DNSAI
INTELLIGENCE

SEPTEMBER 2022 REPORT

DNS ABUSE INSTITUTE

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This report is the first publication from the DNS Abuse Institute's measurement initiative: DNSAI Intelligence.

The intention of this report is to establish a credible source of metrics for addressing DNS Abuse. We hope this will enable focused conversations, and identify opportunities for reducing abuse across the DNS ecosystem.

We aim to create an interactive, evidenced data set that provides sufficiently granular information to improve how the industry and the wider DNS community understands, prevents, and mitigates DNS Abuse.

We hope future iterations of this report create an opportunity to celebrate and recognize good practice, as well as shine a spotlight on potential for areas of improvement in the industry. We hope to understand through these reports which factors, policies, and processes are effective, and empower the industry with evidence.

DNS Abuse impacts everyone. We want to use this understanding to improve the overall health of the DNS ecosystem. Fundamentally, we want to prevent or quickly mitigate harm to end users, businesses, governments, civil society organizations, public services, and the general public while preserving the benefits and principles of an open Internet.

While we expect to offer much greater levels of granularity in our upcoming reports, our first report focuses on higher level aggregate data from May, June, and July 2022. We have this higher level approach for this initial report to allow for further data collection and to gather feedback as to what would be most helpful in our future reports.

We also provide interactive dashboards that are available on our **website** and will be updated monthly.

One of the focuses of the data we gather is abuse up-times or "time to live." This is an important metric because it explains how long DNS Abuse is present on a domain before the harm is mitigated. In order to measure abuse up-times, we track each domain name for 30 days which means the data in our reporting is delayed by one month.

In creating these reports, we have optimized for accuracy and reliability. This means that some of our numbers will necessarily be lower than some other reports on phishing and malware prevalence which may look to the URL level rather than the domain name level.

Because our data collection efforts are just beginning, we do not attempt to make any conclusions about the data at this time. We look forward to reviewing data as time goes on and patterns cement into trends. However, we do offer commentary on how our methodology captures data and will provide a foundation for understanding this complex problem going forward. We encourage all readers to consider the detailed methodology and contact us with questions, ideas, or suggestions to help us improve this initiative. After all, we are here to support the DNS Community and make it better equipped to tackle DNS Abuse.

The DNS Abuse Institute will periodically publish reports on **DNSAI Intelligence**.

ABOUT

Mission

The **DNS Abuse Institute** (DNSAI or the "Institute") was created in 2021 by **Public Interest Registry** (PIR) in pursuit of its non-profit mission. At the Institute, we have one simple mission: reduce DNS Abuse. To do so, we need a reliable, independent, transparent, and sufficiently granular way of measuring DNS Abuse in order to reduce it at the DNS level.

The DNS community does not currently have tools readily available to understand both the prevalence and mitigation of DNS Abuse. Similarly, we have only a limited understanding of the impact of registrar and registry policies and processes on DNS Abuse. If DNS Abuse were a disease, the community has only an anecdotal view of the symptoms, without knowing the causes or the opportunities for treatment.

Our purpose for measuring DNS Abuse is to increase our understanding of the problem and bring greater sophistication to community discussions. With the ultimate goal of reducing abuse in mind, mitigation should still take place at the appropriate level.[1]

We hope this work will build on the existing laudable community initiatives and elevate discussion by providing a source of data specific to the needs of the DNS community. Some existing initiatives include ICANN's **DAAR** project[2] which provides important insights into high level trends, indicating the incidence of DNS Abuse[3] across multiple blocklists and identifying changes over time. There are also security companies that produce reports focusing more on network protection rather than threats at the DNS level, that are typically utilized in enterprise environments for intranet protection. These reports are no doubt useful for network protection, they are not always fit for purposes for addressing abuse via the DNS.

Our ambition is to provide a transparent and replicable resource for DNS level discussions about the prevalence and mitigation of phishing and malware across the open Internet. Our methodology and reporting have been specifically designed to support DNS community discussions on this important issue.

The goal of these reports is to advance community understanding of DNS Abuse and, ultimately, improve industry practice. We hope it will provide insights into how DNS Abuse is changing over time and allow concrete, specific conversations about how abuse is impacting the domain registration industry.

We want to drive changes in the domain registration industry based on this evidence by improving our understanding of where DNS Abuse is concentrated and what works to prevent and mitigate it. We want to highlight good practices and applaud improvements, while also taking note of areas that should be addressed.

This project will start by focusing on the harms of malware and phishing. We have chosen phishing and malware because generally there is sufficient verifiable evidence of the security threat. We may expand to other forms of DNS Abuse in the future, provided that doing so is consistent with our mission and the priorities we have chosen for this initiative.

This initiative will work in tandem with our other initiatives to reduce DNS Abuse: Collaboration and Education. It will guide us toward areas ripe for educational best practices and collaboration within and outside the industry.

[1] See DNS Abuse Definition: Attributes of Mitigation for further information <https://dnsabuseinstitute.org/dns-abuse-definition-attributes-of-mitigation/> Published 24 August 2021.

