Building a Resilient Domain Whitelist to Enhance Phishing Blocklist Accuracy

Jan Bayer*§, Sourena Maroofi§, Olivier Hureau*, Andrzej Duda*§, Maciej Korczynski*§

**Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, France*
firstname.lastname@univ-grenoble-alpes.fr

§*KOR Labs Cybersecurity, France*
firstname.lastname@korlabs.io

Abstract—**Phishing attacks constitute a significant threat to Internet users. One strategy for mitigating this threat involves the use of blocklists by Internet Service Providers (ISPs), companies, and organizations to block potentially malicious traffic. Phishing blocklists offer protection by continuously publishing a multitude of malicious URLs. However, community-supported and automated methods for constructing these blocklists may occasionally result in false positives, erroneously flagging benign domains or URLs as malicious.**

This paper addresses the challenge of reducing false positives in blocklists by proposing a robust scheme for constructing a domain whitelist containing domain names that are highly unlikely to be involved in malicious activities. We mitigate the risk of false negatives, referring to instances where malicious domains or URLs are inaccurately labeled as benign within the whitelist. Our approach is grounded on two key principles: i) the selection of meticulously curated seed domain names encompassing high-profile domains and ii) a careful procedure for validating disputed and defensively registered domains, ensuring their inclusion in the whitelist meets rigorous criteria. The scheme uses four methods for including a domain in the whitelist based on several publicly available data sources: i) reports published by approved dispute resolution service providers, ii) the information on shared in-bailiwick name servers, iii) the domain name WHOIS information, and iv) the information in TLS certificates. We evaluate the quality of our scheme by applying the constructed whitelist to various blocklists to detect false positives.

Index Terms—**phishing, whitelist, blocklist, DNS**

I. Introduction

Nowadays, both enterprises and individual users heavily rely on the use of blocklists as a defensive measure against cyber threats such as phishing and spam. Blocklists consist of compiled sets of URLs or domain names associated with malicious activities or exhibiting indicators of malicious behavior. The lists are often sourced from reputable third-party organizations such as the Anti-Phishing Working Group (APWG) [1] or OpenPhish [2]. Alternatively, they can come from community-driven platforms like PhishTank [3] or are directly integrated into web browsers (e.g., Google Safe Browsing [4] and Microsoft SmartScreen [5]).

Given a vast quantity of URLs and the registration of over 200 thousand new domains per day [6] as well as an important number of phishing attacks in 2022 documented by APWG [7], manual creation of blocklists becomes nearly impossible for security organizations—identifying such a vast number of blocklisted URLs requires an automated approach capable of evaluating tens of thousands of URLs daily. Previous research showed promising results by using machine learning techniques for identifying abusive domain names and URLs [8–18]. These systems generally involve four steps: i) fetching data from feeds, ii) excluding benign domains and URLs from the input data, iii) extracting features to generate a feature vector and iv) analyzing the feature vector using a (machine-learning) model.

To enhance the exclusion of benign domains and avoid false positives, researchers commonly rely on ranking lists [19–23] such as Alexa[1] [24] or Tranco [25] (e.g., URLHaus [26] uses the Tranco list). While these lists effectively reduce the initial dataset by excluding numerous seemingly benign domains, they lack the robustness required for a dependable use as a whitelist or an exclusion list because their primary focus is on measuring domain popularity rather than assessing domain benignness [27], including high-profile or critical domains. For example, **`franklinbnk.com`** (Franklin Bank) is not ranked in Tranco 1M, yet it is frequently targeted in phishing attacks.

Moreover, the presence of nonexisting (expired) domain names within these ranking lists creates an exploitable opportunity for malicious actors. Phishers can register recently expired domains and leverage their appearance in popularity-based lists, thereby evading detection in various security systems. For instance, in May 2023, out of the top 500k domains listed in Tranco, a total of 465 domains remained unregistered and susceptible to misuse.

Reliable whitelists help prevent false positives in blocklist feeds and, therefore, play a crucial role for three main reasons. First, current anti-phishing engines, such as APWG, may classify URLs and defensively registered domain names as malicious due to similarities in extracted features, which may lead to the risk of blocking legitimate URLs for end-users and lower confidence in blocklist feeds. Defensive domain name registration is a strategy in which individuals or organizations register similar domain names to protect their brand, trademarks, and online identity from potential misuse or infringement by others.

[1] Alexa ranking is discontinued since February 1, 2023: https://tranco-list.eu

Second, whitelists can alleviate processing burdens by excluding benign domains that are highly unlikely to engage in malicious activities from subsequent analysis and classification processes. Third, registrars and Top-Level Domain (TLD) registries frequently verify the accuracy of blocklisted domain names manually upon receiving abuse notifications. When in doubt, they may contact the domain owner, a process that is both time-consuming and costly. When such false positive situations occur, the investment in time and resources is unproductive.

In this paper, we propose a robust scheme for generating a reliable domain whitelist that can be used in the preprocessing analysis step for evaluating candidate URLs for malicious activities. The scheme eliminates the risk of false positives in the whitelist with two principles: i) the selection of meticulously curated seed domain names encompassing high-profile domains and ii) a rigorous procedure for validating disputed and defensively registered domains, ensuring their inclusion in the whitelist meets stringent criteria.

Unlike previous methods (e.g., those proposed by Burton et al. [28]), we do not use popularity lists as the initial data source. While our whitelist is not an exhaustive list of all defensive registrations and brand names, it guarantees that legitimate companies own all entries and are not maliciously registered or expired. This assurance comes from our conservative construction approach, founded on active Internet measurements and meticulous analysis.

Theoretically, an attacker could compromise a domain that appears on the whitelist at the DNS or website levels. However, such an event is highly unlikely because of stringent security verifications of the listed domains.

In summary, our contributions are as follows:

- We propose a scheme for generating a reliable domain name whitelist based on multiple domain name sources such as public DNS, domain registration data (WHOIS), TLS certificates, and data from UDRP (*Uniform Domain-Name Dispute-Resolution Policy*) [29] dispute-resolution service providers.
- The scheme involves selecting high-profile brand names and generating variants through techniques like typosquatting, bitsquatting, combo-squatting, and identifying homographs and homophones. We then query these variants for their `NS` records and include in the whitelist the domains whose name servers are *in-bailiwick* and *in-domain*.
- Additionally, we incorporate into the whitelist the domains involved in domain dispute processes, registered by well-known defensive registrars, and present in TLS certificates.
- We evaluate the quality of our scheme by applying the constructed whitelist to various blocklists to detect false positives.

Researchers can access the proof-of-concept whitelist, which includes 17,215 entries, at the provided address: **`https://whitelst.com`**.

II. Methodology

This section outlines our approach to gathering the initial datasets and offers an intricate breakdown of the techniques for creating a reliable domain name whitelist. Figure 1 presents the modules used in our scheme with overarching data sources incorporated into our proposed methodology to pinpoint the domains eligible for inclusion in the whitelist.

Our data collection spans several sources: we use a lexical module to target the domains identified within existing phishing blocklists ①, we examine domain disputes reported through the ICANN *Uniform Domain-Name Dispute-Resolution Policy* (UDRP) [29] ②, undertake active DNS measurements ③, extract registration details from WHOIS records ④, and inspect the fields related to domain names in Transport Layer Security (TLS) certificates ⑤.

A. Gathering Initial Datasets

The objective of the lexical module (① in Figure 1) is to curate a comprehensive repository of domains encompassing the namespace that could potentially attract squatting. We achieve this goal by combining the most frequently targeted high-profile domain names, squatting candidates extrapolated from them, distinctive keywords extracted from domains used in phishing attacks, and suffixes sourced from the Public Suffix List (PSL) [30].

Table I shows the lexical datasets we manage, with each row corresponding to a phase in the generation of the final candidate dataset. We describe below all the datasets.

Table I
Datasets

DS name	Description	Data source	Example	Size
$DS_{refDomains}$	high-profile domains targeted by phishing attacks	APWG [1], OpenPhish [2], PhishTank [3]	**`google.com`** **`ebay.com`**	338
$DS_{refBrands}$	brand names, i.e. e2LL	$DS_{refDomains}$	**`google`** **`ebay`**	305
DS_{kWords}	special keywords from domains used in phishing attacks	*Bayer et al.* [31]	**`support`** **`online`** **`secure`**	28
$DS_{candRoots}$	candidate roots generated by **`dnstwist`** (①)	$DS_{refBrands}$ [30]	**`góóglë`** **`ebay-help`**	1.69M
DS_{suffix}	active suffixes and TLDs	PSL [30]	**`org`** **`co.br`**	5,308
DS_{cand}	final candidates for squatting domains	$DS_{candRoots} \times DS_{suffix}$	**`gogle.org`** **`ebay.co.br`**	8.9B

1) Reference Domain Names: The initial step involves compiling a list of major companies frequently targeted by phishing attacks. Previous studies commonly relied on ranking lists such as Alexa or Tranco [25] to identify prominent domains and brands [21, 32–35].

