NetBeacon Measurement and Analytics Platform (MAP): Methodology

Sourena Maroofi, Maciej Korczyński
KOR Labs
contact@korlabs.io

1 Data Collection and Processing

1.1 URL Blocklists

We initially selected phishing and malware delivery abuse types because they generally provide sufficient verifiable evidence of the security threat. The availability of verifiable evidence is typically not the case for other types of abuse, such as spam or botnet command-and-control domain names [1]. To measure the prevalence (i.e., DNS Abuse rate) and persistence (i.e., uptime) of abusive domain names involved in phishing and malware delivery, we use four reputable URL blocklists provided to us by the Anti-Phishing Working Group (APWG),[1] PhishTank,[2] OpenPhish[3] and ABUSE.ch (URLhaus feed[4]). We may include more data sources in the future. The selected providers supply URLs in near real time via APIs. How often we download them depends on how often the feed is updated or on restrictions imposed by their providers.

- **APWG** provides phishing URLs submitted by accredited users via the eCrime Exchange (eCX) platform.[5] We download the abusive URLs every minute.
- **PhishTank** feed is a community phishing verification system, which contains phishing URLs submitted and verified by its contributors as abusive. We gather abusive URLs every one hour.
- **OpenPhish** dataset publishes URLs identified by or reported to OpenPhish and verified as phishing. We use the premium feed to download malicious URLs every five minutes.

[1] `http://antiphishing.org`
[2] `http://www.phishtank.com`
[3] `https://openphish.com`
[4] `https://urlhaus.abuse.ch`
[5] `https://apwg.org/ecx/`

- **URLHaus** is a community service operated by `abuse.ch` that provides URLs (containing either domains or IP addresses) used for malware delivery. We download the malware delivery URLs every five minutes.

Note that no known blocklists are free of false positives, i.e., legitimate URLs incorrectly flagged as malicious. However, the here-proposed method is designed to reduce the impact of false positives on the uptime metrics (cf. Section 1.4).

From the obtained blocklists, we exclude all URLs containing IP addresses rather than domain names (e.g., `hxxp://59.92.45.214:49492/Mozi.m`[6]). Using the "ICANN domains" section of the Public Suffix List maintained by Mozilla,[7] we extract registered domain names, i.e., second-level domain names and higher-level domains if a given registry provides such registrations, e.g., `example.co.uk`. Note that all the URL feeds that are being used in this report comprise maliciously registered domains, compromised domains (benign domain names that have been compromised at the website, hosting, or DNS level), and special domain names. We define a special domain as a domain name that provides subdomains or a redirection that can be abused by attackers, but the original purpose of the registered domain name is legitimate. Those domain names are generally registered by operators of URL shorteners (e.g., `bitly.com`) or subdomain providers. For example, dynamic DNS providers (e.g., `duckdns.org`), free subdomain providers (e.g., `000webhost.com`), or file sharing services (e.g., `docs.google.com`). We maintain and manually update a list of special domains and make them available to the research community[8] [9]. We keep only domain names likely to be registered by end users and exclude special domain names, to avoid, for example, `google.com` being flagged as abusive.

1.2 Domain Names

In order to estimate the size (i.e., domains under management) and the number of newly registered domain names monthly per registrar, we first collect the list of domain names for each Top-Level Domain (TLD). We process zone files obtained from the ICANN Centralized Zone Data Service (CZDS)[10] provided by participating generic TLDs (gTLD) that accepted our request access. We also process zone files of some country-code TLDs, e.g., publicly accessible zones of .se, .nu[11], .li, .ch[12] TLDs. We also plan to include .uk[13] TLD kindly provided to us by Nominet for the purpose of this study. We collect zone files on a daily basis. Note that the majority of ccTLD registry operators are under no obligation to

[6] We use "`hxxp`" notation to defang a malicious URL.

