action after 66h.

Mitigation efforts are intricate and involve multiple stakeholders working independently and concurrently (§5). These findings are currently under discussion within the three ccTLDs, spanning legal, management, and abuse handling departments. Consequently, policy adjustments may be implemented to enhance abuse mitigation. Furthermore, our analysis in §5.3.2 highlights the potential for ccTLD registries to reduce phishing uptime by leveraging their own data.

6.2 Limitations

Our study has the following limitations. First, we focus only on phishing seen at three ccTLDs. While many studies have compared and analyzed abuse among multiple TLDs, our focus is on three ccTLDs given we have access to the entire activity of ccTLDs over years and a high quality blocklist and private datasets, providing us with a complete view within these three zones.

We intended to include more ccTLDs in our study; however, obtaining such datasets proved challenging. Many TLD registries lack dedicated research departments or personnel focused on research activities, thereby making it difficult to allocate time for such analyses. The absence of previous research papers combining these datasets underscores this challenge. We intend to share our code with other TLDs and coordinate future research efforts, with the aim of expanding our collaborative network.

Our analysis uses the number of SLDs as the primary metric (as do the majority of studies). However, the ideal metric for assessing the severity of individual phishing attacks would involve evaluating the specific financial losses incurred by each attack. Unfortunately, this information is not accessible to us. In our attempt to gather insights, we contacted representatives from two banks. Regrettably, they informed us that losses resulting from individual phishing attacks are intentionally compartmentalized within their institutions and cannot be disclosed – and they refused even to elaborate on why. Despite this limitation, the recurring nature of phishing attacks shows that phishing remains profitable.

7 Related work

Phishing analysis. There are two primary works closely related to ours. Oest *et al.* [71] focus on analyzing phishing attacks against a large payment provider, spaning over a year, using web server logs. Differently from them, we do not have access to web server logs. We also do not limit ourselves to a specific target organization; we analyze all phishing we see in three ccTLDs, and cover 28k SLDs spanning from four to 10 years.

Bijmans *et al.* [12] focus on phishing attacks against banks based in the Netherlands. They use certificate transparency logs [15] in their detection process, therefore focusing on new domain names. They were able to find 1.3k suspicious SLDs, in a 4 month period. We in turn, focus on any segment of phishing – not only banking – and rely upon a phishing blocklist provider instead spanning 4 to 10 years. We report 28.7k SLDs, 80% being compromised domain names. While they report many TLDs, we cover only three. They find that most phishing on Dutch banks is from .info domain names, which are cheaper than .nl domains.

Malicious domain names detection. Hao *et al.* [28–30] delve into the domain registration process of spam domains, with a focus on identifying features hinting at potential malicious intent during registration, before any attack occurs. Their investigation encompasses a thorough analysis of registrar characteristics, lifecycle, registration surges, and naming patterns.

Malicious domains vs comprised domains. Distinguishing compromised domains from maliciously registered has been covered in many works [9, 18, 55, 61]. These studies are primarily focused on the categorization process and do not extend their analysis to include abuse handling procedures, the impersonation of companies, and lifespan of phishing websites. Our research addresses this gap.

Short-lived domain names. Two studies have shown how domains involved in malicious activities tend to be *deleted* from the DNS zone before their expiration date [3, 8], suggesting this is a consequence of the takedown efforts. We see similar patterns for .be, where 68.2% of maliciously registered SLDs are deleted, while .nl sees half of its domains being deleted. Their research, however, does not delve into the takedown policies of TLD registries or compromised domain names.

Mitigation. One study interviewed 24 experts in taking down domain names and concluded that law enforcement agencies are not very effective in taking websites down compared with the specialist companies [33]. Another study by Moore and Clayton addressed notice and take-down of various types of cybercrime – including phishing attacks, and their impact in the incentive structures involved [64]. Moore and Clayton also measured phishing sites lifetimes to effectiveness of takedown strategies employed by targeted institutions [63].

Phishing lifetimes at TLD level were analyzed by Korczyński et al. [48]. However, the study does not delve deeper into whether these domains were taken down at the DNS or hosting level, nor does it analyze the takedown strategies employed by TLD registries, or hosting providers. The role of hosting providers and web masters in preventing domain names from being compromised and misused for phishing has been addressed as well [90]. Lastly, Alowaisheq et al. undertook a systematic study to offer a detailed perspective on the domain takedown process at the DNS level [4]. They examine the practice of mitigating across various types of organizations, including registries and registrars. In this paper, we approach phishing from a distinctly different perspective—examining the takedown policies of three ccTLD registries concerning compromised and maliciously registered domain names.

Policy change: A previous study has shown the impact of spamming domains when the China's .cn registry restricted its domains registration and increase prices at the same time [59]. It caused spamming domains to move to other TLDs.

8 Conclusions

Phishing on the Web continues to be one of the most devastating cyber threats, contributing to hundreds of millions of US Dollars in losses for individuals and companies. In this paper, for the first time, we perform a longitudinal study in collaboration with three

ccTLDs that spans up to 10 years and considers more than 8 million domains to characterize phishing tactics in recent years. Our analysis of the entire namespace of three ccTLDs shows that cybercriminals use two phishing tactics: compromising pre-existing legit domains and maliciously registering new ones. Further investigation unveils that more than a vast majority of these phishing domains are old domains, likely compromised websites (80% of total) We also show that compromised domains are used primarily to impersonate international companies at a meager cost by parasitizing well-established and legitimate domains, while most of the new maliciously registered domains are used to impersonate local companies. Our findings challenge current best practices that focus on scrutinizing newly registered domains or restricted registration policies by ccTLDs, as they can only be effective for a relatively small fraction of phishing activity (less than 20% based on our study). Fighting against compromised domains is much more challenging and is a call for action for closer collaboration with hosters, registrars, and ccTLDs. Our results have already been taken into consideration in the revision of policies by ccTLDs that we collaborate with. As part of our future agenda, we plan to extend our collaboration with additional ccTLDs to characterize the evolving phishing activity better and enhance the mitigation of phishing on the Web.

Acknowledgments

The authors thank the anonymous CCS reviewers and the paper shepherd for their reviews and insightful comments.

Thomas Daniels received funding from VLAIO (Flemish Innovation & Entrepreneurship) through the Baekeland PhD mandate (HBC.2023.0718). Georgios Smaragdakis was supported by the European Commission under the Horizon Europe Programme as part of the project SafeHorizon (Grant Agreement no. 101168562). The content of this article does not reflect the official opinion of the European Union. Responsibility for the information and views expressed therein lies entirely with the authors.

References

[1] Greg Aaron, Lyman Chapin, David Piscitello, and Colin Strutt. 2021. Phishing Landscape 2021: An Annual Study of the Scope and Distribution of Phishing. https://interisle.net/s/PhishingLandscape2021.pdf. Accessed: June 2, 2024.

[2] Greg Aaron and Rod Rasmussen. 2017. APWG Global Phishing Survey: Trends and Domain Name Use in 2016. https://docs.apwg.org//reports/APWG_Global_Phishing_Report_2015-2016.pdf.

[3] Antonia Affinito, Raffaele Sommese, Gautam Akiwate, Stefan Savage, KC Claffy, Geoffrey M Voelker, Alessio Botta, and Mattijs Jonker. 2022. Domain Name Lifetimes: Baseline and Threats. In *6th Network Traffic Measurement and Analysis Conference, TMA 2022, Enschede, The Netherlands*. International Federation for Information Processing (IFIP).

[4] Eihal Alowaisheq, Peng Wang, Sumayah A Alrwais, Xiaojing Liao, XiaoFeng Wang, Tasneem Alowaisheq, Xianghang Mi, Siyuan Tang, and Baojun Liu. 2019. Cracking the Wall of Confinement: Understanding and Analyzing Malicious Domain Take-downs.. In *The Network and Distributed System Security Symposium (NDSS)*. San Diego, CA, United States of America.

[5] Anti-Phishing Working Group. 2023. eCrime eXchange (eCX). https://github.com/APWG/ecx.

[6] APWG. 2023. *APWG Trends Report Q2 2023.* https://docs.apwg.org/reports/apwg_trends_report_q2_2023.pdf Accessed: Jan. 24, 2024.

[7] Banken.nl. 2024. *Marktaandeel | Banken.nl (In Dutch).* https://www.banken.nl/bankensector/marktaandeel Accessed: 23-Jan-2024.