[2] https://www.icann.org/octo-ssr/daar

[3] DAAR covers: Botnet C&C, Phishing, Malware, Spam

Priorities

Our priorities for DSNAI Intelligence are:

- **Transparency:** The methodology that collects, cleans, and aggregates the data must be as transparent as possible. To the extent that anyone should wish to, they could replicate the process.
- **Credibility and Independence:** We aim to have an academically robust and independent approach, separate from commercial interests.
- **Accuracy and Reliability:** The goal of these reports is to enable focused conversations, and to identify opportunities for abuse reduction. The data needs to be of high enough quality to serve as the foundation for meaningful changes to the ecosystem.

Strategy

To ensure these priorities are reflected in the delivery of this initiative, we have made several decisions in the pursuit of these principles. These decisions go to the heart of the methodology, but are also visible in our operational approach to publication and engagement.

To ensure the report is independent, reliable, and uses academically robust methodology we are working with an experienced independent third party to lay out the methodology and conduct the data gathering. The technical analysis for this project is performed by **KOR Labs**, led by **Maciej Korczynski** from Grenoble INP-UGA. Korczynski's existing work in this field includes providing the Technical Analysis for the recent **European Commission's DNS abuse study**, the **Classification of compromised versus maliciously registered domains (COMAR) study**, **SADAG**(a project commissioned by the ICANN Competition, Consumer Trust, and Consumer Choice Review Team), and a paper on **Reputation Metrics Design to Improve Intermediary Incentives for Security of TLDs**. This data is then provided to the DNSAI. DNSAI then works with PIR's Data Analytics team to create the interactive charts and for the purposes of writing this report.

To ensure transparency we have included the detailed methodology provided by KOR Labs at the end of this report. We encourage readers to contact us if they have questions, comments, or suggestions. We, in collaboration with KOR Labs, expect to iterate the methodology as we identify opportunities to improve our reporting. The methodology contains several important decisions that support our priorities.

In creating these reports, we demanded high standards in our reporting, accuracy and reliability. This means our numbers will necessarily be lower than some other reports on phishing and malware prevalence which may look to the URL level rather than the domain name level.

We use a carefully examined and curated set of blocklists, and once a URL is placed on a blocklist it passes through the KOR Labs deduplication, data enrichment and domain extraction process. The blocklists included are:
Anti-Phishing Working Group (APWG), PhishTank, OpenPhish and ABUSE.ch (URLhaus feed). We may consider adding additional sources in the future, provided they meet our requirements.

We endeavor to have evidenced data for domains included in our reporting. The DNS Abuse included should be actionable; a bare domain name is insufficient for our reporting. Our methodology involves KOR Labs actively collecting information related to abusive URLs and registered domain names, such as the content of the malicious URL, the home page of the registered domain name, DNS, and registration records (excluding any registrant data; we do not access or process any registrant data whatsoever in our methodology or data gathering).

We have also specifically excluded multiple reports of the same domain name to avoid duplication which contributes to our more targeted reported levels of abuse. For example, one domain name was reported 79,931 during one month with a different randomly generated URL path after the same domain name (i.e. example.[TLD]/[randomly generated strings]). These URLs would be counted as one unique domain name in our report, not 79,931.

The methodology includes details of the existing limitations impacting the coverage of the study. Some of these limitations can be mitigated with industry engagement. For example, by providing zone file access to KOR Labs for research purposes, ccTLDs can improve the accuracy of the methodology and help us to estimate the domains under management (DUM) for registrars.

For now, our report focuses on ICANN-accredited registrars. KOR Labs does collect and process information on registrars accredited locally by ccTLD registries, which we will consider for inclusion in future reports, though this will require a manual effort of mapping domain names to their corresponding registrars accredited locally by ccTLD registries. Domain names that cannot be attributed to ICANN-accredited registrars are still included in the TLD-level statistics but are not considered in the registrar analysis.

To improve the accuracy and credibility of the reporting we are publishing the reporting in phases and are conducting significant outreach across the industry to sense check data, review our findings, and exchange insights.

Our initial reports will focus on high-level, aggregate statistics. We expect to provide much more granularity in our data (including mapping to the registrar and TLD level), but we want to continue to observe a few more months of data and discuss our findings with the industry and improve our reporting before we ultimately launch more detailed statistics. Subsequent versions will move towards that more granular approach with the identification of specific actors with good practices among registrars and TLDs, as well as those that have some opportunities for improving their abuse practices according to the data.

Measuring DNS Abuse is only one part of this puzzle, we also want to understand how much abuse is being mitigated and how quickly this is happening. To understand abuse persistence we are also measuring whether mitigation has taken place at specific intervals. These intervals begin with short time periods of minutes acquiring the URL from the blocklist, and become less frequent after 72 hours, finally completing after one month. This means there is a one month delay in our reporting due to the time taken to monitor whether mitigation has occurred, or whether these measures have changed over the month (e.g., been reversed).

Subsequent versions of our reporting will also include an improved methodology for distinguishing between compromised websites and maliciously registered domain names. By compromised domains we are referring to initially benign domain names that have been compromised at the website, hosting, or DNS level. This distinction is crucial for our purposes as the mitigation techniques should be different. A compromised website involves a victim who may need to improve their cyber security practices.[4] Suspension (i.e., removal from the DNS) or deletion of the domain name (without any concurrent engagement with the registrant) is typically not appropriate and could have broad unintended consequences (for example, cutting off the victim's email). KOR Labs is currently refining the methodology for this aspect of reporting, and we expect it to be ready for inclusion in the next iteration of reports.