[7] `https://publicsuffix.org`

[8] https://github.com/korlabsio/urlshortener

[9] `https://github.com/korlabsio/subdomain_providers`

[10] `https://czds.icann.org/home`

[11] `https://internetstiftelsen.se/en/domains/tech-tools/access-to-zonefiles-for-se-and-nu/`

[12] `https://securityblog.switch.ch/2020/11/18/dot_ch_zone_is_open_data/`

[13] `https://registrars.nominet.uk/uk-namespace/the-uk-zone-files/`

make their zone files openly available. Therefore, we use several passive and active measurement methods to obtain a more exhaustive list of domains of ccTLD that do not provide access to zone files. This step is intended to give a comprehensive list of domain names currently registered in all TLDs. The domain names will then be mapped to their registrars using the registration information as set out in Section 1.3 and used to estimate the sizes of registrars. Using our measurement approaches and available zone files, we enumerate each month over 300 million registered domain names. For comparison, in September 2022, DomainTools reported 361M domain names.[14]

1.3 Technical Registration Information

For each collected domain name we attempt to gather registration information using the Registration Data Access Protocol (RDAP[15]) or WHOIS[16] protocols, and extract the name of registrar, registrar identifier, domain creation and expiration dates. We do not process or store any registrant data. We perform scans for all newly registered or observed domains as soon as they are acquired and periodically (at least once per month) for all domain names (e.g., ∼300M domains in June 2022). Each month we can collect and parse technical registration information for about 90% of collected domain names. In June 2022, we collected WHOIS records for ∼258M domain names (∼86% of collected domain names). For the remaining domains, we cannot gather registration data for several reasons, such as the lack of a WHOIS server for a given TLD, as discussed later.

To identify a registrar, for each RDAP/WHOIS record, we first extract IANA ID field if it is present and corresponds to the ICANN-accredited registrar name.[17] If IANA ID is not present, we extract the registrar name from the WHOIS record and, whether possible, we match it with the registrar name in the ICANN-accredited list of registrars, and finally map the domain name to the corresponding IANA ID. The second step requires painstaking manual verification to ensure accuracy of the method. Using this approach, in June 2022, we reliably mapped ∼234M unique domain names to their corresponding ICANN-accredited registrars (∼91% of all domains for which we collected IANA ID or registrar name).

It is common practice that the same corporate entity may have multiple IANA IDs due to, for example, merging registrar companies. At the time of writing, for example, it appears that there are four IANA IDs assigned (accredited) to Alibaba Group:[18] 420, 1599, 3775, and 3819. However, we do not merge entities if the IANA IDs are different, as this is error-prone and requires systematic and continuous manual analysis of the registrar market.

[14] `https://research.domaintools.com/statistics/tld-counts`
[15] `https://datatracker.ietf.org/doc/html/rfc7482`
[16] `https://www.rfc-editor.org/rfc/rfc3912.txt`
[17] `https://www.icann.org/en/accredited-registrars?filter-letter=a&sort-direction=asc&sort-param=name&page=1`
[18] `https://www.alibabagroup.com`

Note that ccTLD registries are under no obligation to use IANA identifier or a particular naming convention for registrars. They may use a completely unique local identifier (e.g. an alpha, numeric or alpha-numeric string) or they may choose to use IANA identifiers for those registrars that are ICANN accredited. The identifier may or may not be displayed on the ccTLD's WHOIS. It is generally unlikely that all registrars for a particular ccTLD are ICANN accredited.

A ccTLD with a numeric registrar ID naming convention may choose to display the corresponding IANA ID for their registrars who are accredited under ICANN. Confusingly, for registrars that are not ICANN accredited, they may display the numeric string labeled as an "IANA ID" but it is not an IANA ID. We suspect this is a result of using open source WHOIS software designed for the gTLD ecosystem and substituting a local identifier.