[8] Timothy Barron, Najmeh Miramirkhani, and Nick Nikiforakis. 2019. 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)*. USENIX Association, 383–397.

[9] Jan Bayer, Ben Chukwuemeka Benjamin, Sourena Maroofi, Thymen Wabeke, Cristian Hesselman, Andrzej Duda, and Maciej Korczyński. 2023. Operational Domain Name Classification: From Automatic Ground Truth Generation to Adaptation to Missing Values. In *Passive and Active Measurement*. Cham, 564–591.

[10] DNS Belgium. 2024. *.be Registrant verification now also with Machine Learning.* https://www.dnsbelgium.be/en/news/registrant-verification-with-machine-learning (Accessed on: Apr 26 2024).

[11] DNS Belgium. 2024. *DNS Belgium - Managing .be, .vlaanderen, and .brussels Domain Names.* https://www.dnsbelgium.be/en/ Accessed: 2024-08-29.

[12] Hugo Bijmans, Tim Booij, Anneke Schwedersky, Aria Nedgabat, and Rolf van Wegberg. 2021. Catching Phishers By Their Bait: Investigating the Dutch Phishing Landscape through Phishing Kit Detection. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, 3757–3774.

[13] Frank Breitinger, F Breitinger, D White, B Guttman, M McCarrin, and V Roussev. 2014. *Approximate matching: definition and terminology.* US Department of Commerce, National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-168.pdf

[14] CENTR. 2024. *Domain Registration Statistics.* https://stats.centr.org/public/registrations Accessed: 2024-07-18.

[15] Certificate Transparency. 2024. Working together to detect maliciously or mistakenly issued certificates. https://certificate.transparency.dev/ Accessed: 2024-01-21.

[16] Igino Corona, Battista Biggio, Matteo Contini, Luca Piras, Roberto Corda, Mauro Mereu, Guido Mureddu, Davide Ariu, and Fabio Roli. 2017. DeltaPhish: Detecting Phishing Webpages in Compromised Websites. In *Computer Security – ESORICS 2017*, Simon N. Foley, Dieter Gollmann, and Einar Snekkenes (Eds.). Springer International Publishing, Cham, 370–388.

[17] L. Daigle. 2004. *WHOIS Protocol Specification.* RFC 3912. IETF. http://tools.ietf.org/rfc/rfc3912.txt

[18] Ravindu De Silva, Mohamed Nabeel, Charith Elvitigala, Issa Khalil, Ting Yu, and Chamath Keppitiyagama. 2021. Compromised or Attacker-Owned: A Large Scale Classification and Study of Hosting Domains of Malicious URLs. In *Proc. of USENIX Security*. 3721–3738.

[19] DENIC eG. 2024. *2022 DENIC Domain Statistics.* https://www.denic.de/en/whats-new/news/article/2022-denic-domain-statistics Accessed: 2024-07-18.

[20] Domain Name Wire. 2024. *Verisign Announces .com Price Hike to $10.26.* https://domainnamewire.com/2024/02/08/verisign-announces-com-price-hike-to-10-26/ Accessed: 2024-07-22.

[21] DomainTools. 2024. *Domain Count Statistics for TLDs - DomainTools.* https://research.domaintools.com/statistics/tld-counts/ Accessed: 2024-07-18.

[22] European Union Agency for Cybersecurity. 2023. ENISA Threat Landscape 2023. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023.

[23] European Union Agency for Cybersecurity. 2024. Malware, Phishing, and Ransomware. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2023.

[24] Europol. 2024. *International investigation disrupts phishing-as-a-service platform LabHost.* https://www.europol.europa.eu/media-press/newsroom/news/international-investigation-disrupts-phishing-service-platform-labhost Accessed: Apr 19, 2024.

[25] R. Fielding and J. Reschke. 2014. *Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content.* RFC 7231. IETF. http://tools.ietf.org/rfc/rfc7231.txt

[26] Google. 2023. Google Safe Browsing. https://safebrowsing.google.com/.

[27] Xiao Han, Nizar Kheir, and Davide Balzarotti. 2016. PhishEye: Live Monitoring of Sandboxed Phishing Kits. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security* (Vienna, Austria) *(CCS '16)*. Association for Computing Machinery, New York, NY, USA, 1402–1413.

[28] Shuang Hao, Nick Feamster, and Ramakant Pandrangi. 2011. Monitoring the initial DNS behavior of malicious domains. In *Proceedings of Internet Measurement Conference*. Association for Computing Machinery.

[29] Shuang Hao, Alex Kantchelian, Brad Miller, Vern Paxson, and Nick Feamster. 2016. 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* (Vienna, Austria) *(CCS '16)*. Association for Computing Machinery, New York, NY, USA, 1568–1579.

[30] Shuang Hao, Matthew Thomas, Vern Paxson, Nick Feamster, Christian Kreibich, Chris Grier, and Scott Hollenbeck. 2013. Understanding the domain registration behavior of spammers. In *Proceedings of the 2013 Conference on Internet Measurement Conference*. Association for Computing Machinery, 63–76.

[31] P. Hoffman, A. Sullivan, and K. Fujiwara. 2019. *DNS Terminology.* RFC 8499. IETF. http://tools.ietf.org/rfc/rfc8499.txt

[32] Joel Hruska. 2008. *Spam sees big nosedive as rogue ISP McColo knocked offline.* https://arstechnica.com/information-technology/2008/11/spam-sees-big-nosedive-as-rogue-isp-mccolo-knocked-offline/ Accessed: Jan 31, 2024.

[33] Alice Hutchings, Richard Clayton, and Ross J. Anderson. 2016. Taking down websites to prevent crime. In *2016 APWG Symposium on Electronic Crime Research*. IEEE, 102–111.

[34] ICANN. 2013. *2013 Registrar Accreditation Agreement.* https://www.icann.org/resources/pages/approved-with-specs-2013-09-17-en Accessed: Jan 30, 2024.