As the ranking lists may not accurately reflect the interest of attackers in a domain name, we extract highly targeted brand names from reputable phishing blocklists, namely APWG [1],

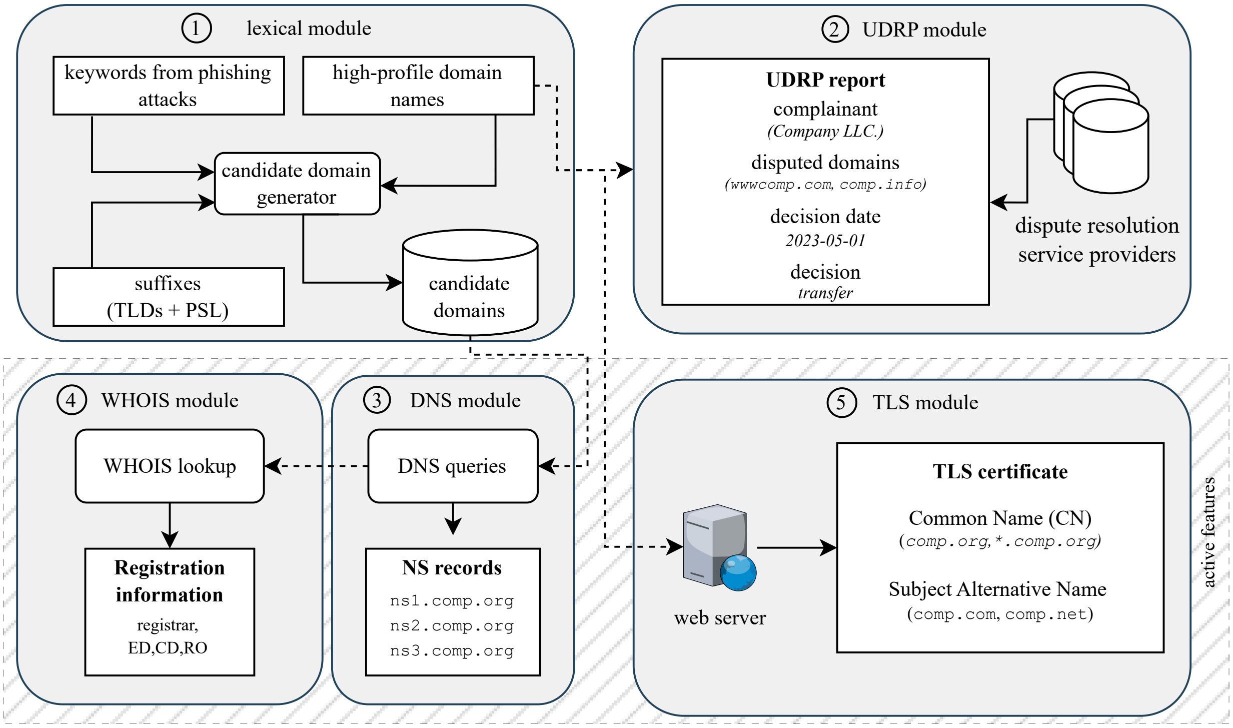

Figure 1. Overview of the modules and data sources for building a domain whitelist.

PhishTank [3], and OpenPhish [2]. These domains are referred to as "reference domains". Our reference domain dataset ($DS_{refDomains}$) consists of 338 unique names, representing 305 unique effective second-level labels (e2LL) ($DS_{refBrands}$).

We ensure their validity as well-known companies by manually inspecting their websites. The selection of reference domains provides us with the initial list, grounded in the assumption that the likelihood of compromising these reference domain names at both the DNS and website levels is minimal. Considering an increasing number of phishing attacks and instances of abuse [36], major companies actively invest in fortifying their infrastructure to enhance security and resilience against such threats.

2) Candidate Domains: The $DS_{refBrands}$ dataset serves as input to our candidate domain generation algorithm. In the initial phase, we use a tailored version of **`dnstwist`** [37], a rule-based tool to generate variants for each reference brand name encompassing homographs, typo-squatting, bitsquatting, and homophones. We also incorporate combo-squatting candidates generated from a curated list of 28 keywords (the DS_{kWords} dataset) observed in previous phishing attacks [31] that includes terms like *login*, *sign*, *secure*, etc. Overall, the $DS_{candRoots}$ dataset contains approximately 1.7 million candidate roots derived from 305 brand names.

The next step involves the generation of a series of suffixes to append to each entry in the set of candidate root and effective second-level labels ($DS_{candRoots}$ and $DS_{refBrands}$). To accomplish this, we use suffixes from the Public Suffix List (PSL) [30]. The PSL contains Top-Level Domains (TLDs) (e.g., **`com`**, **`cz`**, **`dad`**) and higher-level suffixes (e.g., **`com.br`**, **`k12.ca.us`**). We solely use the public portion of the PSL comprising suffixes managed by registry operators as opposed to private suffixes operated by other entities such as cloud providers.

The majority of higher-level suffixes have only a limited number of registered domains and are predominantly linked to highly specialized purposes, such as educational institutions or libraries. For example, a manual examination of 3rd- and higher-level suffixes from the PSL revealed that a significant portion of approximately 2,000 3rd-level suffixes included in the list are associated with Japan (e.g., **`isshiki.aichi.jp`** or **`bunkyo.tokyo.jp`**). Considering the relationship between the suffix level and its specificity of usage, we opt for top-level and second-level public suffixes (e.g., **`com`**, **`com.br`**, **`co.uk`**) that are more likely to be available for registration via public registrars.

3) Domain Exclusion Based on DNS and WHOIS: The constructed dataset comprises numerous domain names for which we must gather additional DNS and WHOIS data. Considering the extensive scale of this process, we make exclusions based on specific criteria. First, we exclude *non-delegated* suffixes corresponding to the cases for which the parent name servers do not provide referrals to child zones of a given suffix as specified in RFC 8499 [38]. Additionally, we omit *catch-all* suffixes, referring to suffixes for which name servers return **`NOERROR`** DNS responses even for non-existing domain names (e.g., **`ws`**).

This step follows Algorithm 1. We perform a query for

Algorithm 1 Algorithm for excluding PSL suffixes

```
isCatchall = False
suffixNSSet = iterative NS scan
if rcode of suffixNSSet is NOERROR then
    isDelegated = True
    for ns in suffixNSSet do
        rndDomain = genRndDomain() + "." + suffix;
        ARec = dnsA(rndDomain, ns);
        if rcode of ARec is NOERROR then
            isCatchall = True
            break;
else
    isDelegated = False
```

the **NS** (name server) records of each suffix in our designated subset of the PSL suffixes. If the response code is **NXDOMAIN**, indicating that the domain does not exist, we mark the suffix as *non-delegated* and exclude it from the DS_{suffix} dataset. Among the suffixes that remain, we generate a lengthy and randomly composed domain name that is highly likely to be nonexistent and we initiate a query to retrieve its **A** record. If the name servers return a **NOERROR** response code, we mark the suffix as a *catch-all*. With this algorithm, we successfully identify and remove 86 *catch-all* and *non-delegated* suffixes from the dataset resulting in a refined list of 5,308 suffixes (DS_{suffix}).

By performing the Cartesian product of $DS_{candRoots}$ and DS_{suffix}, we generate a comprehensive list of candidate domain names for the next stage. This final dataset of candidate domains is an extensive collection of 8.9 billion records (DS_{cand}).

For each record, we maintain the information on the root (brand name) used for its generation. Occasionally, two brands may generate the same candidate domain using distinct methods for candidate generation. For instance, the candidate domain **bisa.com** resulted from two seeds: **bisa** using the original root and **visa** as a typosquatting candidate replacing character **v** by **b**. To ensure traceability, we incorporate the origin of the generated name into our algorithm, which enables the mapping of such candidates back to their corresponding original brand names.

B. Methods for Constructing a Reliable Domain Name Whitelist

We have designed five methods for constructing a reliable domain name whitelist that take advantage of publicly available data sources: i) **DD** (Disputed Domains), based on reports published by approved dispute resolution service providers, ii) **SINS** (Shared In-Bailliwick Name Servers), based on the information on shared in-bailiwick name servers, iii) **RO** (Registrant Organization), based on the WHOIS information on the registrant organization, iv) **DR** (Defensive Registrars), based on the WHOIS information on the defensive registrars, and v) **DC** (Domain Certificates), based on the information in domain certificates. After conducting experiments with all five methods, we have opted to exclude the RO method from our scheme due to its susceptibility to being circumvented by malicious actors, which might lead to false negatives in the whitelist.