This means, for ccTLDs WHOIS lookups: (i) some will display no identifier at all, (ii) some will display a local identifier that is unrelated to the IANA ID, (iii) some will display an identifier labeled as "IANA ID", but it is unlikely that all of these will actually be IANA IDs, some may look like they could be IANA IDs but are a local identifier. Sometimes the identifier is intentionally chosen to exist in a range outside of IANA IDs to prevent colliding with another registrar identifier. The result of this is that it is particularly challenging to map all ccTLD registrars against a centralized database.

For example, at the time of writing, the analysis of the WHOIS record of the domain name 'baba.in', shows that it was registered with 'PDR Ltd. d/b/a PublicDomainRegistry.com' with an IANA ID 303. However, the WHOIS record shows the IANA ID as 801217, which is not the valid registrar IANA ID based on the list published by ICANN. We have extensively analyzed WHOIS data, identified cases where an identifier labeled as "IANA ID" does not correspond with the IANA ID list, and removed such domain names from the analysis of registrars.

Note that different ccTLD registries operate under different jurisdictions and may or may not provide specific fields in WHOIS. Some do not provide the registrar's name, registrar's abuse email address, or the creation date of the domain name. Some registry operators instead of providing query-based WHOIS/RDAP service ensure a web-based WHOIS service that may be protected by CAPTCHA. In such cases, we cannot map at scale domain names to the relevant registrars in order to estimate the number of domains under their management, nor can we map abusive domain names to registrars. Despite the limitations described above, each month, we are able to precisely identify ICANN-accredited registrars for about 90% of the collected WHOIS records.

Currently, statistics are calculated only for ICANN-accredited registrars, but we also collect and process information on registrars accredited locally by ccTLD registries, which we consider for inclusion in future reports. For reporting by TLDs, abuse identified in domains managed by local registrars is included in the total numbers reported for that ccTLD zone.

Finally, we attempt to map all domain names found in the abuse feeds to the corresponding registrar names in the same way as described above, using RDAP/WHOIS records collected and parsed as soon as we acquire malicious URLs.

1.4 Uptime Measurements

For each unique abusive domain name, we measure the uptime (also referred to as persistence of abuse), defined as the time between the malicious URL has been blocklisted and abuse has been mitigated (i.e., maliciously registered domain and/or hosting service has been suspended and/or abusive content has been removed from the website). We consider that the abuse has been mitigated, even if only the malicious content has been removed.[19] This determination stems from our observation that the same entity may provide domain registration and hosting services. In order to minimize the damage to victims and the potentially harmless domain name registrant, it is common practice to first remove the malicious content and then gather evidence to determine whether the domain name is registered by the attacker or is a legitimate registration that has been the subject of some other compromise. Depending on the assessment, the company may also suspend the registered domain name if it is malicious. To accommodate such cases, we mark the domain name abuse as remediated, even if the mitigation action took place only at the hosting level. Given that for maliciously registered domain names mitigation is typically accomplished at the registrar level, we measure and calculate uptimes only for registrars rather than TLD registry operators.

We actively collect various information related to abusive URLs and registered domain names, namely the content of the malicious URL and the home page of the registered domain name, DNS, and WHOIS records. We extract features used to determine whether the maliciously registered domain has been removed from the zone and/or hosting service has been suspended and/or abusive content has been removed from the website. After the initial measurement, performed at the time of acquiring the malicious URL, we repeat the measurements for one month: 5 minutes after blocklisting, 15m, 30m, 1 hour, 2h, 3h, 4h, 5h, 6h, 12h, 24h, 36h, 48h, and then once every 12 hours. Typically malware delivery and phishing attacks are mitigated within the first day after blocklisting [2]. Therefore, we perform more granular scans at the beginning of the measurements and less frequent measurements later.

Even though some of the URLs which appear on the blocklist remain accessible after one month, we do not continue the measurement and set the uptime to one month. Some URLs obtained from blocklists are already mitigated at the time of the first scan. If our system detects such cases, we calculate the time between listing and the first measurement, which is usually very short and provides a good approximation of the mitigation time.