[35] ICANN. 2023. *2023 Global Amendments to the Base gTLD Registry Agreement (RA), Specification 13, and 2013 Registrar Accreditation Agreement (RAA).* https://www.icann.org/en/system/files/files/proposed-base-gtld-ra-global-amendment-11sep23-en.pdf Accessed: Jan 30, 2024.
[36] ICANN. 2024. *About gTLD Compliance Program.* https://www.icann.org/resources/pages/compliance-2012-02-25-en [Online; accessed 30-January-2024].
[37] ICANN. 2024. *ccTLDs.* https://www.icann.org/resources/pages/cctlds/cctlds-en [Online; accessed 30-January-2024].
[38] ICANN. 2024. *ICANN's Contracted Parties Approve New Obligations to Mitigate DNS Abuse.* https://www.icann.org/en/blogs/details/icanns-contracted-parties-approve-new-obligations-to-mitigate-dns-abuse-13-12-2023-en [Online; accessed 30-January-2024].
[39] IE Domain Registry. 2022. *Dispute Resolution.* https://www.weare.ie/dispute-resolution/ Accessed: Apr 4, 2024.
[40] IE Domain Registry. 2023. How to register a domain. https://www.weare.ie/how-to-register-a-domain/.
[41] IE Domain Registry. 2023. *Registration and Naming Policy.* https://www.weare.ie/wp-content/uploads/2023/12/Registration-and-Naming-2023.pdf Accessed: Apr 4, 2024.
[42] IE Domain Registry. 2024. *We Are IE - The Official Registry for .ie Domain Names.* https://www.weare.ie Accessed: 2024-08-29.
[43] Internet Corporation for Assigned Names and Numbers (ICANN). 2024. . https://www.icann.org/ Accessed: Feb 15, 2024.
[44] Interpol. 2024. Cybercrimes cross borders and evolve rapidly. https://www.interpol.int/en/Crimes/Cybercrime.
[45] Japanese National Police Agency. 2022. Threats in Cyberspace in 2022. https://www.npa.go.jp/english/bureau/cyber/document/threatsincyberspace2022.pdf.
[46] Joint Cybersecurity Advisory (CSA). 2022. 2022 Top Routinely Exploited Vulnerabilities. https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-215a.
[47] Kevin Lane Keller. 2019. Consumer Research Insights on Brands and Branding: A JCR Curation. *Journal of Consumer Research* 46, 5 (11 2019), 995–1001. https://doi.org/10.1093/jcr/ucz058 arXiv:https://academic.oup.com/jcr/article-pdf/46/5/995/50265152/jcr_46_5_995.pdf
[48] Maciej Korczyński, Samaneh Tajalizadehkhoob, Arman Noroozian, Maarten Wullink, Cristian Hesselman, and Michel van Eeten. 2017. Reputation Metrics Design to Improve Intermediary Incentives for Security of TLDs. In *IEEE EuroS&P*. 579–594.
[49] Maciej Korczyński, Maarten Wullink, Samaneh Tajalizadehkhoob, Giovane C.M. Moura, Arman Noroozian, Drew Bagley, and Cristian Hesselman. 2018. Cybercrime After the Sunrise: A Statistical Analysis of DNS Abuse in New gTLDs. In *ACM ASIACCS*.
[50] Brian Krebs. 2008. Host of Internet Spam Groups Is Cut Off. https://www.washingtonpost.com/wp-dyn/content/article/2008/11/12/AR2008111200658.html Accessed on: April 17, 2024.
[51] M. Kucherawy and E. Zwicky. 2015. *Domain-based Message Authentication, Reporting, and Conformance (DMARC).* RFC 7489. IETF. http://tools.ietf.org/rfc/rfc7489.txt
[52] KvK. 2023. *Netherlands Chamber of Commerce KVK.* https://www.kvk.nl/en/ Accessed: Jan. 24, 2024.
[53] SIDN Labs. 2024. *Abuse Statistics.* https://stats.sidnlabs.nl/en/abuse.html Accessed: 2024-07-19.
[54] SIDN Labs. 2024. *Assessing the Risk of New .nl Registrations Using RegCheck.* https://www.sidnlabs.nl/en/news-and-blogs/assessing-the-risk-of-new-nl-registrations-using-regcheck Accessed: Jan 30, 2024.
[55] Sophie Le Page, Guy-Vincent Jourdan, Gregor V. Bochmann, Iosif-Viorel Onut, and Jason Flood. 2019. Domain Classifier: Compromised Machines Versus Malicious Registrations. In *Web Engineering*, Maxim Bakaev, Flavius Frasincar, and In-Young Ko (Eds.). Springer International Publishing, Cham, 265–279.
[56] J. Levine. 2010. *DNS Blacklists and Whitelists.* RFC 5782. IETF. http://tools.ietf.org/rfc/rfc5782.txt
[57] Vector Guo Li, Gautam Akiwate, Kirill Levchenko, Geoffrey M Voelker, and Stefan Savage. 2021. Clairvoyance: Inferring blocklist use on the Internet. In *Passive and Active Measurement: 22nd International Conference, PAM 2021, Virtual Event, March 29–April 1, 2021, Proceedings 22*. Springer, 57–75.
[58] Daiping Liu, Zhou Li, Kun Du, Haining Wang, Baojun Liu, and Haixin Duan. 2017. Don't Let One Rotten Apple Spoil the Whole Barrel: Towards Automated Detection of Shadowed Domains. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (Dallas, Texas, USA) *(CCS '17)*. Association for Computing Machinery, New York, NY, USA, 537–552.
[59] He Lonnie Liu, Kirill Levchenko, Márk Félegyházi, Christian Kreibich, Gregor Maier, and Geoffrey M Voelker. 2011. On the effects of registrar-level intervention. In *4th USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET 11)*. USENIX Association, Boston, MA, 1–8.
[60] Marktplaats. 2024. *Marktplaats - De plek om nieuwe en tweedehands spullen te kopen en verkopen (In Dutch).* https://www.marktplaats.nl/ Accessed: Jan. 24, 2024.
[61] Sourena Maroofi, Maciej Korczyński, Cristian Hesselman, Benoît Ampeau, and Andrzej Duda. 2020. COMAR: Classification of Compromised versus Maliciously Registered Domains. In *2020 IEEE European Symposium on Security and Privacy (EuroS P)*. 607–623.
[62] P. Mockapetris. 1987. *Domain names - concepts and facilities.* RFC 1034. IETF. http://tools.ietf.org/rfc/rfc1034.txt
[63] Tyler Moore and Richard Clayton. 2007. Examining the Impact of Website Take-down on Phishing. In *Anti-phishing Working Groups 2nd Annual eCrime Researchers Summit*. ACM, 1–13.
[64] Tyler Moore and Richard Clayton. 2009. *The Impact of Incentives on Notice and Take-down.* Springer, 199–223.
[65] Giovane C. M. Moura, Thomas Daniels, Maarten Bosteels, Sebastian Castro, Moritz Müller, Thymen Wabeke, Thijs van den Hout, Maciej Korczyński, and Georgios Smaragdakis. 2024. Characterizing and Mitigating Phishing Attacks at ccTLD Scale. In *Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security (CCS '24)* (Salt Lake City, UT, USA). ACM, New York, NY, USA, 15. https://doi.org/10.1145/3658644.3690192
[66] Ximena J Nelson and Robert R Jackson. 2009. Aggressive use of Batesian mimicry by an ant-like jumping spider. *Biology letters* 5, 6 (2009), 755–757.
[67] Netcraft. 2024. *Cybercrime Detection.* https://www.netcraft.com/platform/cybercrime-detection/ Accessed: 2024-07-18.
[68] Netcraft. 2024. *Netcraft.* https://www.netcraft.com Accessed: 2024-07-18.
[69] Nominet. 2024. *.UK Register Statistics 2024.* https://www.nominet.uk/news/reports-statistics/uk-register-statistics-2024/ Accessed: 2024-07-18.
[70] NOS. 2024. *Internationaal phishing netwerk opgerold, vijf arrestaties in Nederland (In Dutch).* https://nos.nl/artikel/2517210-internationaal-phishing-netwerk-opgerold-vijf-arrestaties-in-nederland (Accessed on: Apr 18 2024).
[71] Adam Oest, Penghui Zhang, Brad Wardman, Eric Nunes, Jakub Burgis, Ali Zand, Kurt Thomas, Adam Doupé, and Gail-Joon Ahn. 2020. 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)*. USENIX Association, 361–377.
[72] Krebs on Security. 2023. *Sued by Meta, Freenom Halts Domain Registrations.* https://krebsonsecurity.com/2023/03/sued-by-meta-freenom-halts-domain-registrations/ Accessed: 2024-07-19.
[73] Robert Poulin. 2007. *Evolutionary Ecology of Parasites.* Princeton University Press, Princeton. https://doi.org/doi:10.1515/9781400840809
[74] Koushyar Rajavi, Tarun Kushwaha, and Jan-Benedict E M Steenkamp. 2019. In Brands We Trust? A Multicategory, Multicountry Investigation of Sensitivity of Consumers' Trust in Brands to Marketing-Mix Activities. *Journal of Consumer Research* 46, 4 (06 2019), 651–670. https://doi.org/10.1093/jcr/ucz026
[75] Revenue. 2023. *Welcome to revenue.ie.* https://www.revenue.ie/en/home.aspx Accessed: Jan. 24, 2024.
[76] Nikolaos Sarantinos, Chafika Benzaïd, Omar Arabiat, and Ameer Al-Nemrat. 2016. Forensic Malware Analysis: The Value of Fuzzy Hashing Algorithms in Identifying Similarities. In *2016 IEEE Trustcom/BigDataSE/ISPA*. 1782–1787. https://doi.org/10.1109/TrustCom.2016.0274
[77] SIDN. 2023. *General Terms and Conditions for .nl Registrants.* https://www.sidn.nl/downloads/5sWqyY0sTKHoWlCTm9RmZt/ef98ec32612ff200cfa94efe64b7341c/General_Terms_and_Conditions_for_nl_Registrants_20231001.pdf Accessed: Jan 30, 2024.
[78] SIDN. 2024. *Abuse Prevention.* SIDN. https://www.sidn.nl/en/cybersecurity/abuse-prevention Accessed: 2024-07-19.
[79] SIDN. 2024. *Complaining about the content of a website.* https://www.sidn.nl/en/nl-domain-name/complaining-about-the-content-of-a-website Accessed: Jan 30, 2024.
[80] SIDN. 2024. *SIDN - The Foundation for Internet Domain Registration in the Netherlands.* https://sidn.nl/en Accessed: 2024-08-29.
[81] SIDN. 2024. *SIDN For confidence online.* https://www.sidn.nl/en/internet-security/transparency-report [Accessed on 25th April 2024].
[82] SIDN. 2024. *SIDN's Policy on Taking Down Domain Names - Part 2.* https://www.sidn.nl/en/news-and-blogs/sidns-policy-on-taking-down-domain-names-part-2 Accessed: Jan 30, 2024.
[83] SIDN. 2024. *Taking Down Domain Names as a Sanction of Last Resort - Part 1.* https://www.sidn.nl/en/news-and-blogs/taking-down-domain-names-as-a-sanction-of-last-resort-part-1 Accessed: Jan 30, 2024.
[84] Ravindu De Silva, Mohamed Nabeel, Charith Elvitigala, Issa Khalil, Ting Yu, and Chamath Keppitiyagama. 2021. Compromised or Attacker-Owned: A Large Scale Classification and Study of Hosting Domains of Malicious URLs. In *30th USENIX Security Symposium (USENIX Security 21)*. USENIX Association, 3721–3738.
[85] Check Point Software. 2023. *Microsoft Dominates as the Most Impersonated Brand for Phishing Scams in Q2 2023.* https://www.checkpoint.com/press-releases/microsoft-dominates-as-the-most-impersonated-brand-for-phishing-scams-in-q2-2023/ Accessed: Jan. 24, 2024.
[86] Raffaele Sommese, Mattijs Jonker, Jeroen van der Ham, and Giovane C. M. Moura. 2022. Assessing e-Government DNS Resilience. In *2022 18th International Conference on Network and Service Management (CNSM)*. 118–126.
[87] Spamhaus. 2024. Strengthening trust and safety across the internet | Spamhaus. https://www.spamhaus.org Accessed: April 22, 2024.