1) ***DD****: Disputed Domains:* The domains that resemble trademarks or personal names may lead to legal disputes known as 'cybersquatting' [39]. To simplify the resolution process for trademark owners and avoid lengthy lawsuits, ICANN introduced UDRP [29]. Trusted legal organizations, such as the World Intellectual Property Organization (WIPO), act as dispute resolution service providers with panelists from different countries. UDRP reports include the details about the involved domain names, the complainant (trademark owner), the respondent (violating registrant), the evaluation of cybersquatting, and the panel decision with its date. The UDRP fees range from $1500 to $5000 per complaint, making it a more cost- and time-effective option than Anticybersquatting Consumer Protection Act (ACPA) lawsuits [40]. However, companies must still allocate significant resources to domain security, making them high-profile participants in dispute processes.

Our UDRP module (② in Figure 1) uses the data from four dispute resolution (D-R) service providers: WIPO [41], the Alternative Dispute Resolution (ADR) Forum [42] formerly known as National Arbitration Forum, the Asian Domain Name Dispute Resolution Center (ADNDRC) [43], and the Canadian International Internet Dispute Resolution Center (CIIDRC) [44]. We extracted a set of 135,043 unique disputed domains that were i) transferred to the complainant and ii) for which we could collect the decision date.

The UDRP policy states that the decision is implemented ten business days from the decision date unless the respondent files a lawsuit against the complainant [29]. We consider a resolved UDRP complaint effective thirty days after the decision date, i.e., we add twenty days of margin to cover possible differences in the definition of business days as seen at the ICANN principal office.

While the manual analysis indicates that a substantial portion of the disputed domains are converted into defensive registrations, we refine our focus and exclusively include in the whitelist the domains associated with the meticulously curated set of prominent brands denoted as $DS_{refBrands}$.

UDRP reports contain information about the complainant, i.e., the legal name of a company or a physical person that filed the complaint with one of the Dispute Resolution providers. As an example, in November 2022, *PayPal* filed a complaint[2] with the ADR Forum regarding the **ppbpaypal.com** domain. The panelist ruling resulted in transferring the contested domain to *PayPal*. At the time of writing, this domain continues to be defensively registered and remains under the supervision of *PayPal*, making it a valuable addition to our whitelist.

Note that the *complainant* value in UDRP reports may contain different legal names linked to one organization. For

[2] Case no. 2020641: https://www.adrforum.com/DomainDecisions/2020641.htm

example, reports in which *PayPal* figures as the complainant contain different values:

`PayPal, Inc.`
`PayPal Inc.`
派普尔公司 **`(Paypal, Inc.)`**

Algorithm 2 Algorithm for domains from UDRP reports

```
$brand\_RO$, *RO → brand mapping*
$brand\_KW$, *unique keywords for each brand*
$mes\_dt$, measurement date
procedure UDRPFilter($compl, decision, dec\_dt, rds$)
    brand = null
    if $dec\_dt > (mes\_dt - 30\ days)$ then
        return decision not yet implemented
    $brand = extract\_brand(compl)$
    for $domain$ in $rds$ do
        cd, is_defreg = analyzeWhois($domain$)
        if $cd \geqslant dec\_dt$ then
            continue    ▷ re-registered after decision date
        else if $brand \neq null$ then
            $addToWhitelist(rd, brand)$
        else if $is\_defreg = true$ then
            $ro = roFromThinWhois(domain)$
            $brand = extract\_brand(ro)$
            if $brand \neq null$ then
                $addToWhitelist(rd, brand)$
            continue
procedure EXTRACT_BRAND($ro\_or\_com$)
    if $ro\_or\_com$ in $brand\_RO$ then
        return $brand\_RO[compl]$
    else
        for $kword$ in $brand\_KW$ do
            if $kword$ in $ro\_or\_com$ then
                return $brand\_KW[kword]$
    return null
```

Algorithm 2 outlines our approach for processing UDRP reports, addressing several issues identified in the reports. First, we gather the **`Registrant Organization`** entries from the collected WHOIS records for the $DS_{refDomains}$. Subsequently, we normalize these entries by solely retaining alphanumeric characters and spaces, while eliminating the values related to privacy protection, such as **`REDACTED FOR PRIVACY`**. We conduct a thorough manual analysis of the acquired list and create a compilation of unique keywords that align with each brand name. Table II shows examples of the mapping of brands to keywords.

We exclusively take into account the UDRP reports that have been fully executed, meaning that their decision date precedes the measurement date by a minimum of 30 days. For each domain listed in UDRP reports, we next try to determine if its corresponding complainant is among our collection of high-profile brands (performed by the **`EXTRACT_BRAND`** function in Algorithm 2).

Following a method similar to that used for **`Registrant Organization`** entries obtained from the WHOIS records, we normalize the complainant value extracted from UDRP reports. We then search for an exact match of the (normalized) complainant value within the list of (normalized) registrant organization values.

Table II
EXAMPLES OF THE KEYWORD-TO-BRAND MAPPING

Name	Domain name	Keyword values

Facebook	**`facebook.com`**	**`facebook`**
		`meta platforms`

PayPal	**`paypal.com`**	**`paypal`**
Fifth Third Bank	**`53.com`**	**`fifth third`**
Absa bank	**`absa.co.za`**	**`absabank`** **`absa`**

If an exact match is not found, we proceed to search for a partial word match using unique keywords associated with each company. When we successfully identify a brand using either of these methods, we extract the creation date from the WHOIS data of the *transferred* disputed domain (if available). We only consider domains whose creation date is before the dispute decision date, which indicates that the domain was registered prior to the decision date and underwent no ownership changes other than the transfer to the complainant.

*2) **SINS**: Shared In-Bailliwick Name Servers:* Companies may use the same authoritative name servers for their primary domains, alternative or localized domains, defensive registrations, and other domains to streamline and simplify their domain management processes. For example, the **`NS`** record of **`example.com`** (the primary domain), **`example.cn`** (a localized domain), and **`exmple.com`** (a defensive registration) for the *Example* company may point to the same set of name servers, **`ns1.example.com`** being one of them for instance.

To add domains to the whitelist for a specific reference domain name based on this assumption, we perform a DNS scan (module ③ in Figure 1) on the DS_{cand} dataset using the Cloudflare public DNS resolver, while adhering to the terms and conditions of the service [45]. We query all domain names for their **`NS`** records and seek the domains with the **`NOERROR`** response code, indicating that a domain name exists in the zone file and is, therefore, registered. The result of the scan is the following: 2.5M domains returned **`NOERROR`** and 522.4k had a non-empty answer.

We also collect the **`NS`** records for the 338 reference domains ($DS_{refDomains}$). However, for this method, we only keep reference domains whose name servers are *in-bailiwick* and *in-domain*, that is, the name server domain name is a subdomain as the origin of the zone [38]. For example, we consider **`google.com`** as *in-baliwick* as its name servers **`ns[1-4].google.com`** are subdomains of the origin **`google.com`**. On the other hand, the **`NS`** records for **`visa.net`** are **`ns[1-3].dnsvisa.com`** indicating *out-of-bailiwick* name servers.

The reason for such strict filtering is that certain high-profile companies outsource the DNS management to DNS providers such as *Cloudflare* or *Akamai*. In these cases, we cannot automatically distinguish between an *out-of-bailiwick* name server whose name was registered by a high-profile company itself and a name server shared by regular customers of DNS providers. Furthermore, attackers could use the same strategy to whitelist a malicious domain by registering a domain and using one of the DNS providers accessible to the public (e.g., Cloudflare).

The last verification step is the resolution of all generated domains for each brand name by their originally collected name servers. All name servers from `NS` records of the reference domain name must return the `NOERROR` response code for the corresponding candidate domain. This step is necessary to prevent attackers from setting false `NS` records for malicious domain names: an attacker could register a malicious domain, for example, `google-support.net`, and set its `NS` record to `ns1.google.com`. However, if such a domain name is not present in the Google zone file, a DNS query for an `A` record to the Google authoritative name servers would result in `REFUSED` or `NXDOMAIN` response codes depending on the name server configuration. We further discuss evasion techniques in Section IV.

3) ***RO****: Registrant Organization:* ICANN requires accurate registrant organization information for domain registration to address legal disputes and prevent ownership issues [46]. However, the General Data Protection Regulation (GDPR) implementation in 2018 limited public access to registration data [47]. Despite this challenge, we have collected registration organization information for 215 out of 338 reference domains. Among them, 37 domains used privacy protection services, while 178 brands included their legal names in the WHOIS data.