Our approach is one of collaboration and engagement, and we endeavor to speak to all interested parties and provide them with early access to data that concerns their organization. We are committed to refining this project as work continues and welcome insights from across the industry to help us iterate and improve. If you would like to review your data, please do contact us.

For clarity, **DNSAI Intelligence** exists completely independently of **NetBeacon**, the centralized abuse reporting tool we created for the benefit of the DNS. Reports from NetBeacon do not go into our measurement work for Intelligence. This is a conscious choice to optimize and encourage usage of NetBeacon and prevent any abuse of NetBeacon as an attempt to influence DNSAI Intelligence data. See the methodology for more information on how domains are included in our measurement work.

[4] In the majority of cases, a compromised domain is a domain that has been registered for legitimate use and has become compromised through the website associated with that domain name, for example as a result of unpatched Content Management Systems (CMS). See our Best Practice for further information on securing your website to prevent compromise: https://dnsabuseinstitute.org/secure-your-website-save-the-internet

Understanding this Report

This report is the first publication from the DNS Abuse Institute's measurement initiative: DNSAI Intelligence.

This report shows high level aggregate data from **May, June, July 2022**.

It focuses on phishing and malware:

- **Phishing** is an attempt to trick people into sharing important personal information— banking information, logins, passwords, credit card numbers.
- **Malware** is malicious software designed to compromise a device on which it is installed.

It includes the following charts:

- **Chart 1: Overall Aggregate Trends**
- **Chart 2: Mitigation**
- **Chart 3: Time to Mitigation**
- **Chart 4: Malicious vs. Compromised**

Our methodology provides important context and we recommend it is read in full.

Each chart is accompanied by:

- '**About this Chart**' to help the reader understand the data being displayed, and;
- '**Commentary**' where we have added any observations on the data.

Where we are showing data over time, the intent is to try and demonstrate seasonality, year over year, and we are therefore hoping to be able to display about 2 years of data depending on functionality and viewability.

GENERAL DNS ABUSE TRENDS

These charts provide a broad overview of DNS Abuse trends.

Chart 1: Overall Aggregate Trends

About this Chart

This chart provides a high level view on how much DNS Abuse has been identified by our methodology, and how it's changing over time.

It shows the absolute volume of unique domains our methodology has identified are engaged in phishing and malware, broken out by category.