As the phishing attacks grow in sophistication and use evasion techniques to avoid detection and tracking of malicious websites [3], our measurement platform

[19] While having only the content removed counts as mitigation for our report, a more complete remedy would be to suspend the domain name as well, because otherwise the domain name might be reused by the attacker in other phishing or malware delivery campaigns.

may not always be able to determine whether abuse has been mitigated or not. Previous work revealed that client-side evasion techniques, known as cloaking grew from 23% to 33% between 2018 and 2019 [3]. Some phishing attacks serve the phishing website only to specific regions or specific browser types. Some of them prevent the end user from visiting the phishing site more than once. Such cases are excluded from the uptime analysis and investigated manually. The measurement platform constantly evolves to account for evasion techniques and minimize the number of undetermined cases over time.

We manually analyze a sample of URLs that were not mitigated within one month and confirm that some were false positives, i.e. legitimate websites and domain names incorrectly labeled as malicious. In order to systematically minimize or eliminate their impact on the overall uptime metric, we calculate only the median uptime, which is less susceptible to skewing caused by false positives than the mean.

Finally, the obtained results (median uptime) may reflect the mitigation policies of some individual registrars, i.e. the maximum time they process phishing or malware delivery reports and mitigate abuse (e.g., within 12 hours of being blocklisted). We plan to contact the relevant registrars to validate our results.

1.5 TLD Sizes

To obtain a meaningful, quantitative metric, representing the relative distribution of abusive domains per TLD, we first need to estimate their sizes, or in other words, the number of domains under management (DUM). Whenever possible, we calculate the number of domains directly from available zone files. For all other TLDs, similarly to the previous work [4], we use approximate sizes estimated made public by DomainTools.[20] For example, in September 2022, there were approximately 6,271,000 .nl domain names registered,[21] while DomainTools reported approximately 5,955,000 .nl domains[22] (~95% of all registered .nl domain names).

1.6 Malicious versus Compromised Domains

While some domains are registered purely for malicious purposes, others are benign but compromised (e.g., by exploiting website security vulnerabilities [5] or misconfigured nameservers [6]). In either case, such domain names affect the reputation of all intermediaries involved in hosting, content distribution or domain registration, including TLD registries and registrars. Distinguishing between these two classes of abuse is crucial for mitigation efforts. A registrar, upon receiving a report of abuse and confirming that the domain name appears to be maliciously registered and engaged in phishing or malware distribution, should take mitigation action at the DNS level. Domain names compromised at the hosting or website level should generally not be mitigated at the DNS level

[20] `https://research.domaintools.com/statistics/tld-counts/`
[21] `https://stats.sidnlabs.nl/en/registration.html`
[22] `https://research.domaintools.com/statistics/tld-counts/`

to avoid collateral damage to the registrant and website visitors. Instead, the registrar should forward the complaint to the hosting provider, which should remove the abusive content and patch the vulnerable hosting.

Existing methods for categorizing domain names are based on a set of predefined heuristics (such as the method used in Global Phishing Survey [7]) or on machine learning-based approaches such as the COMAR classifier [8]. Previous work has shown that simple heuristics-based methods provide relatively high accuracy, but can result in a high rate of false positives (maliciously registered domain names classified as compromised) and are much easier to evade [8]. The machine learning approach has proven to be very accurate with a very low rate of false positives [8].

In this study, we use a hybrid method based on the MalCom classifier developed for research purposes by KOR Labs–conceptually similar to COMAR, achieving very high accuracy–and on mitigation actions taken by registrars or TLD registries at the DNS level. MalCom, like COMAR, is based on a large set of pre-selected features and automatically generated models based on ground-truth data (automatically and manually labeled maliciously registered and compromised domain names). MalCom uses a new set of features and active learning, i.e., the models are periodically updated to account for changes in attackers' behavior, making it harder to evade over time.