We further collect WHOIS data for 2.5M domains that returned a `NOERROR` response (see Section II-B2). We extract the registrant organization information and filter out domains that use privacy protection services. Then, we compare the registrant organization values with those of the original brand to check for exact matches.

Nevertheless, the practical implementation of this approach faces challenges. ICANN requires accredited registrars to verify the contact information such as email or phone numbers (ICANN Registrar Accreditation Agreement [48]). However, the verification of the Registrant Name or Organization is not mandated, enabling the inclusion of false information. To investigate this issue further, we have registered a domain with a false but famous organization name at one of the biggest registrars. The false information was propagated to WHOIS, which confirms the lack of additional verification.

Based on these experiments, we have chosen not to incorporate the identified 8.5k domains into our whitelist due to potential exploitation by attackers. Nevertheless, we make use of the collected WHOIS data for the next method that focuses on identifying defensive registrars (see Section II-B4).

Alternatively, contact details such as mailing addresses, phone numbers, and email addresses, or all registrant organization information, could have been extracted from WHOIS and compared between candidate and reference domain names. However, it is important to note that certain registrars might not consistently verify the authenticity of registrant fields. Furthermore, the process of collecting, parsing, and comparing the registrant organization fields in WHOIS (assuming they are not redacted) can be cumbersome.

Table III
DEFENSIVE REGISTRARS (DR)

Registrar name	IANA ID	name servers of the registrars
MarkMonitor	292	`markmonitor.com`
CSC Corporate Domains	299	`cscdns.{uk,net}`

Com Laude	470	`comlaude.{co.uk,ch,net}`
		`comlaude-dns.{net,eu,co.uk,com}`

RegistrarSEC	2475	`facebook.com`

Safenames	447	`safenames.{co.uk,info,com,net,org}`
		`idp365.net`

Nameshield	1251	`observatoiredesmarques.fr`
		`nameshield.net`
		`ns1.f`

IP Twins	1728	`iptwins.{com,net}`
SafeBrands	1290	`mailclub.{com,eu,fr}`
Hogan Lovells	1526	`lovellsnames.org`

4) ***DR****: Defensive Registrars:* For certain domain names, the set of authoritative name servers may differ from the reference domain name (*out-of-bailiwick* name servers). For instance, if the name servers for the defensively registered domain `instagraam.com` are `[a-d].ns.facebook.com`, the **SINS** method described in Section II-B2 would not categorize it as benign. Therefore, we have designed another method that leverages information extracted from WHOIS data to identify defensive registrations (module ④ in Figure 1).

First, we have identified nine reputable defensive registrars that collaborate with high-profile companies. By examining the top 30 registrars that appeared most frequently in the dataset of disputed domains used in the **UDRP** method (see Section II-B1), we have manually selected nine defensive registrars presented in Table III.

For each of the 2.5 million registered domains, we have retrieved the registrant organization name (as described for the **RO** method in Section II-B3) and registrar information from WHOIS. Domains enter the whitelist if the registrant organization in WHOIS corresponds to one of the original values in our collected list of reference domain names, and the registrar IANA ID and registrar name indicate one of the nine defensive registrars.

However, some TLDs provide limited or no WHOIS information and it is impossible to extract the domain registration information. To tackle this problem, we combine the results from the **SINS** method based on shared name servers described in Section II-B2. The domain names that defensive registrars use to host authoritative name servers for the domains of their customers can be used as an indicator of a defensively registered domain name. For example, the WHOIS data for `docusign.com.bo` does not include information about the registrar and is thus impossible to be marked as a defensive registration but querying for this domain `NS` record reveals

Table IV
VALIDITY PERIODS OF DOMAINS IN THE WHITELIST

Source	Valid from	Valid until
DD	decision date plus 30 days	domain expiration date
SINS	discovery date	domain expiration date
DR	discovery date	domain expiration date
DC	first occurrence in CT Log	$min_{date}\{notAfter, expirationdate\}$

that it is the domain managed by *MarkMonitor* as the origin of all returned name servers (`ns1.markmonitor.com` and `ns3.markmonitor.com`) is `markmonitor.com`. We thus use the mapping of the name servers to their respective defensive registrars shown in Table III and include such identified domains in the whitelist.

*5) **DC**: Domain Certificates:* Transport Layer Security (TLS) certificates are additional sources for domain analysis and potential inclusion in the whitelist. When creating a certificate, two options exist for specifying the domain name. The first option is to include the domain name or a wildcard in the `Common Name` (CN) field of the Subject. The second option is to use the `subjectAltName` extension to specify a domain name or a wildcard. Note that companies have full autonomy for specifying the covered domains and prioritizing their specific needs.

In our method, we leverage the domain information in TLS certificates. We have gathered the certificates for the 338 reference domains with our TLS module (module ⑤ in Figure 1) to obtain a dataset of 6,946 entries, including 6,071 fully qualified domain names (FQDNs) and 875 wildcards. Due to the inability to assess the significance of all domains covered by wildcards, we exclude them from the dataset.

The use of this dataset is the following. We generate a domain list for each brand using the candidate domain list derived from its respective seed. We then cross-reference this list with the domains listed in the certificate to determine their inclusion (e.g., `amazion.com`). The reason for this check is that we have observed cases in which, for example, banks specify domains used at different stages of their service development, such as `www-dev.bank.com` or `www.staging.bank.com`. Some of these domains may host services with unresolved bugs or vulnerabilities, putting them at risk of compromise.

C. Whitelist Validity Period

Whitelists require regular updates to account for domain expiration and potential re-registration by attackers. Each data source in our whitelist has its own method for determining the validity period based on its specific characteristics and procedures.

Table IV shows the validity periods for each proposed method. For **SINS** and **DR**, we consider that a domain name enters the whitelist at the instant of discovery, i.e., the day of the measurement and inclusion in the whitelist. A domain will stay on the list until the domain expiration date. For the domains found in TLS certificates, we consider them as whitelisted from the moment they appeared in the Certificate Transparency Log [49, 50]. These domains remain valid either until the date specified in the `notAfter` field or until the expiration date of the respective domain, depending on which date occurs earlier, which results in different validity periods spanning months to years, depending on the certificate (e.g., paid or free), its issuing authority, or the registration period. The domains extracted from UDRP reports are considered whitelisted thirty days following the decision date, and they will stay on the whitelist until the domain name expiration date.

Such an approach allows us to monitor the expiration date of each domain on the whitelist. Once the expiration date is due for a domain name, we verify the initial reason for whitelisting, i.e., we re-run the corresponding method based on the source of the domain. Consequently, the validity period of the whitelist is the closest to the domain expiration.

D. Two Types of Entries in the Whitelist

Our list of 338 most targeted brands contains the domain names that might provide subdomains that host content not fully under the control of a company, which means that the subdomains may be compromised and they should not be included in the whitelist. For instance, *Free*, a French telecommunication company, offers personal websites for its customers hosted at its subdomain, e.g., `alice.free.fr`, while there are often phishing attacks that impersonate the *Free* homepage containing the login form. Another example is *GitHub* which provides a similar service at `github.io` where people can host the documentation of projects.

On the other hand, the domains of companies providing financial services do not provide such services for security reasons and they only have a few subdomains. In this case, we can safely include the domains in the whitelist.

To distinguish between these two cases, our whitelist consists of two types of entries: a) exact FQDN match and b) wildcards. The first one corresponds to the domains for which the subdomains may be malicious, i.e., they may provide services prone to be used by attackers to host phishing web pages (e.g., a web page hosted at `github.io`). The latter signifies that a domain and its subdomains are safe and whitelisted.

We manually investigated each of the 338 brands and their websites to determine the type of entry for their original domains as well as for the domains generated from the domain name. For example, `google.com` and its defensive registration `gogle.com` appear in the list as type a) while the domains related to *Amazon* appear with type b), i.e., `*.amazon.com`, `*.amazonpphp.com`, etc.

III. EVALUATION

In this section, we evaluate our proposed scheme by analyzing how each method contributes to the final whitelist. We compare the constructed whitelist to different existing systems by analyzing the overlap with ranking lists and by applying the constructed whitelist to various blocklists retrospectively.

A. Composition of the Whitelist

In total, we have included 17,215 domains in the final whitelist created on September 1, 2023. Figure 2 presents its composition and the overlap of the domains coming from different sources and included in the whitelist by different methods.

We can observe that disputed domains included by the **DD** method contribute the most to the final whitelist with 7,405 (43.9%) domains uniquely discovered by this method. This result shows that UDRP reports are a valuable source of information for creating domain whitelists and could be considered by the cybersecurity and research community for creating datasets of benign domains. The **DR** method based on defensively registered domains is the second most significant contributor with 23.6% (3,979) of domains not generated by other sources. The domains whose name servers share the same origin as their seed domain or name servers used for defensive registrations (**SINS** method) account for 9.4% of domains in the whitelist. Note that the significant overlap (19.3%) between the list of defensively registered domains and domains found based on the name server match is due to the fact that defensive registrations are often registered on behalf of the requestor (trademark owner or company), but they can be delegated to one of the defensive registrar name servers.