[88] Yongjun Sung and Jooyoung Kim. 2010. Effects of brand personality on brand trust and brand affect. *Psychology & Marketing* 27, 7 (2010), 639–661. https://doi.org/10.1002/mar.20349
[89] Tom Symonds. 2024. *Police bust global cyber gang accused of industrial-scale fraud.* https://www.bbc.com/news/uk-68838977 Accessed: Apr 19, 2024.
[90] Samaneh Tajalizadehkhoob, Tom Van Goethem, Maciej Korczyński, Arman Noroozian, Rainer Böhme, Tyler Moore, Wouter Joosen, and Michel van Eeten. 2017. Herding Vulnerable Cats: A Statistical Approach to Disentangle Joint Responsibility for Web Security in Shared Hosting. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (Dallas, Texas, USA) *(CCS '17)*. ACM, New York, NY, USA, 553–567. https://doi.org/10.1145/3133956.3133971
[91] tdebatty. 2024. *java-spamsum.* https://github.com/tdebatty/java-spamsum Accessed: 2024-02-13.
[92] US Federal Bureau of Investigation, Internet Crime Complaint Center. 2023. Internet Crimer Report. https://www.ic3.gov/Media/PDF/AnnualReport/2023_IC3Report.pdf.
[93] Marie Vasek, John Wadleigh, and Tyler Moore. 2016. Hacking Is Not Random: A Case-Control Study of Webserver-Compromise Risk. *IEEE Transactions on Dependable and Secure Computing* 13, 2 (2016), 206–219. https://doi.org/10.1109/TDSC.2015.2427847
[94] Verisign. 2024. *Zone File.* https://www.verisign.com/en_US/channel-resources/domain-registry-products/zone-file/index.xhtml Accessed: 2024-07-18.
[95] Maarten Wullink, Giovane CM Moura, and Cristian Hesselman. 2018. Dmap: Automating Domain Name Ecosystem Measurements and Applications. In *Proceedings of the IEEE Network Traffic Monitoring and Analysis Conference*. IEEE, Vienna, Austria, 1–8.
[96] Qinge Xie and Frank Li. 2024. Crawling to the Top: An Empirical Evaluation of Top List Use. In *Passive and Active Measurement*, Philipp Richter, Vaibhav Bajpai, and Esteban Carisimo (Eds.). Springer Nature Switzerland, Cham, 277–306.
[97] Penghui Zhang, Adam Oest, Haehyun Cho, Zhibo Sun, RC Johnson, Brad Wardman, Shaown Sarker, Alexandros Kapravelos, Tiffany Bao, Ruoyu Wang, Yan Shoshitaishvili, Adam Doupé, and Gail-Joon Ahn. 2021. CrawlPhish: Large-scale Analysis of Client-side Cloaking Techniques in Phishing. In *2021 IEEE Symposium on Security and Privacy (SP)*. 1109–1124.

Appendices

A Examples of phishing domain types

We show the main types of phishing domains and their associated DNS traffic observed at the `.nl` authoritative servers, as shown in Figure 15.

Maliciously registered domains are those in which attackers register it themselves and take care of the associated steps – configuring its DNS and hosting the phishing website. We show an example of a maliciously registered domain in Figure 15a. This domain was used to impersonate a Dutch credit card issuer, and had words in Dutch language related to payment in its name. We see that the domain was registered after 12:00 UTC, and the first queries arrive a little after that (it may take up to an hour for a domain to be published in the `.nl` zone after registration). We see many hundreds of resolvers querying this domain after its registration, which is not usually seen in most domains. After 15:00, the domain is flagged as malicious and notified by Netcraft, and the domain is ultimately mitigated around 10:00 the day after, by being deleted from the `.nl` zone. We can see its steps in Figure 1.

Some attackers prefer to *age* the domain name before carrying out a phishing attack, *i.e.,*to wait a while (Dormant in Figure 1). Still, those aged domains are seen as maliciously registered, given their real intent to cause harm. Figure 15b shows an example of an aged domain, in which the attackers waited 75 days to impersonate multiple banks (Bank of Montreal, for instance, and several Dutch banks). We see a spike in the number of resolvers per hour when the attack is carried out. After that, the domain is mitigated by being deleted from the `.nl` zone.

Compromised domains, on the other hand, occur when attackers exploit someone else's website to host a phishing site. This is typically achieved by exploiting vulnerabilities in websites, particularly those using susceptible content management systems (CMSes) such as WordPress and Joomla [90, 93]. We show an example of a compromised domain in Figure 15c: a 21-year old domain hosting a small company selling vitamins, with very low DNS traffic, suddenly experiences a spike on resolvers seen by the `.nl` authoritative servers, around 14:00. Around 17:00, it is detected by Netcraft, and mitigated by having its website cleaned up at 18:30. The domain then resumed its normal traffic and remains with the same registrant. The phishing site hosted in its subdirectories was used to impersonate a webmail provider.

B ccTLDs abuse mitigation procedures

Next we discuss in detail the abuse handling policies of each registry, which we summarized for breivity in §2.6.

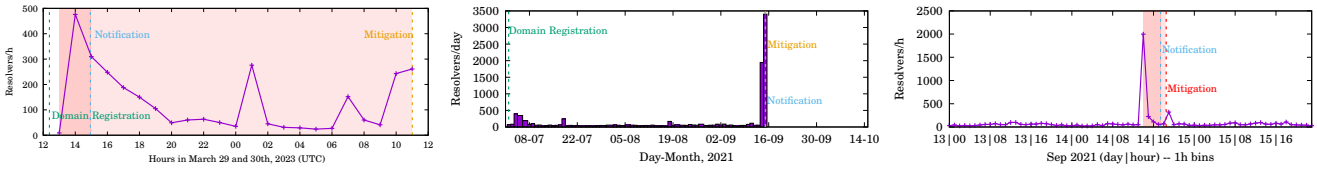
B.1 `.nl` abuse handling procedure

SIDN, the `.nl` registry, establishes its general terms and conditions for domain registrations as delineated in [77]. The policy for the takedown of malicious domain names is expounded in [82, 83].

Detection of malicious domains: `.nl` registry implements multiple methodologies to identify malicious domain names within its zone. These include the utilization of third-party blocklist providers (Netcraft), the application of internal systems that leverage machine learning to detect suspicious registrations [54], and the incorporation of reports received from external entities via its notice-and-take-down procedures [79].

Notification: `.nl` registry employs two strategies to notify the parties involved in malicious domain names. The first strategy involves the use of its blocklist provider notification service, which disseminates notifications via email to hosting providers, registrars, DNS providers, webmasters, and upstream providers. The second strategy involves direct contact with the registrant and registrar, either through corporate email or phone. The latter is utilized when contacting registrant data or when registrars or hosting providers fail to respond to Netcraft notifications.