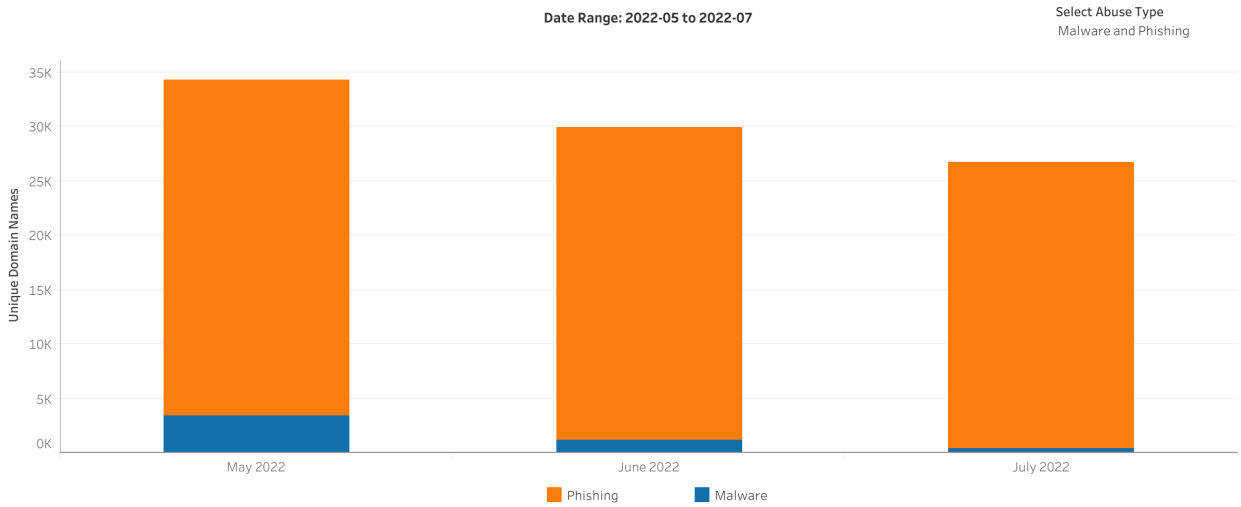

*Figure 1: Overall Aggregate Trends - **Phishing and Malware***

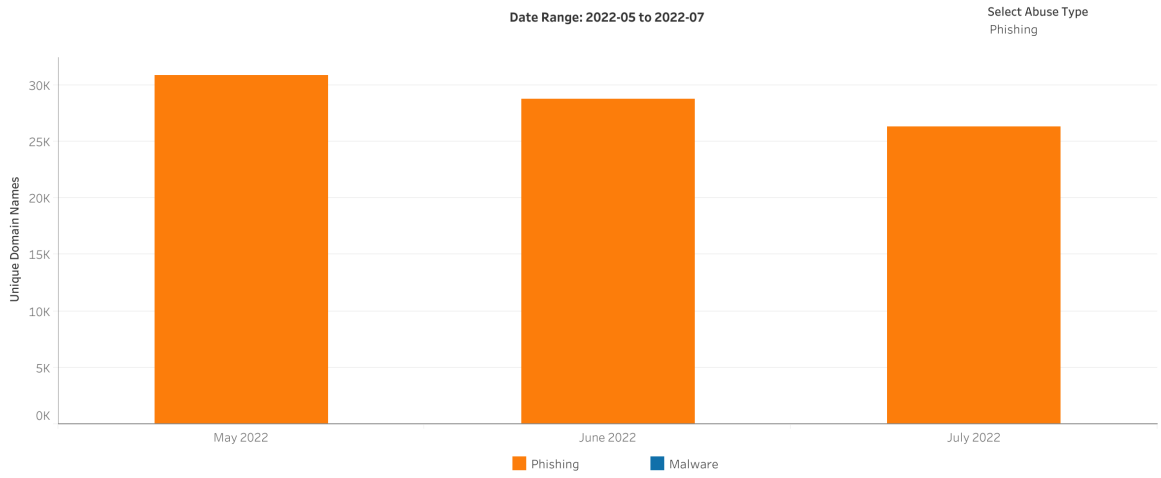

*Figure 2: Overall Aggregate Trends - **Phishing***

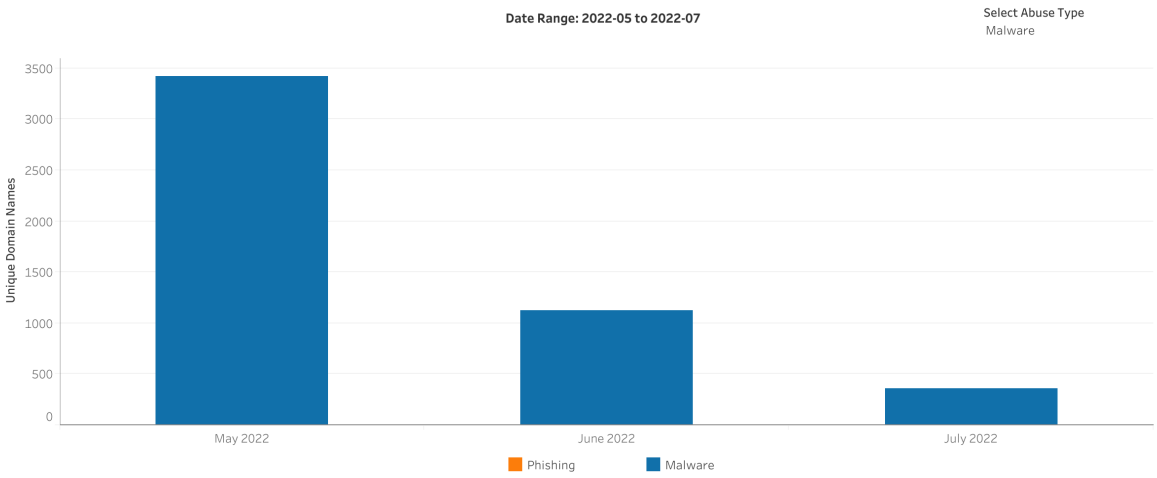

*Figure 3: Overall Aggregate Trends - **Malware***

Commentary

Our data set currently only includes three months, which is not enough data to determine a trend over time.

Our methodology identified more phishing than malware. This is inline with existing measurement reporting, such as ICANN's DAAR project[5], which typically reports more phishing than malware.

We've noted there is a considerable drop in the reported numbers of malware and we'll continue to investigate.

[5] https://www.icann.org/octo-ssr/daar

Chart 2: Mitigation

About this Chart

This chart is intended to demonstrate how much DNS Abuse we observe as being mitigated on a monthly basis.

The methodology includes a process to determine whether any mitigation has been observed. This involves taking an initial measurement of various factors related to the URL and repeating these measurements for one month. Further details are set out in the methodology.

This results in four labels:

- **Mitigated**: We believe a mitigating action has occurred. This action could be taken by a registrar, registry, a hosting provider, or another relevant actor.
- **Not Mitigated:** We did not detect any indication of mitigation.
- **Uncategorized**: We were unable to determine whether or not mitigation occurred.
- **Unprocessed**: The domains were not processed due to network connectivity or server problems.