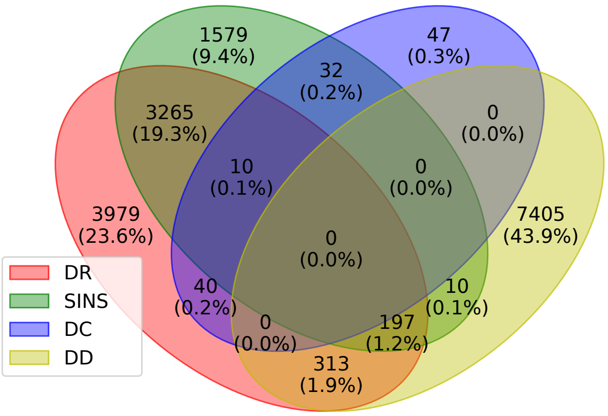

Figure 2. Overlap of whitelisted domains coming from different sources

In the next step, we have analyzed to what extent different brands contribute to our whitelist. The domains generated for *Amazon*, *Google*, *Microsoft*, *Instagram*, and *Facebook* together account for 32.3% of the whole whitelist with 1,377 (8%), 1,190 (6.9%), 1,105 (6.4%), 981 (5.7%), and 911 (5.3%) domains, respectively (see Figure 3).

Table V shows the top 10 brands for which their corresponding domains cover the highest number of TLDs. The domain names related to *Amazon* cover 372 TLDs (23% of all TLDs) and were registered by four different defensive registrars: *Com Laude*, *CSC Corporate Domains*, *Hogan Lovells*, and *MarkMonitor*. Altogether, *Amazon* covered all types of cybersquatting, i.e., change in TLD, typosquatting, combosquatting, bitsquatting, homoglyphs, homophones, and other types of squatting candidates such as hyphenation or transposition.

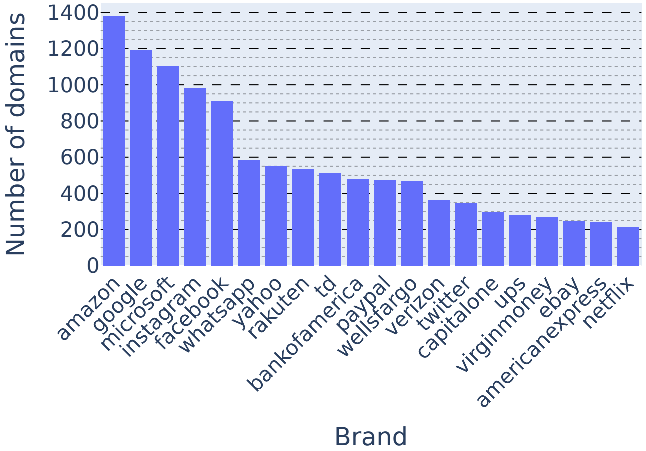

Figure 3. Top 20 brands with the most important contribution to the whitelist

Table V
TOP 10 BRANDS WITH THE NUMBER OF COVERED TLDS AND CORRESPONDING DEFENSIVE REGISTRARS

brand	# of TLDs	Defensive registrars	
Amazon	372	**Com Laude** **CSC**	**Hogan Lovells** **MarkMonitor**
Instagram	273	**Com Laude** **Hogan Lovells**	**MarkMonitor** **RegistrarSEC**
Microsoft	264	**Com Laude** **CSC**	**MarkMonitor**
Facebook	251	**Com Laude** **Hogan Lovells**	**MarkMonitor** **RegistrarSEC**
Twitter	194	**CSC**	**MarkMonitor**
Rakuten	180	**Com Laude** **MarkMonitor**	**Safebrands**
Google	175	**MarkMonitor**	
Paypal	175	**MarkMonitor**	
Ebay	150	**MarkMonitor**	
Vodafone	106	**Com Laude**	**CSC**

Table VI
CONFIRMED FALSE POSITIVES IN EXISTING BLOCKLISTS

blocklist	# of unique domains	# of reported URLs
APWG	73	1,154,248
OpenPhish	15	25
Phisthank	5	12
Google Safe Browsing	21	–

B. Comparison with Existing Systems

To evaluate how our proposed whitelist contributes to enhancing precision and reducing the occurrence of false positives, we compared it with well-established systems used by both industry and the research community.

1) Application to Blocklists: We have applied the whitelist to existing blocklists to demonstrate its effectiveness. We have gathered URLs and domains that appeared between August 2022 and August 2023 in three phishing blocklists: APWG, OpenPhish, and PhishTank, and selected the URLs for which their FQDN or the registered domain (depending on the type of whitelist entry as described in Section II-D) appeared on our whitelist and were blocklisted. Table VI presents the results: the number of confirmed false positives in the existing blocklists. We have identified 73 unique wrongly classified domain names that appeared in the APWG blocklist containing 1.1 million URLs. This important number of URLs comes from a small number of domains that generate a unique URL when the APWG automated system visits a given page. An example of such a domain is **`absabank.mu`** with over a million blacklisted URLs. The majority of such URLs are similar to the following example:

`https://online.absabank.mu/air/feature/loginprocess?execution=e1s1&_t=1674345386992`

This behavior has been observed for phishing web pages that try to avoid detection systems by creating a unique URL on each victim visit [51, 52]. Such behavior can also be observed on the official login pages of some financial institutions which, in combination with a benign domain wrongly classified as malicious and with an automated approach, results in an important number of false positive URLs. The number of false positives for other blocklists is less significant with 15 wrongly classified domains for 25 URLs in OpenPhish, and 12 URLs containing 5 unique domains for PhishTank including domains like **`walletconnect.com`**, **`bancobpm.it`**, or **`orange.fr`**.

We applied Google Safe Browsing to our whitelist, which resulted in 21 domains marked as a threat. We manually investigated these domains and found that 19 out of 21 domains went through a UDRP dispute process. However, all of them were already transferred to the complainant at the time of the measurement, indicating a false positive in GSB. For example, a UDRP complaint filed at WIPO by *Twitter, Inc.* contained 62 domains out of which 6, e.g., **`twitter-warning.com`**, **`twitter-safety.com`**, or **`twitter-processing.com`**, were labeled as malicious by GSB. Such cases reveal a common problem of blacklists when a domain name is correctly identified as malicious at the time of blacklisting but remains on the blacklist even after the domain does not represent a threat anymore, i.e., is transferred to the trademark owner as in the reported cases. The remaining 2 domains, **`restorebankofamerica.com`** and **`wëllsfargo.com`**, were defensive registrations of *Bank of America* and *Wells Fargo* at the time of measurement.

2) Comparison with Ranking Lists: The cybersecurity community tends to use popularity ranking lists to either create their "ground truth" datasets of benign domains [53], the lists of popular domains to exclude from training when developing domain classification systems, or the exclusion lists used as a step in pipelines in the blocklist construction [25, 26]. Our list complements these approaches with a reliable list of the domains that are either i) directly the target of phishing attacks or ii) defensively registered by trademark owners and thus will not be involved in phishing attacks or other malicious activities.

Out of the 17,215 whitelisted domains, only 627 (3.6%) domains have a Tranco rank, i.e., appeared in the Tranco 1M list as of September 24, 2023. Eighteen original reference domains do not have a rank in Tranco while being regularly targeted by phishing attacks. For example, the official domain of *Absa Bank (Mauritius) Limited*, **`absabank.mu`**, does not appear in the Tranco list but was found as a false positive in the previous section with over one million reported URLs. Our whitelist can thus be used to complement existing popular ranking lists when creating exclusion lists while building phishing detection systems.

Table VII
TIMELINE OF THE **`VERIFYISSUE-META.CLICK`** DOMAIN

#	Event	Date
1	Domain registration	**`05-05-2022`**
2	UDRP complaint filed by *Meta Platforms*	**`28-10-2022`**
3	Domain appeared in phishing blocklists	**`11-11-2022`**
4	UDRP decision to transfer to *Meta Platforms*	**`24-01-2023`**
5	Domain included in the whitelist	**`01-09-2023`**
6	Expiration date (as of the time of writing)	**`05-05-2025`**

C. Case Studies

In this section, we present two case studies of patterns related to the benignness of domains that appear in blocklists.