Mitigation: The mitigation policy of `.nl` registry is contingent on the type of reported abuse. Reports of phishing and malware are accorded higher priority. The approach of `.nl` registry is to initially allow other involved parties, such as the hosting provider and registrars, to attempt mitigation of the malicious website. If these attempts are unsuccessful, then `.nl` registry will remove the domain name from the DNS zone (§6 and §18 in [77]) after 66 hours from the initial contact [82] with the registrant. The domain will remain registered but will be marked as inactive, thus preventing further registration while in this state. The rationale for this approach is discussed in detail in [82, 83], and stems from the principle that `.nl` registry does not assume the role of a content moderator or judge. The registrant is only contacted after the domain is marked as inactive.



(a) Maliciously registered: 1 day old **(b) Maliciously registered but aged: 75 days old** **(c) Compromised domain: 21 years old**

Figure 15: Phishing domains categories.

B.2 `.ie` abuse handling procedure

The `.ie` registry publishes its Registration Policy in [41]. The policy explaining the process for dispute resolution, which includes abusive domains, is available in [39].

Detection of malicious domains: `.ie` includes in its Registration Policy the need for registrant's to show a substantive connection to the country, and produce evidence supporting that relationship, as explained in §2 and §3 of [41]. All documentation provided is reviewed by humans, who decide if the registration should proceed, making the creation of a domain for malicious purposes difficult. `.ie` also uses third-party feed providers (Netcraft) in agreement with registrars for expedited remediation.

Notification: When a malicious domain name is identified by Netcraft, `.ie` will directly contact registrars, hosting and DNS companies directly asking for remediation. In particular, registrars working with `.ie` have guidelines designed in partnership on how to handle these reports.

A second path of notification can come from the public, by engaging the legal authorities, with `.ie` having an expedite Complaint Protocol. If a case comes through this channel, the Registry is in charge of contacting the Registrar and cases are handled within 1 or 2 business days.

Mitigation: Depending on how the malicious domain was reported, the mitigation will take place at different levels. A case coming from the legal authorities is handled with the highest priority, the registrant is contacted by all means possible. The domain could be suspended by the registry while pending investigation if deemed necessary. Cases observed by the blocklist provider are usually handled directly by the registrar or the web hosting company.

B.3 `.be` abuse handling procedure

DNS Belgium, the `.be` registry, employs both proactive and reactive measures to prevent abuse within the `.be` zone.

Detection of malicious domains: Already during the registration of a new domain name, a model determines the chance that a domain name is registered with malicious intent. If the predicted chance is above a predetermined threshold, the registrant will need to prove her identity before the domain name is delegated. Next to the automated assessment, all newly registered domain names are also subjected to manual screening, even if the registrant has verified his data with Registrant Verification. Note that this system has only been active since November 2020, so not during the whole period studied in this paper. Please also note that the detection system is improved constantly and a rule-based model has recently been replaced by a Machine Learning model.

Notification: As a reactive measure, DNS Belgium also continuously monitors various third party abuse feeds and notifies both registrar and registrant when phishing and/or malware is detected.

Mitigation: According to the Terms & Conditions of DNS Belgium, the registrant has to provide correct contact data during the registration of a domain name. If the provided data is suspected to be false or incomplete, the registrant is given 14 days to correct and verify the information. If the contact details are not updated within 14 days, the domain name will be removed from the TLD zone so that the domain will no longer work. However, when phishing or malware is detected, DNS Belgium will manually assess the registration and weigh the risk of mistakenly suspending a legitimate domain name against the risk of harm caused to internet users by a phishing attack. DNS Belgium will immediately suspend the domain from the `.be` zone if the risk of harm caused to users is deemed higher. Also in this case the registrant has 14 days to respond.

Next to the measures described above, DNS Belgium also has Notice & Action agreements with governmental entities: if for instance the Ministry of Economy reports a serious infringement, DNS Belgium will immediately make the `.be` domain name in question inaccessible.

C Most impersonated companies

Table 10 shows the top 10 impersonated companies per ccTLD, as measured by the number of SLDs used in phishing attacks in each dataset.

Microsoft tops the list. We see that Microsoft ranks number one by having most SLDs trying to impersonate it. Microsoft's products and services, such as Office 365, are used by millions of individuals and businesses worldwide. This large user base makes it an attractive target, and access to such accounts can provide them with e-mails and also a range of internal documents (Sharepoint) and multiple other systems used by victims' organizations. Microsoft also tops the list of other industry phishing reports [85]. For all ccTLDs, the median Microsoft SLD phishing domain name is at least 5 years, suggesting they are indeed compromised domain names.

Local banks and financial services. We see that in `.nl` and `.be`, their local banks or financial services companies also draw many attacks – but their median domain age is much lower than other international brands. We also see that for `.nl`, the most impersonated local company in Table 10 is a bank, which has also the largest

market share (41% in 2022) in the country [7]. For .be, we see the there is only one local company in the top 10 list.

D Extra figures

Figure 16 shows a histogram for the top 20 targeted countries, based on the targeted companies' country of origin. We see that even though we have three European ccTLDs, most of the target companies are US-based. These include companies like Microsoft and Netflix. Then, for only .nl and .be, the second most popular are their respective countries. For .ie, Irish companies rank number 8 among the top 20. For the rest, we see many European countries.

Consistency over time: we see that this exploitation of local companies is consistent over time. Figure 17 shows the results of median age per year, for each ccTLD.

Local companies . Now we focus on the local companies. For each ccTLD, we extract all their local companies and compute the median age of the SLDs used against the companies.

Figure 18 shows the results. In this figure, the x axis shows each individual local company, ordered by the number of domains (SLDs) they have been attacked with (left y axis). On the right y axis, we show these SLDs' median age. We see for both .nl and .be that few companies concentrate most of the SLDs, and they have a very low median age at notification time – indicating these are maliciously registered domain names. However, we still see some local companies that are rarely attacked, say, using a single SLD, with a high median age of more than 3 years.

Figure 19 shows the CDF for domains older than 7 days being deleted or records changed at the DNS level for mitigation. We do not see any indication from this graph that domain names are actively being deleted as a mitigation strategy. We see that these deletes occur mostly after 60 days of the notification, so they are likely to be due to the natural expiration times of domain names, and the domain name natural recycling. We see that 30–40% of the old domains have their NS records within the first day, for both .nl and .be. We see that .ie NS changes for old domains take longer to take place.

Figure 20–Figure 22 show the mitigation results for old domains (>7 days) for the three ccTLDs. We see that most domains are mitigated.

E Impersonated companies

Table 11 shows the list of company, sector, and coutnry of origin found accross all ccTLD, sorted by country of origin. We see mostly US-based companies.

F ccTLDs and impersonated companies compared

Table 12 shows the full table of all subsets, in addition to those discussed in Figure 7.

F.1 Remaining intersections

For the intersections with .ie, we see the same patterns: old domain names being used to target mostly international companies.

For the Dutch and Irish companies, we see 29 companies that only have phishing SLDs in the .nl and .ie zone. We see in Figure 7 that 26 of them are international, one Dutch and two Irish. The Dutch company is the national chamber of commerce ([52]), and the Irish company are the tax authority ([75]) and a road toll payment company.

We see how the Dutch company is impersonated using 12 .nl domain names (1 day median age), while the same company is impersonated using two old .nl domain names. This shows that the same companies can be exploited using different strategies, depending on who is carrying out the attack.

G Correlation between attacks across ccTLDs

G.1 Dutch bank case

Next we focus on phishing against one of the largest banks in the The Netherlands. This bank also operates in Belgium, but we account them as different entities.

Figure 23a shows the number of monthly SLDs on both .nl and .be for the Dutch branch of the bank. We see that .nl domains are far more used than .be domains for phishing attacks against the Dutch branch of the bank. We compute the Pearson coefficient (0.62) and then the R^2 value (0.39) between both curves and find a weak correlation.

In Figure 23b, we see phishing SLDs on the Belgian branch of the bank. We see that some months we see more SLDs on the .be while others more SLDs on the .nl zone. We compute the R^2 between both curves and find even a weaker correlation: 0.09.

We look into individual days of attacks to determine if there was coordinated campaigns on the same dates. Between 2019-08-29 and 2023-06-05, the Dutch branch of the bank has been targeted on .nl on 734 days and on the Belgian branch 112 days on .be, 90 of which overlap between the two TLDs. So we found no clear evidence of coordinated attacks. What we see, however, is the same SLD being used against multiple banks, given the usage of phishing kits.

G.2 Case study: Microsoft

Microsoft is the number one targeted company across all three ccTLDs (Table 7).

In this section, we investigate the monthly distribution of SLDs per ccTLD attacking Microsoft. Figure 24 shows the month SLDs per ccTLD, in stacked bar plot. We see no direct correlation in the attack monthly distribution: the highest R^2 is 0.3, and Pearson coefficient is 0.53, between .nl and .be, which is also a week one.