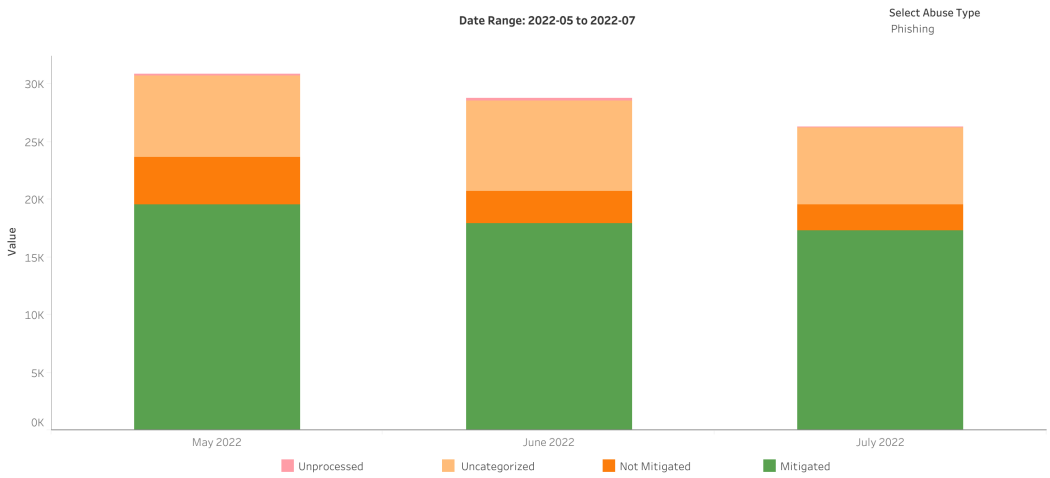

*Figure 4: Overall Mitigation Trends - **Phishing***

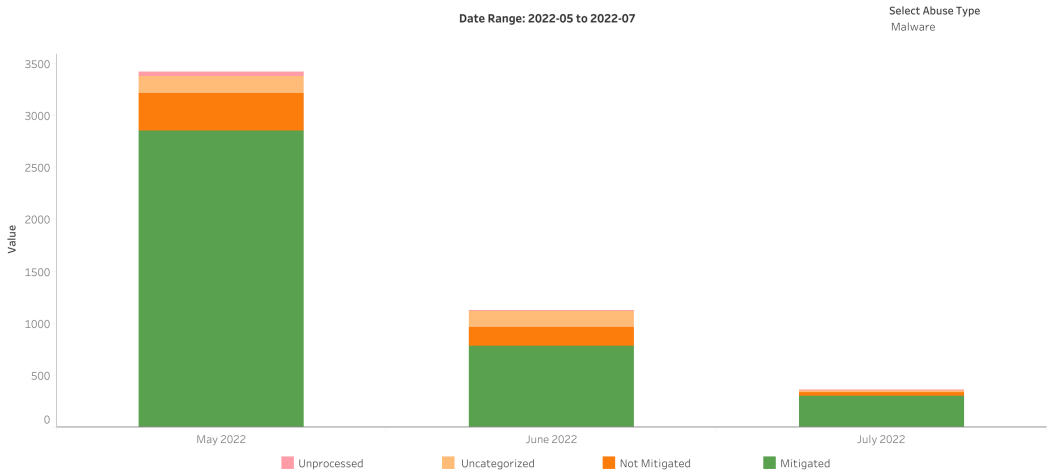

*Figure 5: Overall Mitigation Trends - **Malware***

Commentary

Our data set currently only includes three months, which is not enough data to determine a trend over time. We've noted there is a considerable drop in the reported numbers of malware and we'll continue to investigate.

The proportion of domains for which we were unable to categorize is higher for phishing than malware. One possible reason for this is the evasion techniques outlined in the methodology.

Missing or limited data is challenging to manage in any data driven project. In the pursuit of transparency, we have clearly identified the number of domains that we were unable to categorize or unable to process.

For future reports, we are working to balance principles of accuracy and reliability with the desire to compare trends over time. We want this to be a project of iterative improvement, with accuracy increasing over time. However, we also want the ability to compare trends over months and years. We are working on improving the breadth of coverage for categorization of mitigation activity, while avoiding significant changes to the existing categorization methodology for domains that could be categorized. See the methodology for further details.

Chart 3: Time to Mitigation

About this Chart

This chart is intended to show how the observed time taken to mitigate phishing and malware is changing over time.

For the domains that our methodology determined were mitigated, this chart shows how many registrars had a median time to mitigation in each category.

After an initial measurement, KOR Labs repeats measurements for one month to determine if mitigation has occurred. The intervals used are (starting at the time of acquiring the URL from the blocklist): 5m, 15m, 30m, 1hr, 2hr, 3hr, 4hr, 5hr, 6hr, 12hr, 24hr, 36hr, 48hr, and then once every 12 hours for one month.

While we are describing this information as a "median registrar mitigation time" it should be noted that we do not know definitively that it was the registrar that took action. This data could include mitigation taken by the registry, the host, or any other relevant party. The reference to a registrar is indicative that the domain is under their management.