1) Transferred Blocklisted Domains: We have observed that some domains whitelisted by the **DD** method did appear in blocklists. After manual investigation, we have found that some of these domains are not false positives (we did not count them as false positives in Table VI), and they were involved in phishing attacks at the time of blocklisting. Later on, their owners discovered such cases either through monitoring of blocklists or by internal processes for active discovery of cybersquatting and started a UDRP process to obtain such domains from the original registrant. Once the domain transfer (if applicable) is implemented, such domains become benign and can thus be included in our whitelist.

Table VII gives an example of such a domain: **`verifyissue-meta.click`** was registered on May 5, 2022, impersonating *Meta Platforms*, and appeared in APWG and OpenPhish six months later. On October 28, 2022, *Meta Platforms* filed a complaint with the WIPO Arbitration and Mediation Center and the assigned panelists confirmed that the original registrant used the domain in an improper manner, and decided to transfer the ownership of the disputed domain to the complainant, i.e., *Meta Platforms*, on January 24, 2023. A domain becomes a valid entry in our whitelist from the moment it is transferred to *Meta Platforms* as a defensive registration and remains valid until its expiration date.

2) ICANN New gTLDs Programs: Registries regularly introduce new top-level domain name programs for public registrations. During the so-called 'Sunrise Period' spanning a minimum of 30 days, trademark owners are granted a prior opportunity to register domain names that match their trademarks before these names become accessible to the broader public [54]. In April 2023, the *Google* Registry launched 8 new gTLDs: **`.zip`**, **`.mov`**, **`.prof`**, **`.nexus`**, **`.esq`**, **`.dad`**,

.foo, and **.phd** with the 'Sunrise Period' starting on April 2, 2023. *Com Laude*, one of the defensive registrars, registered domains on behalf of *Rakuten* with all 8 TLDs that became blocklisted by APWG two days later. Such domains cannot by no means be marked as phishing URLs and should not appear in the blocklist. With the 'Sunrise Period' of **.box** active from August 12 to September 12, 2023, and three new gTLDs to be launched in September 2023, these cases may appear on a regular basis [55]. Our whitelist identifies such domains as benign and, if applied in the automated pipelines of systems for detecting phishing URLs, it helps prevent such errors due to its regular updates.

IV. Discussion

In this section, we engage in an in-depth discussion of various pivotal aspects related to our proposed scheme. We address the conservative approach we adopted, the challenges associated with domain registration data access, and the discovery of defensively registered but non-resolving domains.

A. Conservative Approach

One of the design goals of the proposed scheme was the resilience to possible evasion by attackers while keeping coverage as high as possible. Having this objective in mind, we have chosen a conservative approach to the inclusion of domains in our whitelist even at the cost of a shorter list. Given that blocklist providers can use the proposed whitelist, it is important to keep a close-to-zero false negative rate. As no ground truth data exists for this type of whitelist, we have deliberately eliminated some sources despite their promising results due to the risk of introducing false negatives. For example, the condition of strict *in-bailiwick* and *in-domain* match for the SINS method reduces the number of included domains by almost 40% because the companies such as *Microsoft* set the **NS** record for their main domains to *out-of-bailiwick* name servers (the registered domain of **NS** for **microsoft.com** is **azure-dns.com**). Future work on the construction methods may consider including such cases in our whitelist without risking wrong decisions.

B. Domain Registration Data

Access to domain registration information has been a challenge. The format of the WHOIS information is inconsistent across different registries and registrars. ICANN requires that accredited registrars of new gTLDs follow a defined format [48]. However, ccTLDs and non-accredited registrars are not required to follow these guidelines and often, they implement their own formats, more or less complete and sometimes missing important information like the creation or expiration date. Furthermore, parsing such information is tedious and requires implementing specific parsers or retrieving already parsed data from paywalled third parties [56]. The Registration Data Access Protocol (RDAP) provides responses to queries in a standardized JSON format for registration data, accessible through HTTP. However, at the time of writing, only 27 (9%) of 308 active ccTLDs have implemented an RDAP service [57].

Another challenge for the security community is related to the terms and conditions of the WHOIS and RDAP usage that often prohibits automated processing of registration information imposing rate limits. At the same time, the data from WHOIS or RDAP is often crucial to building systems similar to ours. We use the best-effort strategy to acquire as much data as possible but we, as well as other researchers, investigators, and other involved parties, would benefit from less restricted access to the registration information, and the proposed whitelist would gain in precision and coverage.

C. Registered Non-Resolving Domains

We have found some defensively registered domains but not discovered by any of the proposed methods. A manual investigation has revealed that some domains are registered, i.e., their WHOIS data is accessible but defensive registrars disable their delegation making them inaccessible in public DNS (i.e., the query response is **NXDOMAIN**).

In order to evaluate the effectiveness of our proposed methods, we delved into these cases to determine their prevalence. *Google* delegates its defensive registrations to *MarkMonitor*. Given that the previously observed case of a defensively registered but not delegated domain name happened to be done by this registrar, we have listed all candidate domain names from the previously presented dataset related to *Google*, collected the WHOIS data, and performed active DNS measurements querying the **NS** record. In total, we found 701 domains defensively registered by MarkMonitor. 68 (9.7%) domains returned **NXDOMAIN** out of which 56 (82%) domains were either put on hold by the registry (the Extensible Provisioning Protocol **serverHold** status) or were inactive with no delegation [58]. Four domains did not contain any status that would indicate a problem in delegation and for the remaining eight domains, we could not collect the EPP status code as they were registered at ccTLDs that do not include this information in WHOIS.

The solution to this limitation would be to collect WHOIS data systematically for all candidate domains. However, as mentioned previously, the acquisition of WHOIS data for a large number of domains is difficult due to the terms and conditions of the registrar, rate limits, and incompleteness of data due to GDPR.

V. Related Work

In this section, we examine three categories of related work: typosquatting along with other forms of domain squatting, defensive registrations, and the construction of whitelists.

Based on the recurrent appearance of defensive registrations in prior work, we chose to follow successfully used approaches to generate candidate domains. While prior work focused on identifying and analyzing potential cybersquatting namespace, we identify the domains that appear as squatting ones but are under the control of the trademark owners in the form of *defensive registrations*.

A. Typosquatting

Wang et al. [35] defined five models for generating typos and proposed a system to identify typo domains for a limited number of domains from the Alexa list. Their research revealed that most typo domains were either parked or redirected to parking services. Holgers et al. [34] defined a *homograph attack* and showed that confusable domains often have only one character replaced by their Unicode confusable character.

In 2010, Moore et al. [59] used the Damerau-Levenshtein distance [60] to identify squatted domains and found that hundreds of thousands of typo domains were registered under `.com`, the majority concentrated among a few large registrars and advertisement platforms. They also identified 4,133 defensively registered typo domains that shared the same name servers as the domains from which they were generated.

Nikiforakis et al. [61] studied the phenomenon of *bitsquatting* [62] for the Alexa top 500 domains and showed that it is used as the known way for benefitting from domain squatting with parking services, malware distribution, or affiliate abuse. They claimed that 3.9% of the 5,366 identified bitsquatting domains were defensively registered.

Szurdi et al. [33] proposed the *YATT* (Yet Another Typosquatting Tool) framework and estimated that approximately 20% of all `.com` domains were typo domains. They categorized defensive registrations as typo domains directly or indirectly redirected back to the original domain, excluding them from the final list of typosquatting domains. Halvorson et al. [63] analyzed the `.xxx` new gTLD namespace and estimated that 92% of registrations were made for defensive purposes while only 5.9% were registered to serve content.

In 2015, Agten et al. conducted a seventh-month-long content-based study investigating the longitudinally of typosquatting namespace targeting the top 500 popular domains from the Alexa rating list. They found that 477 out of those 500 domains have at least one typosquatting domain while only 156 domains (32.2%) used defensive registration as a protective measure against abuse. Interestingly, some of the domains that use this strategy chose not to renew their defensive registrations. Kintis et al. [21] studied the prevalence of combosquatting, i.e., abusive domain registration that combines a trademark with one or more phrases.

Liu et al. [64] gave insight into the usage of Internationalized Domain Names (IDNs) based on visual similarity and Quinkeer et al. [65] assessed the scale of homograph squatting domains. Le Pochat et al. [66] focused on typosquatting related to international keyboard layout and found that companies have paid more attention to defensively registering such domains but failed at covering all possibilities.

B. Defensive Registrations

Only a little research concerns the defensive registration ecosystem. Quinkert et al. [65] found that 8% of the 2,895 candidate domain names were registered for defensive purposes by 23 distinct brands. They identified defensive registration either by extracting an e-mail address from WHOIS information or name servers using public DNS and determining if they belong to the original (reference) brand. However, such an approach suffers from several flaws. First, WHOIS information often contains fields redacted for privacy due to GDRP. The values of the registrant contact email address are replaced by URLs leading to the registrar's website with a contact form. Second, shared name servers may indicate that a domain name relates to the original brand name, however, without a second necessary step of validation if such a domain exists in the company zone file, querying these name servers may lead to incorrect classification as described in Section II-B2. Moreover, the problem of shared third-party DNS providers should be carefully considered.