While machine learning-based approaches are highly accurate and can support registrars and TLD registries regarding the type of mitigation actions to take, they might still provide incorrect classification results due to, for example, missing values (e.g., calculating the age of a domain name is only possible if the creation date in WHOIS can be retrieved). To further increase the classification accuracy, we collect *a posteriori* evidence indicating malicious registration based on mitigation actions. Specifically, we flag a domain as malicious if the domain name was removed from the zone file or the hosting service was suspended for a registered domain. Note that even if we detect a mitigation action at the level of the malicious site rather than at the registered domain name level, we continue our measurements because the domain name may also be blocked later.

Finally, if, based on the mitigation action, we determine that the domain has been maliciously registered, we will categorize it as such, otherwise we will use the classification results obtained from the MalCom classifier.

2 Security Metrics

We use two types of security metrics [9] in the reports: *i)* distributions of abusive domain names (occurrence) and *ii)* persistence of abuse (uptimes). They provide a complementary view of the DNS Abuse problem, prevention, and mitigation. The distributions may indicate the preferences of malicious actors (that may choose to abuse, for example, one registrar and not the other) and can be driven by the registration policies of registrars and TLD registries. The persistence of abuse shows how promptly intermediaries mitigate abuse once it has occurred.

In our previous work [4, 10, 11], we proposed three complementary occurrence metrics: distributions (or rates) of abusive domain names, fully-qualified domain names (FQDNs) and URLs. While the distribution of domain names is the most intuitive metric, it comes with a limitation: it may not always reflect the “the amount of abuse” associated with a given domain name. One domain name can be used in one phishing attack and another in multiple attacks causing more harm to end-users. However, measuring “the amount of abuse” or, in other words, harm caused to the victims is very challenging and the two additional metrics must be carefully interpreted. Our manual analysis reveals important limitations of the previously proposed two complementary occurrence metrics. For example, we observe that each time the victim (or a crawler) visits some malicious websites, unique URLs are being generated and labeled as abusive. The domain ‘serverss-kundenserverss.xyz’ (maliciously registered with ‘1API GmbH’ with IANA ID 1387) was reported to our system 79,931 times from the APWG feed during the May 2022 period, each time with a different randomly generated URL path, but with the same fully qualified domain name. In such a case, the URL-based occurrence metric may over-count malicious resources and affect the accuracy of security metrics. Therefore, we measure and calculate the occurrence metric only for unique abusive domain names, not for URLs or FQDNs.

While the absolute number of abusive domains by intermediary gives insights into DNS Abuse, distributions relative to the number of domains under management by TLD registries or registrars allow more reliable comparisons. Therefore, the reports will show the number of abused domains normalized by TLD or registrar sizes.

Given the variety of intermediaries involved in the domain name registration process and hosting, as well as multiple options an attacker has in abusing domain names, TLD security metrics reflect the “healthiness of a TLD ecosystem” rather than the security performance of individual TLD registries. That said, voluntary security practices or registration policies of TLD registries can reduce DNS Abuse (e.g., early detection systems). Note that even benign domain names (registered by legitimate users), with websites that have been compromised, can be abused and become a vehicle for phishing or malware distribution attacks. Those abuse the reputation of legitimate businesses and the reputation of all intermediaries involved, such as TLD registries and registrars, even if they might not be best positioned to mitigate it. More importantly, victims (and even domain name registrants) often do not distinguish between the intermediaries involved in domain registration and hosting and can not identify the right entity to contact about abuse. Still, victims can eventually identify an abuse contact of TLD registries, which, once notified, may forward abuse complaints to intermediaries better positioned to mitigate it. Therefore, for TLDs, we calculate the abuse rates using the following formula:

$$Rate = \frac{Occurrence}{DUM} \times 100 \ [\%] \tag{1}$$

It expresses the percentage of all abusive domains (cf. Section 1.1) to domain names under management (DUM) for each TLD in a given month as explained in Section 1.5.