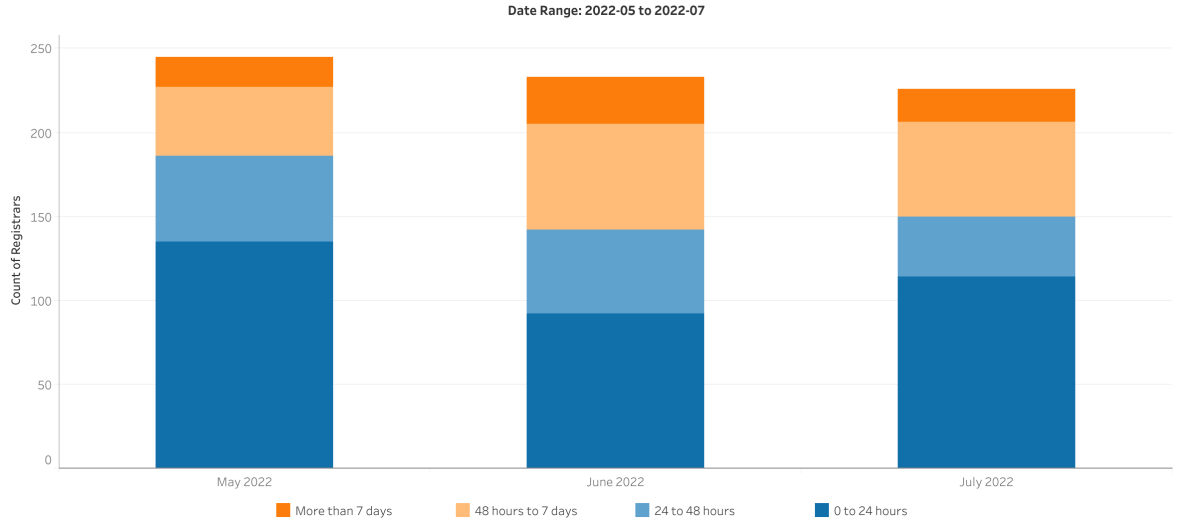

*Figure 6: Time to Mitigation - **May, June, July 2022***

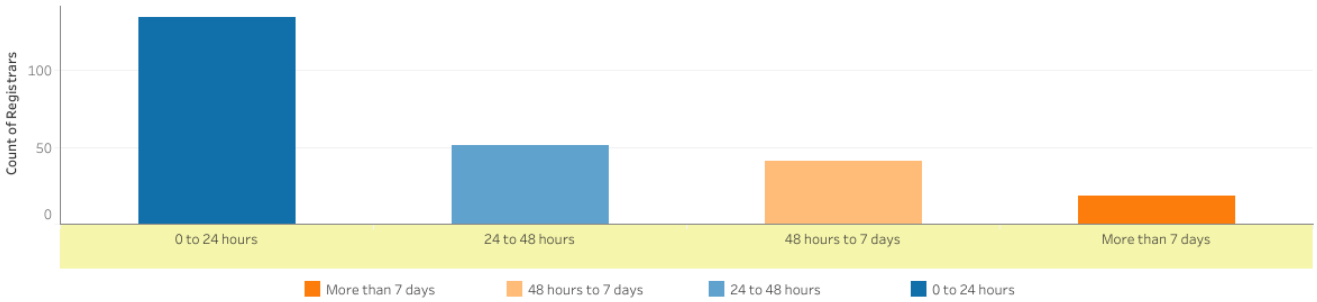

*Figure 7: Time to Mitigation - **May 2022***

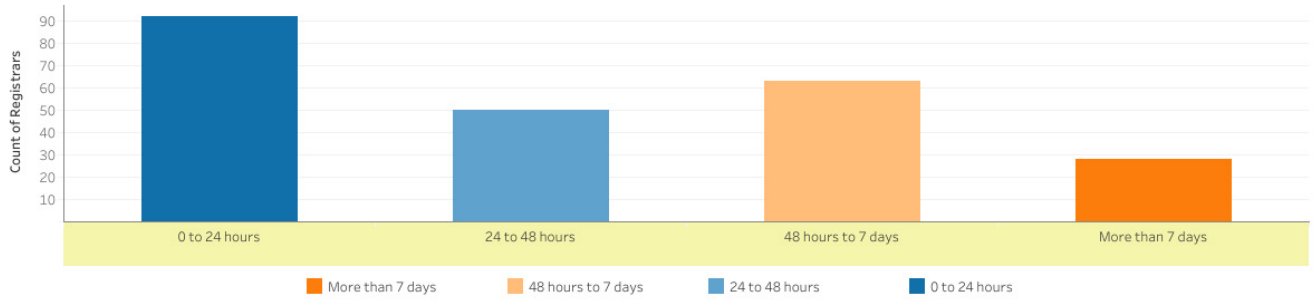

*Figure 8: Time to Mitigation - **June 2022***

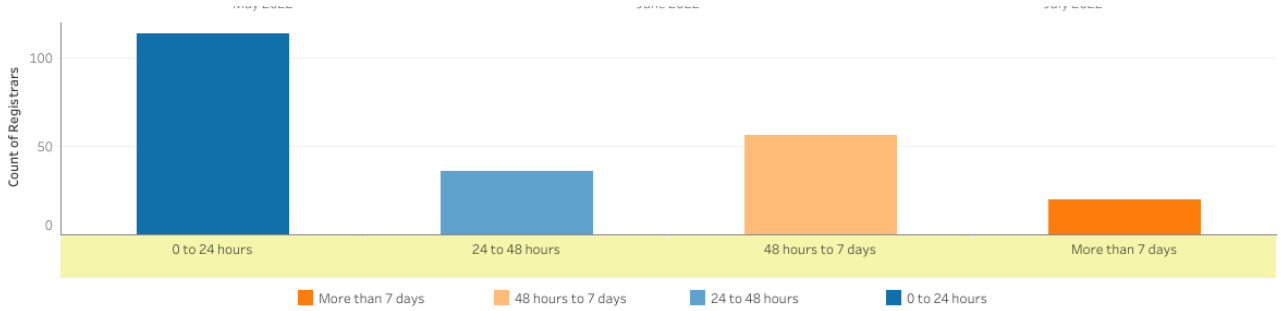

*Figure 9: Time to Mitigation - **July 2022***

Commentary

Our data set currently includes three months, which is not enough data to determine a trend over time.