Maroofi et al. [67] listed over 55 thousand defensive registrations to study the deployment of the Sender Policy Framework (SPF) [68] and Domain-based Message Authentication, Reporting, and Conformance (DMARC) [69] for the top 500 domains from the Alexa list. They considered a domain as defensively registered if it shares the same `A` or `NS` record as one of the reference domains or if the domain homepage URL is the same as one of the reference domains. We argue that such a classification may lead to errors as an attacker can temporarily set the `A` or `NS` records to point to one of the reference domains. Moreover, recent research showed that phishing web pages can redirect to a benign website in case of cloaking [70], which may lead to a malicious domain being considered as defensively registered.

C. Whitelists

The creation and maintenance of whitelists have received only limited attention from the research community. Burton et al. [28] proposed a method for the automated creation of whitelists using Bayesian statistical learning applied to popularity ranking lists such as the Alexa Top 1M domain list or the Majestic Million.

Known blacklists such as SURBL [71] or Spamhaus [72] follow the structure of a DNS blacklist (DNSBL) introduced in the RFC 5782 [73]. Besides blocklists, RFC 5782 establishes a standard structure for whitelists, however, it does not provide any specific policy for the content of a whitelist.

VI. Ethical Considerations

Our research encompasses large-scale network measurements and adheres to community best practices [74–76], which involves extensive collection of DNS and WHOIS data. To ensure efficiency and ethical conduct, we have randomized our input list for DNS scans, distributing the workload among authoritative name servers (TLDs). Moreover, for DNS measurements, we have leveraged the infrastructure of the Cloudflare public resolver, known for its capacity to handle high query rates [77]. Finally, we have conducted our scans over multiple days to further distribute the workload.

In our WHOIS measurement process, we initially used DNS scans to filter domains and focused on collecting the WHOIS data for active domain names. In particular, we have adhered to the limits of queries to various WHOIS servers specified in operator specifications or identified empirically.

Throughout the scanning period, we maintained a website on our IP address with the information about the study and an opt-out mechanism but we did not receive any complaints.

VII. Conclusion

Effective construction of domain name whitelists is important for reducing the volume of all domain names to analyze when looking for maliciously registered or compromised domain names and for limiting the number of false positives in blocklist feeds.

In this paper, we have proposed a scheme for generating a resilient domain whitelist based on four types of domain name sources: public DNS, domain registration data (WHOIS), TLS certificates, and data from UDRP dispute-resolution service providers. In the construction of the whitelist, we first extract reference domains, a selection of the brand names most targeted by phishing, and generate their variants by applying typo-squatting, bitsquatting, combo-squatting, and finding homographs and homophones. We also identify the suffixes that generate a **NOERROR** response code for non-existent domains.

Based on this dataset of domain names, we query them for their **NS** records and seek the domains with the **NOERROR** response code. We only keep the domains whose name servers are *in-bailiwick* and *in-domain*. We then collect their WHOIS data and extract the registrant organization to filter out the domains that use one of the privacy protection services and compare the collected values of the registrant organization with the values of the original brand. Other sources for the whitelist are dispute-resolution service providers that give us the companies involved in domain dispute processes considered high-profile. We extract a set of unique disputed domains transferred to the complainant and for which we could collect the decision date. We have also identified several well-known defensive registrars that cooperate with high-profile companies and label as whitelisted the domains registered with one of the nine defensive registrars. Finally, we leverage the domains that appear in TLS certificates of the most targeted brand names.

To evaluate the constructed whitelist, we have analyzed the proportion of domains entered into the whitelist with a given proposed method: most of the whitelisted domains are registered with one of the nine defensive registrars and discovered by *in-bailiwick* and *in-domain* checks. We have also applied the whitelist to historical data and revealed several confirmed false positives in existing blocklists.

References

[1] Anti-Phishing Working Group, “Global phishing survey: Trends and domain name use in 2016,” https://docs.apwg.org/reports/APWG_Global_Phishing_Report_2015-2016.pdf, 2016.

[2] OpenPhish, https://openphish.com/, 2023.

[3] D. Ulevitch, “PhishTank,” https://phishtank.org/, 2006.

[4] Google, “Google Safe Browsing,” https://safebrowsing.google.com/, 2023.

[5] Microsoft, “Microsoft Defender SmartScreen,” https://learn.microsoft.com/en-us/windows/security/threat-protection/microsoft-defender-smartscreen/microsoft-defender-smartscreen-overview, 2023.

[6] Zonefiles.io, https://zonefiles.io, 2023.

[7] Anti-Phishing Working Group, “Phishing Activity Trends Reports,” https://apwg.org/trendsreports/, 2023.

[8] A. Blum, B. Wardman, T. Solorio, and G. Warner, “Lexical feature based phishing url detection using online learning,” in *AISEC*, 2010, pp. 54–60.

[9] C. Whittaker, B. Ryner, and M. Nazif, “Large-scale automatic classification of phishing pages,” in *NDSS*, 2010.

[10] J.-H. Li and S.-D. Wang, “PhishBox: An approach for phishing validation and detection,” in *IEEE DASC/PiCom/DataCom/CyberSciTech*, 2017, pp. 557–564.

[11] K. Tian, S. T. K. Jan, H. Hu, D. Yao, and G. Wang, “Needle in a haystack: Tracking down elite phishing domains in the wild,” in *IMC*, 2018.

[12] G. Xiang, J. Hong, C. P. Rose, and L. Cranor, “Cantina+: A feature-rich machine learning framework for detecting phishing web sites,” *ACM TISSEC*, 2011.

[13] B. Kondracki, B. A. Azad, O. Starov, and N. Nikiforakis, “Catching transparent phish: Analyzing and detecting mitm phishing toolkits,” in *CCS*, 2021, p. 36–50.

[14] R. Liu, Y. Lin, X. Yang, S. H. Ng, D. M. Divakaran, and J. S. Dong, “Inferring phishing intention via webpage appearance and dynamics: A deep vision based approach,” in *USENIX Security*, 2022, pp. 1633–1650.

[15] G. Ho, A. Cidon, L. Gavish, M. Schweighauser, V. Paxson, S. Savage, G. M. Voelker, and D. Wagner, “Detecting and characterizing lateral phishing at scale,” in *USENIX Security*, 2019, pp. 1273–1290.

[16] S. Marchal, K. Saari, N. Singh, and N. Asokan, “Know your phish: Novel techniques for detecting phishing sites and their targets,” in *IEEE ICDCS*, 2016.

[17] R. S. Rao and A. R. Pais, “Detection of phishing websites using an efficient feature-based machine learning framework,” *Neural Computing and Applications*, vol. 31, pp. 3851–3873, 2019.

[18] I. Corona, B. Biggio, M. Contini, L. Piras, R. Corda, M. Mereu, G. Mureddu, D. Ariu, and F. Roli, “Deltaphish: Detecting phishing webpages in compromised websites,” in *ESORICS*, 2017, pp. 370–388.

[19] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster, “Building a Dynamic Reputation System for DNS,” in *USENIX Security*, 2010.

[20] V. G. Li, M. Dunn, P. Pearce, D. McCoy, G. M. Voelker, and S. Savage, “Reading the tea leaves: A comparative analysis of threat intelligence,” in *USENIX Security*, 2019, pp. 851–867.

[21] P. Kintis, N. Miramirkhani, C. Lever, Y. Chen, R. Romero-Gómez, N. Pitropakis, N. Nikiforakis, and M. Antonakakis, “Hiding in plain sight: A longitudinal study of combosquatting abuse,” in *CCS*, 2017.

[22] Sean Gallagher, “New “Quad9” DNS service blocks malicious domains for everyone,” https://arstechnica.

com/information-technology/2017/11/new-quad9-dns-service-blocks-malicious-domains-for-everyone/, Nov. 2017.

[23] H. Tupsamudre, S. Jain, and S. Lodha, "Phishmatch: A layered approach for effective detection of phishing urls," *arXiv preprint arXiv:2112.02226*, 2021.

[24] Amazon, "Alexa," https://www.alexa.com/, 2022.

[25] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen, "Tranco: A research-oriented top sites ranking hardened against manipulation," in *NDSS*, 2019.

[26] R. Hüssy, "URLhaus Malware URL exchange," https://urlhaus.abuse.ch/, 2018.

[27] V. L. Pochat, T. V. Goethem, and W. Joosen, "Evaluating the long-term effects of parameters on the characteristics of the tranco top sites ranking," in *USENIX CSET*, 2019.