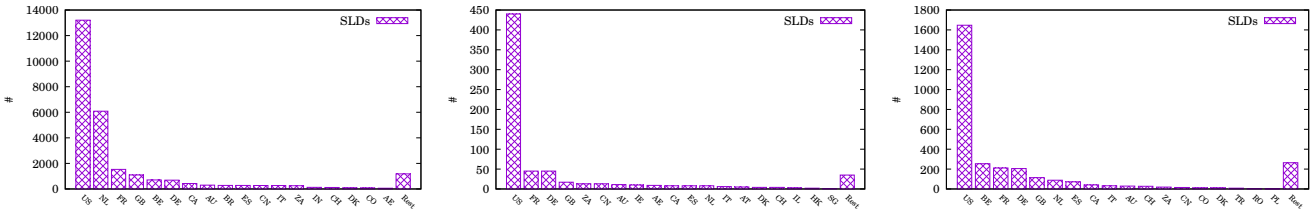
In this section, we investigate the monthly distribution of SLDs per ccTLD attacking Microsoft. Figure 24 shows the month SLDs per ccTLD, in stacked bar plot. We see no direct correlation in the attack monthly distribution: the highest correlation is 0.3, between .nl and .be, which is also a week one.

H Phishing domain screenshots

Figure 25 shows the screenshot of the French bank phishing described in §5.1, while Figure 26 shows the Tinder phishing mittigation.

Table 10: Top 10 impersonated companies for `.nl`, `.ie` and `.be`. Yellow rows are from local companies.

Rank	Company	URLs	SLDs	Median Age (days)	Avg. Age (days)	Country	Sector
The Netherlands ccTLD (`.nl`)							
1	Microsoft	11860	2319	2251.0	3028.0	US	Technology
2	PayPal	14574	2134	1751.0	2270.0	US	Financial Services
3	ING Netherlands	11906	1815	1.0	73.0	NL	Banking
4	International Card Services	3599	1410	2.0	360.0	NL	Financial Services
5	Apple	6019	1276	1775.0	2283.0	US	Technology
6	ABN AMRO	2800	1259	1.0	156.0	NL	Banking
7	Google	3489	1236	1416.0	1877.0	US	Technology
8	Rabobank	4704	1222	1.0	160.0	NL	Banking
9	Webmail Users	4022	1054	2247.0	2695.0	US	Internet Services
10	Netflix	5763	756	1653.0	2365.0	US	Entertainment
Ireland ccTLD (`.ie`)							
1	Microsoft	884	135	2598.0	3365.0	US	Technology
2	Webmail Users	205	60	1921.0	2760.0	US	Internet Services
3	Netflix	389	46	2792.0	2966.0	US	Entertainment
4	PayPal	337	24	2279.0	2519.0	US	Financial Services
5	DHL Worldwide Express	151	21	2301.0	3049.0	DE	Logistics
6	Amazon	188	15	3430.0	3359.0	US	E-commerce
7	United States Postal Service	65	14	2559.0	3032.0	US	Postal Services
8	La Banque Postale	284	13	2430.0	3218.0	FR	Banking
9	JP Morgan Chase	91	11	2784.0	3318.0	US	Banking
9	BNP Paribas Nickel	17	11	1855.0	2597.0	FR	Banking
Belgium ccTLD (`.be`)							
1	Microsoft	2496	424	2016.0	2909.0	US	Technology
2	Webmail Users	1270	284	1368.0	2629.0	US	Internet Services
3	Netflix	1839	170	3039.0	3594.0	US	Entertainment
4	PayPal	1174	133	1829.0	2651.0	US	Financial Services
5	DHL Worldwide Express	1184	118	3239.0	3627.0	DE	Logistics
6	ING Belgium	295	87	1.0	142.0	BE	Banking
7	Amazon	494	71	1318.0	2511.0	US	E-commerce
8	Apple	2220	66	820.0	1747.0	US	Technology
9	Google	355	65	288.0	1869.0	US	Technology
10	JP Morgan Chase	589	58	3904.0	3644.0	US	Banking



(a) `.nl` (b) `.ie` (c) `.be`

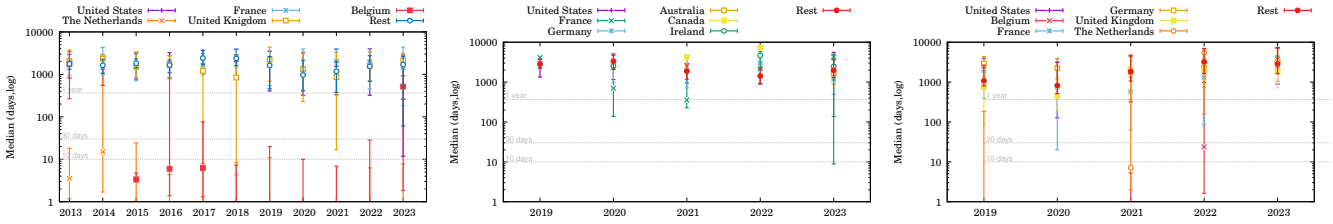
Figure 16: Top 20 countries from targeted companies, ordered by number of used SLDs in attacks.

I Extra phishing SLDs timelines

Not all phishing cases we found in our 234 new domains we have web crawling data have a clear cut pattern. Here we show a few cases.

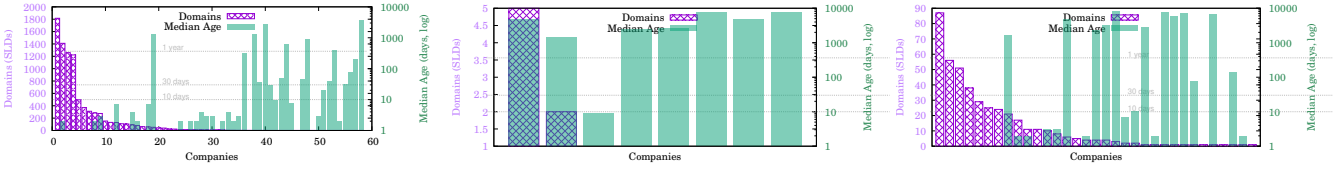
For each SLD, we combine data points from Netcraft and ours to build a timeline, as shown in Figure 9a, where we show a SLD which was used to impersonate the dating site Tinder. From Netcraft, we take notification and mitigation times (dashed lines). We plot the x axis relative the the mitigation time.

The figure shows the domain registration time (from the `.nl` registration database), active websites (from our in-house crawler). Our first web data point is immediately after registration, and it shows as "Unreachable". In this case, the domain was registered but not published in the `.nl` zone file, so it could not be resolved – the `.nl` zone is updated every 30 minutes, so there is a delay between registration and visibility. Later, our crawler finds web



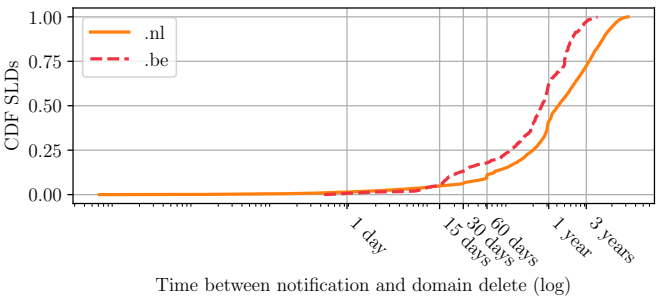
(a) `.nl` (b) `.ie` (c) `.be`

Figure 17: SLDS age over time, per brand country. Error bars shows the first and third quantiles.

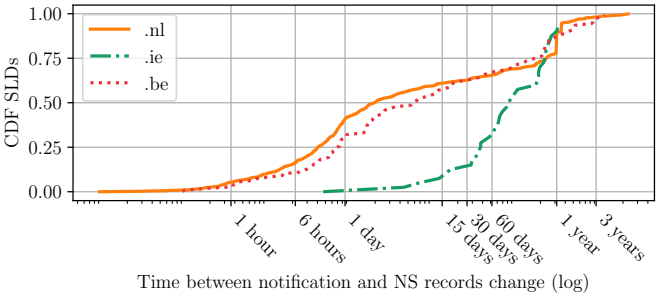


(a) `.nl` (b) `.ie` (c) `.be`

Figure 18: Local companies: number of SLDs and median age (days) using in phishing attacks.