There is no official industry standard for how quickly mitigation should occur. This makes the presentation of mitigation time challenging. We believe there is a general industry view that mitigation within 24 hours is thought to be a quick response to evidence of phishing or malware. As phishing and malware are quite time sensitive issues with most harm happening at the start of the attack, we believe that mitigation after 7 days is not quick enough to prevent and disrupt harm, which is why we have included it as a specific category.

As this is our first report we are reviewing the presentation of this data and considering whether there is a more meaningful way that we can present the results.

Chart 4: Malicious vs. Compromised

About this Chart

This chart is intended to show how any trends in malicious vs. compromised domains are changing over time. A compromised domain is a benign domain name that has been compromised at the website, hosting, or DNS level. The 'uncategorized' label refers to domains that our methodology was unable to categorize for a number of reasons, including problems in collecting the metadata necessary to classify domain names accurately.

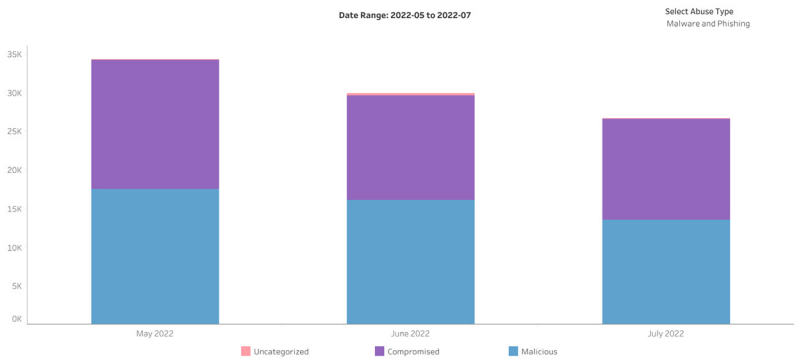

*Figure 10: Compromised vs Malicious - **Phishing and Malware***

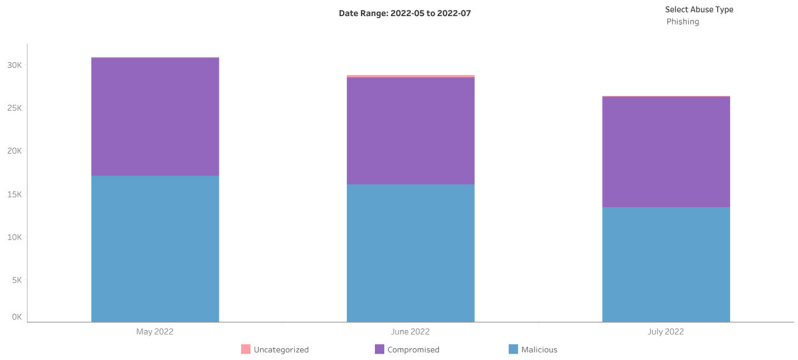

*Figure 11: Compromised vs Malicious - **Phishing***

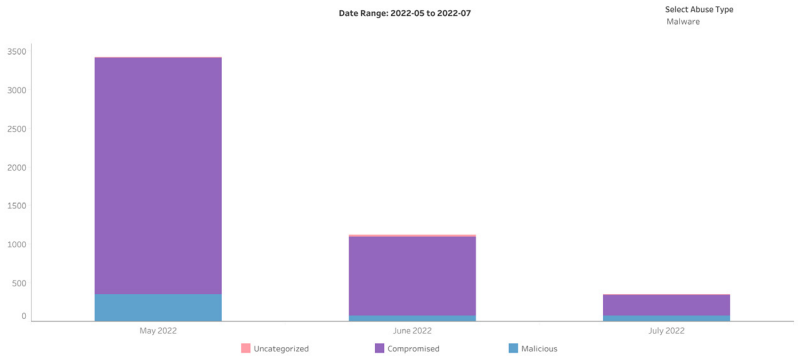

*Figure 12: Compromised vs Malicious - **Malware***

Commentary

Our data set currently includes three months, which is not enough data to determine a trend over time.

The distribution between domains identified as malicious or compromised is different for phishing and malware over this initial period. The data shows more domains identified as maliciously registered for phishing. For malware, more domains were identified as compromised.

DNS Abuse Institute Intelligence Platform: Methodology

V1.0

KOR Labs
maciej.korczynski@korlabs.io

The DNS Abuse Institute and KOR Labs are collaborating to publish the DNS Abuse Institute Intelligence reports, aimed at providing reliable and actionable data on the state of DNS Abuse. In this paper, KOR Labs explains the methodology it designed in helping the Institute develop these reports.

1 Data Collection and Processing

1.1 URL Blocklists

While there are various forms of DNS Abuse, for purposes of this report we initially selected phishing and malware because they generally provide sufficiently verifiable evidence of the security threat. The availability of verifiable evidence is typically not the case for other types of abuse, such as botnet command-and-control domain names or spam [1]. To measure the prevalence (i.e., DNS Abuse rate) and persistence (i.e., uptime) of abusive domain names involved in phishing and malware, 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 but will be selective in doing so. The chosen providers supply URLs in near real time via APIs. How often we download URLs 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.