[28] R. Burton and L. Rocha, "Whitelists that work: Creating defensible dynamic whitelists with statistical learning," in *APWG eCrime*, 2019, pp. 1–10.

[29] ICANN, "Uniform Domain Name Dispute Resolution Policy," https://www.icann.org/resources/pages/policy-2012-02-25-en, 1999.

[30] Mozilla, "List of Top-Level Domains," https://publicsuffix.org/list/, 2023.

[31] J. Bayer, B. C. Benjamin, S. Maroofi, T. Wabeke, C. Hesselman, A. Duda, and M. Korczyński, "Operational domain name classification: From automatic ground truth generation to adaptation to missing values," in *PAM*, 2023, p. 564–591.

[32] P. Agten, W. Joosen, F. Piessens, and N. Nikiforakis, "Seven months' worth of mistakes: A longitudinal study of typosquatting abuse," in *NDSS*, 2015.

[33] J. Szurdi, B. Kocso, G. Cseh, J. Spring, M. Felegyhazi, and C. Kanich, "The long "Taile" of typosquatting domain names," in *USENIX Security*, 2014, pp. 191–206.

[34] T. Holgers, D. E. Watson, and S. D. Gribble, "Cutting through the confusion: A measurement study of homograph attacks," in *USENIX ATC*, 2006.

[35] Y.-M. Wang, D. Beck, J. Wang, C. Verbowski, and B. Daniels, "Strider Typo-Patrol: Discovery and analysis of systematic Typo-Squatting," in *SRUTI*, 2006.

[36] Anti-Phishing Working Group, "Phishing Activity Trends Report: 3Q 2022," https://docs.apwg.org/reports/apwg_trends_report_q3_2022.pdf, 2022.

[37] Marcin Ulikowski, "dnstwist," https://github.com/elceef/dnstwist, 2023.

[38] P. E. Hoffman, A. Sullivan, and K. Fujiwara, "DNS Terminology," RFC 8499, 2019.

[39] ACPA, "Anticybersquatting consumer protection act," https://www.govinfo.gov/content/pkg/PLAW-106publ113/html/PLAW-106publ113.htm, 1999.

[40] WIPO, "Schedule of Fees under the UDRP," https://www.wipo.int/amc/en/domains/fees/, 2002.

[41] World Intellectual Property Organization, "WIPO Domain Name Decisions," https://www.wipo.int/amc/en/domains/decisionsx/, 2023.

[42] ADR Forum, "Domain Name Dispute Proceedings & Decisions," https://www.adrforum.com/domain-dispute/search-decisions, 2023.

[43] Asian Domain Name Dispute Resolution Center, "UDRP decisions," https://www.adndrc.org/decisions/udrp, 2023.

[44] Canadian International Internet Dispute Resolution Center, "CIIDRC Decisions," https://ciidrc.org/domain-name-disputes/ciidrc-decisions/, 2023.

[45] Cloudflare, "Cloudflare 1.1.1.1," https://developers.cloudflare.com/1.1.1.1/, 2023.

[46] C. Alvarez, "Good Practices for the Registration and Administration of Domain Name Portfolios (Part II)," https://www.icann.org/en/blogs/details/good-practices-for-the-registration-and-administration-of-domain-name-portfolios-part-ii-23-6-2017-en, 2017.

[47] C. Lu, B. Liu, Y. Zhang, Z. Li, F. Zhang, H. Duan, Y. Liu, J. Q. Chen, J. Liang, Z. Zhang *et al.*, "From WHOIS to WHOWAS: A large-scale measurement study of domain registration privacy under the GDPR." in *NDSS*, 2021.

[48] ICANN, "Registrar Accreditation Agreement," https://www.icann.org/resources/pages/approved-with-specs-2013-09-17-en, 2013.

[49] Google, "Certificate Transparency," https://certificate.transparency.dev/, 2023.

[50] Sectigo, "Crt.sh," https://crt.sh/, 2023.

[51] KOR Labs, "DNS abuse institute intelligence platform: Methodology," https://dnsabuseinstitute.org/wp-content/uploads/2022/10/DNSAI-Compass-Methodology.pdf, 2022.

[52] J. Bayer, Y. Nosyk, O. Hureau, S. Fernandez, S. Paulovics, A. Duda, and M. Korczynski, *Study on Domain Name System (DNS) abuse – Technical report. Appendix 1.* Publications Office of the European Union, 2022.

[53] D. Hasselquist, E. K. Gawell, A. Karlström, and N. Carlsson, "Phishing in style: Characterizing phishing websites in the wild," in *TMA*, 2023, pp. 1–4.

[54] ICANN, "Trademark Clearinghouse Rights Protection Mechanism Requirements," https://newgtlds.icann.org/sites/default/files/rpm-requirements-14may14-en.pdf, 2023.

[55] ——, "Sunrise Calendar," https://www.trademark-clearinghouse.com/gtld-calendar, 2023.

[56] S. Liu, I. Foster, S. Savage, G. M. Voelker, and L. K. Saul, "Who is .com? learning to parse whois records," in *IMC*, 2015, p. 369–380.

[57] ICANN, "Bootstrap Service Registry for Domain Name Space," https://www.iana.org/assignments/rdap-dns/rdap-dns.xhtml, 2023.

[58] S. Hollenbeck, "Extensible Provisioning Protocol (EPP) Domain Name Mapping," RFC 5731, 2009.

[59] T. Moore and B. Edelman, "Measuring the perpetrators and funders of typosquatting," in *Financial Cryptography and Data Security*, 2010, pp. 175–191.

[60] F. J. Damerau, "A technique for computer detection and correction of spelling errors," *Commun. ACM*, p.

171–176, 1964.

[61] N. Nikiforakis, S. Van Acker, W. Meert, L. Desmet, F. Piessens, and W. Joosen, “Bitsquatting: Exploiting bitflips for fun, or profit?” in *WWW*, 2013, p. 989–998.

[62] A. Dinaburg, “Bitsquatting: DNS Hijacking without exploitation,” in *BlackHat Security*, 2011.

[63] T. Halvorson, K. Levchenko, S. Savage, and G. M. Voelker, “Xxxtortion? inferring registration intent in the .xxx tld,” in *WWW*, 2014, p. 901–912.

[64] B. Liu, C. Lu, Z. Li, Y. Liu, H. Duan, S. Hao, and Z. Zhang, “A reexamination of internationalized domain names: The good, the bad and the ugly,” in *IEEE/IFIP DSN*, 2018, pp. 654–665.

[65] F. Quinkert, T. Lauinger, W. Robertson, E. Kirda, and T. Holz, “It's not what it looks like: Measuring attacks and defensive registrations of homograph domains,” in *IEEE CNS*. IEEE, 2019.

[66] V. Le Pochat, T. Van Goethem, and W. Joosen, “A smörgåsbord of typos: Exploring international keyboard layout typosquatting,” in *IEEE EuroS&PW Workshops*, 2019, pp. 187–192.

[67] S. Maroofi, M. Korczyński, and A. Duda, “From Defensive Registration to Subdomain Protection: Evaluation of Email Anti-Spoofing Schemes for High-Profile Domains,” in *TMA*, 2020.

[68] S. Kitterman, “Sender Policy Framework (SPF) for Authorizing Use of Domains in Email, Version 1,” RFC 7208, 2014.

[69] M. Kucherawy and E. Zwicky, “Domain-based Message Authentication, Reporting, and Conformance (DMARC),” RFC 7489, 2015.

[70] P. Zhang, A. Oest, H. Cho, Z. Sun, R. Johnson, B. Wardman, S. Sarker, A. Kapravelos, T. Bao, R. Wang, Y. Shoshitaishvili, A. Doupé, and G.-J. Ahn, “Crawlphish: Large-scale analysis of client-side cloaking techniques in phishing,” in *IEEE S&P*, 2021, pp. 1109–1124.

[71] SURBL, https://surbl.org/.

[72] Spamhaus, http://www.spamhaus.org/.

[73] J. R. Levine, “DNS Blacklists and Whitelists,” RFC 5782, 2010.

[74] D. Dittrich and E. Kenneally, “The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research,” https://catalog.caida.org/paper/2012_menlo_report_actual_formatted, 2012.

[75] C. Partridge and M. Allman, “Ethical considerations in network measurement papers,” *Commun. ACM*, vol. 59, no. 10, p. 58–64, sep 2016.

[76] Z. Durumeric, E. Wustrow, and J. A. Halderman, “ZMap: Fast Internet-Wide Scanning and Its Security Applications,” in *USENIX Security*, 2013, pp. 605–620.

[77] Cloudflare, “1.1.1.1 - Rate Limiting,” https://developers.cloudflare.com/1.1.1.1/infrastructure/network-operators/, 2023.