(a) Domain delete



(b) Domain suspension

Figure 19: Phishing Mitigation of old domains (> 7 days old) at DNS level.

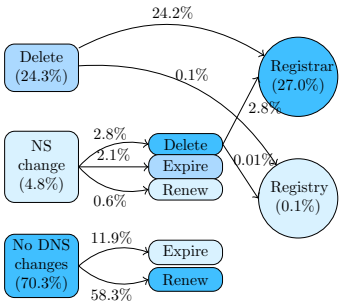


Figure 20: `.nl` (19,160 domains): DNS level mitigation of old domains.

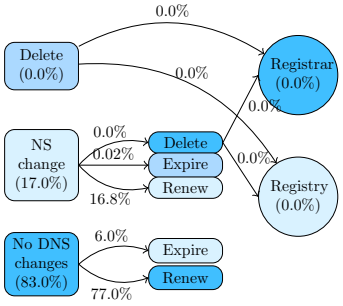


Figure 21: `.ie` (534 domains): DNS level mitigation of old domains.

pages. We classify them into two websites, each determined by a context triggered piecewise hashes value (CTPH, also known as fuzzy hashes [13, 76, 91]).

Table 11: Common 139 companies found in all three ccTLDs.

Country	Company	Sector	Country	Company	Sector
AE	Aramex	Logistics	IT	Aruba SPA	Web Services
AE	Emirates post Group	Postal Services	KR	Naver	Internet Services
AT	Bank Austria	Banking	NL	SNS Bank	Banking
AT	Bank99	Banking	NL	WeTransfer	File Sharing
AT	Paylife	Financial Services	NZ	Bank of New Zealand	Banking
AU	ANZ	Banking	RO	ING Romania	Banking
AU	Australian Government	Government Services	RO	Raiffeisen Bank	Banking
AU	Services Australia	Government Services	SE	Spotify	Music Streaming
AU	Australia Post	Postal Services	SG	Singapore Post	Postal Services
BE	Argenta	Banking	UK	Postmaster	Postal Services
BE	Government of Belgium	Government Services	US	Bank of America	Banking
CA	Royal Bank of Canada	Banking	US	Citizens Bank	Banking
CA	Shopify	E-commerce	US	Fifth Third Bank	Banking
CA	Neteller	Financial Services	US	First Horizon National Corporation	Banking
CA	Canada Post	Postal Services	US	Huntington Bank	Banking
CA	Bell Canada	Telecom	US	JP Morgan Chase	Banking
CH	Swiss Post	Postal Services	US	Keybank	Banking
CH	Swiss Federal Railways	Transportation	US	M&T Bank	Banking
CL	Banco Estado	Banking	US	Regions Bank	Banking
CN	Made-in-China.com	B2B Marketplace	US	Truist	Banking
CN	Alibaba	E-commerce	US	Wells Fargo	Banking
CN	NetEase	Internet Services	US	Discover	Banking/Credit Cards
CN	SF Express	Logistics	US	Rackspace	Cloud Services
CN	EMS China	Logistics/Shipping	US	Dropbox	Cloud Storage
CO	BBVA Colombia	Banking	US	Box	Cloud Storage/Collaboration
CO	Bancolombia	Banking	US	Alaska FCU	Credit Union
DE	Deutsche Kreditbank AG	Banking	US	America First Credit Union	Credit Union
DE	ING Germany	Banking	US	Idaho Central Credit Union	Credit Union
DE	Sparkasse	Banking	US	Mountain America	Credit Union
DE	Targo Bank	Banking	US	Navy Federal Credit Union	Credit Union
DE	Volksbanken Raiffeisenbanken	Banking	US	RBFCU	Credit Union
DE	postbank.de	Banking	US	SECU Credit Union	Credit Union
DE	DHL Worldwide Express	Logistics	US	Blockchain	Cryptocurrency
DE	IONOS	Web Hosting	US	Amazon	E-commerce
DE	GMX	Web Services	US	DocuSign	Electronic Signature
DK	Danske Bank	Banking	US	Netflix	Entertainment
DK	Maersk	Logistics/Shipping	US	American Express	Financial Services
DK	Nets.eu	Payment Solutions	US	Cash App	Financial Services
ES	Banco Santander España	Banking	US	PayPal	Financial Services
ES	La Caixa	Banking	US	IRS	Government Services
ES	Correos Spain	Postal Services	US	America Online	Internet Services
FR	BNP Paribas	Banking	US	EarthLink	Internet Services
FR	BNP Paribas Nickel	Banking	US	Webmail Users	Internet Services
FR	BRED Banque Populaire	Banking	US	Yahoo	Internet Services
FR	Credit Agricole	Banking	US	FedEx	Logistics
FR	Credit Mutuel de Bretagne	Banking	US	UPS	Logistics
FR	La Banque Postale	Banking	US	United States Postal Service	Postal Services
FR	Societe Generale	Banking	US	Facebook	Social Media
FR	OVH	Cloud Computing	US	LinkedIn	Social Media
FR	Cetelem	Financial Services	US	Adobe	Technology
FR	Impot (French Tax Authority)	Government Services	US	Apple	Technology
FR	L'Assurance Maladie (French Public Healthcare)	Healthcare/Insurance	US	Google	Technology
FR	Chronopost	Logistics	US	Microsoft	Technology
FR	DPD Express Parcel Delivery	Logistics	US	ATT	Telecom
FR	La poste	Postal Services	US	Xfinity	Telecom
FR	Orange	Telecom	US	Godaddy	Web Services
GB	Lloyds Bank	Banking	US	weebly	Website Development
GB	TSB	Banking	ZA	ABSA	Banking
GB	DVLA	Government Services	ZA	First National Bank of South Africa	Banking
GB	Gov.uk	Government Services	ZA	Nedbank	Banking
GB	HM Revenue & Customs	Government Services	ZA	Standard Bank	Banking
GB	TV Licensing	Government Services	ZA	Sars	Government Services
GB	National Health Service	Healthcare	ZA	Post Office South Africa	Postal Services
GB	British Telecom	Telecom	NaN	Luno	Cryptocurrency
HK	Global Sources	Online Marketplace	NaN	Metamask	Cryptocurrency
HU	Magyar Posta Zrt.	Postal Services	NaN	Pancakeswap	Cryptocurrency
IE	Permanent TSB	Banking	NaN	Trust Wallet	Cryptocurrency
IE	Eircom	Telecom/Internet Services	NaN	track-trace.com	Logistics/Shipping
IT	BNL	Banking	NaN	FluBot Lure Proxy	Unknown
IT	Poste Italiane	Postal Services			

The first website is a landing web page, so it's a benign one. Later, this website is replaced by the actual phishing website (Site #2), and our crawl visits it 7 times before the Netcraft notification. Netcraft detects phishing at -3h, and our next measurement at -1h sees a HTTP 403 error, indicating mitigation. Subsequent measurements are classified as unreachable. This is because this domain was deleted from the `.nl` zone, which occurred at -12 min, shown as a blue line.

Figure 27 shows the timeline for a phishing attack used against an online marketplace. The phishing is placed under a subdirectory (`SLD/pay/brand`). We see that the domain is registered around -100 hours. After that, our `DMap` crawler retries 8 websites from the main page at the SLD level. Our crawler was then redirected to

Table 12: Subsets characteristics. SLDs refers to non-unique SLDs, given a same SLD can be used for multiple companies. Unique SLDs are found in Table 3.