For each registrar, similarly to TLDs, we use Formula 1 to calculate the occurrence metric (abuse rate) as a percentage of abusive domain names to domains under management (cf. Section 1.2 and 1.3). For each registrar, we also calculate the median uptime metric (cf. Section 1.4), which is less susceptible to skewing caused by false positives than the mean uptime. As explained in Section 1.3, if the registration information for a given abusive domain name is not available in the public WHOIS, or it cannot be queried at scale or parsed, we exclude such a domain from further analysis (occurrence and uptime metrics).

Acknowledgments

We would like to thank Anti-Phishing Working Group, OpenPhish, Phishtank, and Abuse.ch for giving access to their ULR blocklist data, Nominet for providing access to the .uk zone file, and reviewers for their constructive and valuable feedback.

References

[1] V. L. Pochat, T. V. hamme, S. Maroofi, T. van Goethem, D. Preuveneers, A. Duda, W. Joosen, and M. Korczyński, "A practical approach for taking down avalanche botnets under real-world constraints," in *27th Annual Network and Distributed System Security Symposium, NDSS*. The Internet Society, 2020.

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

[3] P. Zhang, A. Oest, H. Cho, Z. Sun, R. Johnson, B. Wardman, S. Sarker, A. Kapravelos, T. Bao, R. Wang, Y. Shoshitaishvili, A. Doupe, and G. Ahn, "Crawlphish: Large-scale analysis of client-side cloaking techniques in phishing," *IEEE Security and Privacy*, vol. 20, no. 2, pp. 10–21, 2022.

[4] M. Korczyński, S. Tajalizadehkhoob, A. Noroozian, M. Wullink, C. Hesselman, and M. van Eeten, "Reputation metrics design to improve intermediary incentives for security of tlds," in *2017 IEEE European Symposium on Security and Privacy (Euro SP)*, April 2017.

[5] S. Tajalizadehkhoob, T. van Goethem, M. Korczyński, A. Noroozian, R. Böhme, T. Moore, W. Joosen, and M. van Eeten, "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,*. ACM, 2017, pp. 553–567.

[6] M. Korczyński, M. Król, and M. van Eeten, "Zone Poisoning: The How and Where of Non-Secure DNS Dynamic Updates," in *Proceedings of the 2016 ACM on Internet Measurement Conference*, ser. IMC '16. ACM, 2016, pp. 271–278.

[7] G. Aaron and R. Rasmussen, "APWG Global Phishing Survey: Trends and Domain Name Use in 1H2014," http://docs.apwg.org/reports/APWG_Global_Phishing_Report_1H_2014.pdf.

[8] S. Maroofi, M. Korczyński, C. Hesselman, B. Ampeau, and A. Duda, "COMAR: Classification of Compromised versus Maliciously Registered Domains," in *2020 IEEE European Symposium on Security and Privacy (EuroS&P)*, 2020.

[9] M. Korczyński and A. Noroozian, "Security reputation metrics," in *Encyclopedia of Cryptography, Security and Privacy.* Springer Berlin Heidelberg, 2021. [Online]. Available: https://doi.org/10.1007/978-3-642-27739-9_1625-1

[10] M. Korczyński, M. Wullink, S. Tajalizadehkhoob, G. C. Moura, and C. Hesselman, "Statistical Analysis of DNS Abuse in gTLDs Final Report," Tech. Rep., 2017. [Online]. Available: https://www.icann.org/en/system/files/files/sadag-final-09aug17-en.pdf

[11] M. Korczyński, M. Wullink, S. Tajalizadehkhoob, G. Moura, A. Noroozian, D. Bagley, and C. Hesselman, "Cybercrime after the sunrise: A statistical analysis of dns abuse in new gtlds," in *Proceedings of the 2018 on Asia Conference on Computer and Communications Security.* ACM, 2018, pp. 609–623.