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

- **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.
- **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, our proposed method is designed to reduce the impact of false positives on the uptime metrics (cf. Section 1.4).

From these 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, third-level, and beyond if a given registry provides such registrations, e.g., example.co.uk. Note that all the URL feeds used in this report include 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] We keep only domain names likely to have been registered by end users and exclude special domain names, to avoid, for example, google.com being flagged as abusive.

Finally, not all domain names can be processed, monitored or categorized for several reasons, such as network connectivity issues, blocklist server maintenance, or metadata collection problems. Therefore, certain

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

[7] https://publicsuffix.org

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

domains may be entirely excluded from the study or from some statistical analyses.

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 a list of domain names for each Top-Level Domain (TLD). To collect these lists, we process zone files obtained from the ICANN Centralized Zone Data Service (CZDS)[9] provided by participating generic TLDs (gTLD) that accepted our request for access. We also process zone files of some country-code TLDs, e.g., publicly accessible zones of .se, .nu,[10] .li, .ch[11] TLDs. We also plan to include the .uk[12] TLD zone file which was 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 make their zone files openly available. Therefore, we use several passive and active measurement methods to obtain a more exhaustive list of domains of ccTLDs 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 (gTLDs and ccTLDs). 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 size of the registrar domain portfolio (domains under management). Using our measurement approaches and available zone files, we enumerate over 300 million registered domain names each month. For comparison, in September 2022, DomainTools reported 361 million domain names.[13]

1.3 Technical Registration Information

For each collected domain name, we attempt to gather certain registration information using the Registration Data Access Protocol (RDAP[14]) or WHOIS[15] protocols, and extract the name of registrar, registrar identifier, domain creation and expiration dates. For the avoidance of doubt, we do not access, process, or store any registrant data at any point in our methodology. 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,

[9] https://czds.icann.org/home
[10] https://internetstiftelsen.se/en/domains/tech-tools/access-to-zonefiles-for-se-and-nu/
[11] https://securityblog.switch.ch/2020/11/18/dot_ch_zone_is_open_data
[12] https://registrars.nominet.uk/uk-namespace/the-uk-zone-files/
[13] https://research.domaintools.com/statistics/tld-counts
[14] https://datatracker.ietf.org/doc/html/rfc7482
[15] https://www.rfc-editor.org/rfc/rfc3912.txt

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 RDAP/WHOIS server for a given TLD, as discussed later.

To identify a registrar for a given domain name, for each RDAP/WHOIS record, we first extract the IANA ID field if it is present and corresponds to an ICANN-accredited registrar in the ICANN List of Accredited Registrars.[16] If the IANA ID is not present, we extract the registrar name from the RDAP/WHOIS record and, when possible, we attempt to match it with a registrar name in the ICANN List of Accredited 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:[17] 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.

Note that ccTLD registries are under no obligation to use the IANA identifier or a particular identifier 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 RDAP/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 in RDAP/WHOIS 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 RDAP/WHOIS software designed for the gTLD ecosystem and substituting a local identifier.

[16] https://www.icann.org/en/accredited-registrars?filter-letter=a&sort-direction=asc&sort-param=name&page=1

[17] https://www.alibabagroup.com

This means, for ccTLDs RDAP/WHOIS lookups: (i) some will display no registrar 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 it 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' in the .IN ccTLD, shows that it was registered with 'PDR Ltd. d/b/a PublicDomainRegistry.com' which has the IANA ID 303. However, the .IN 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 RDAP/WHOIS service ensure a web-based domain name registration information lookup 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 RDAP/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 will 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, to calculate the security metrics for registrars described below in Section 2, we attempt to map all domain names found in the abuse feeds (cf. Section 1.1) 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.[18] 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 appears that the common practice is 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 RDAP/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.

[18] 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.

As the phishing attacks grow in sophistication and use evasion techniques to avoid detection and tracking of malicious websites [3], our measurement platform may not 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 phishing attacks 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 incorrectly included in a blocklist. 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.[19] For example, in September 2022, there were approximately 6,271,000 .NL domain names registered,[20] while DomainTools reported approximately 5,955,000 .NL domains[21] (~95% of all registered .NL domain names).

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

1.6 Malicious versus Compromised Domains

While some domains are registered purely for malicious purposes (i.e., to carry out DNS Abuse), 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. Mitigation of maliciously registered domain names confirmed to be engaged in phishing and malware distribution can generally take place at the DNS level (i.e., through action by the registrar or TLD registry). In contrast, domain names compromised at the hosting or website level should generally not be mitigated at the DNS level 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 RDAP/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 suspended or deleted 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, hosting, or content delivery 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 help prevent or reduce DNS Abuse (e.g., early detection systems, and incentive programs). 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. Attacks using compromised websites 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 RDAP/WHOIS, or it cannot be queried at scale or parsed, we exclude such a domain from further analysis (occurrence and uptime metrics).

3 Acknowledgments

We would like to thank Anti-Phishing Working Group, OpenPhish, Phishtank, and Abuse.ch for giving access to their URL blocklist data, Nominet for providing

access to the .uk zone file, and reviewers for their constructive and valuable feedback.

4 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. Bohme, 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.

DNSAI
INTELLIGENCE

www.dnsabuseinstitute.org

© 2022 Public Interest Registry. All rights reserved.