			SLDs				Median age (days)		
Subset	Companies	Countries	SLDs	.nl	.ie	.be	.nl	.ie	.be
∪ all	1,233	86	37,215	32,762	728	3,725	1,641	2,746	2,154
∩ all	139	27	21,175	18,040	651	2,484	1,938	2,739	1,921
intl. companies	133	24	20,090	17,084	636	2,370	1,971	2,748	1,967
Dutch companies	2	1	573	549	6	18	659	3,460	614
Irish companies	2	1	61	31	6	24	3,570	3,535	3,599
Belgian companies	2	1	451	376	3	72	4	1,695	812
(.nl ∩ .ie) - ∩ all	29	15	268	231	37	–	1,891	–	3,205
intl. companies	26	13	247	214	33	–	1,949	–	2,991
Dutch companies	1	1	14	12	2	–	1	–	3,987
Irish companies	2	1	7	5	2	–	1,726	–	6,147
(.nl ∩ .be) - ∩ all	247	51	11,611	10,603	–	1,008	1,394	–	1,969
intl. companies	215	49	3,673	3,064	–	609	1,585	–	2,490
Dutch companies	13	1	7,151	7,061	–	90	1	–	2
Belgian companies	19	1	787	478	–	309	2	–	2
(.ie ∩ .be) - ∩ all	12	10	38	–	18	20	–	2,682	1,842
intl. companies	11	9	34	–	16	18	–	2,744	1,731
Irish companies	1	1	4	–	2	2	–	1,483	2,911
Belgian companies	0	–	–	–	–	–	–	–	–
.nl only	639	70	3,888	3,888	–	–	1,602	–	–
intl. companies	597	69	2,808	2,808	–	–	1,756	–	–
Dutch companies	42	1	1,080	1,080	–	–	4	–	–
.ie only	20	12	22	–	22	–	–	2,398	–
intl. companies	17	11	19	–	19	–	–	2,255	–
Irish companies	3	1	3	–	3	–	–	2,540	–
.be only	147	38	213	–	–	213	–	–	5,671
intl. companies	135	37	164	–	–	164	–	–	6,340
Belgian companies	12	1	49	–	–	49	–	–	108

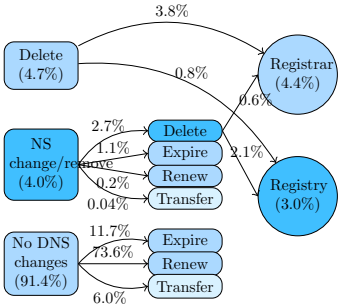
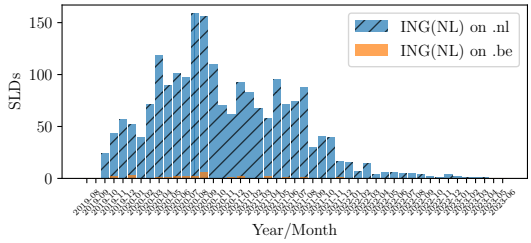
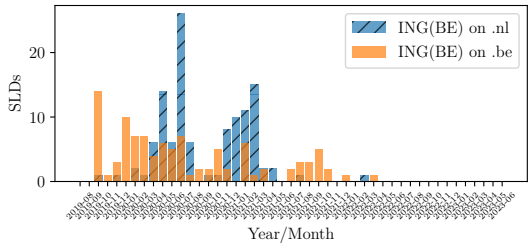


Figure 22: .be (2,359 domains): DNS level mitigation of old domains



(a) Dutch branch



(b) Belgian branch

Figure 23: Case 1: Phishing against a large Dutch bank.

Google URLs. This may be an attempt to pass as a legitimate website while the phishing is hidden in a subdirectory. Still, we see that after mitigation our datasets show the domain has been mitigated – we could no longer resolve the domain name. This domain was ultimately fixed using DNS, but it was left to expire after its one year lease.

Figure 28 shows another SLD whose phishing was hidden in a subdomain (subdir.example.com), thus not directly reachable by our crawler. We see that after registration, our crawler retrieves site #1 , which is the hosting provider's default landing page, which is typically included whenever new domains are registered. After that, we see multiple HTTP404 error codes, which shows that the web site is not found. Then, we see two visits to another side – which seems like the default page of Microsoft ISS web server, judging from its title. From that point on, our measurements see only HTTP 403 error, which can be seen as a sign of mitigation (Forbidden), or cloaking. Regardless, this phishing site hidden under a directory remains unreachable to our SLD crawler, but see that its SLD status change even before mitigation.

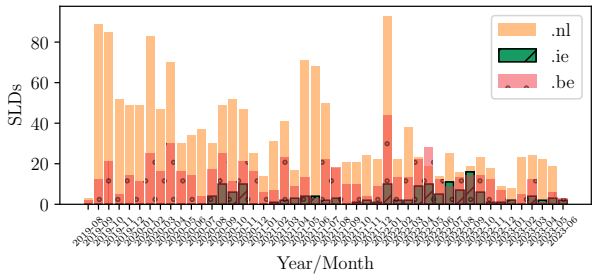


Figure 24: Phishing attacks impersonating Microsoft

Another type of phishing hiding we saw is shown in Figure 29. While the phishing was on the SLD level, it was not at `index.html` – it was on a different file name (`newsci.php`). We see that after registration, the first crawl fails – the domain is not yet published in the DNS zone. After that, the next round of measurements return HTTP 503 errors, which indicate "Service Unavailable". After the notification, however, we see that the website becomes unreachable: the name cannot any longer be resolved, although the domain was not deleted. So we can say that there is DNS level mitigation, but we cannot say if the HTTP503 are simply web cloaking, trying to deceive crawlers like ours.

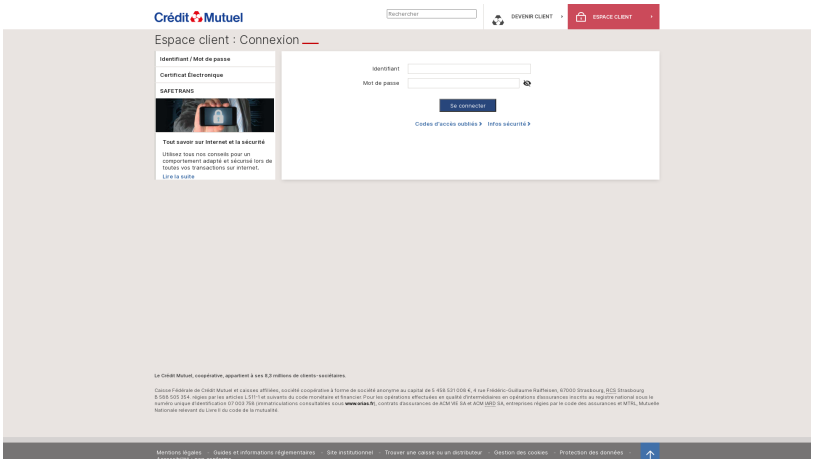


Figure 25: French bank phishing website detected by `.nl` in-house crawler (DMAP) (2022-09-30T04:55:24)

.

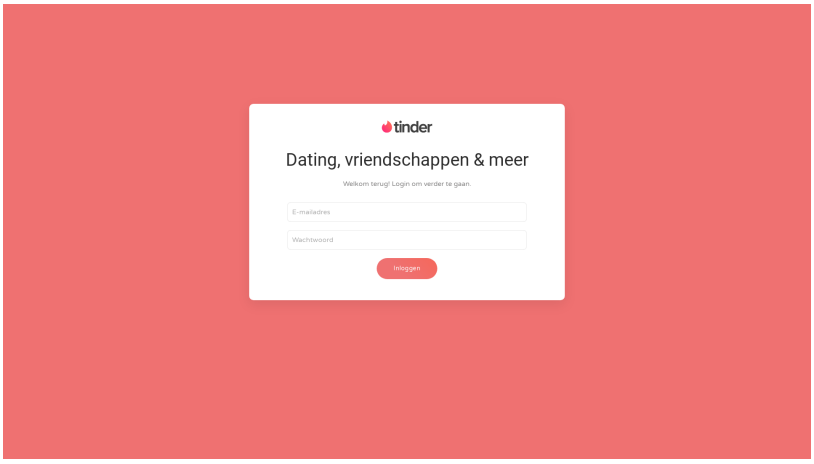


Figure 26: Tinder phishing website detected by `.nl` in-house crawler (DMAP). (2022-12-27T12:01:29)

.

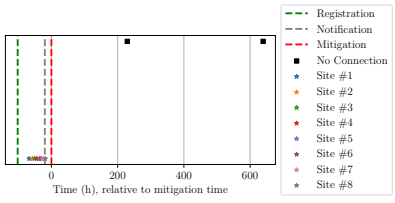


Figure 27: Phishing hidden in a subdirectory, while SLD redirects to Google.

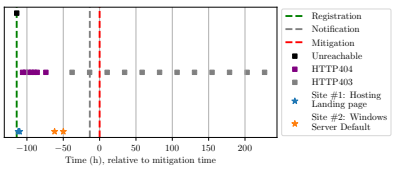


Figure 28: Phishing hidden in a subdomain.

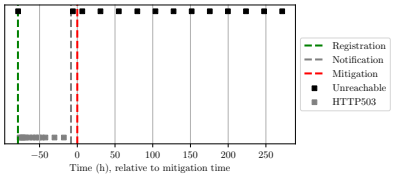


Figure 29: Phishing hidden in a different file at the SLD level.