Why do different DNS Abuse measurement projects result in different numbers?

March 2024

Contents

The Challenge of Measurement

Measuring DNS Abuse[1]—however you define it—is hard and complicated, but why do reasonable minds reach different conclusions on the numbers?

We hear this question frequently since the DNS Abuse Institute ("Institute") developed our measurement initiative: DNSAI Compass ("Compass"), a collaboration with KOR Labs.

The intention of this paper is to create a greater awareness of how DNS Abuse is measured and help the community to understand and interrogate data presented to them. It also highlights the importance of having transparent methodology.[2]

A measurement project should be able to provide you with details on how they reach their final numbers and explain decisions that were made along the way: which source lists were used, how was the data cleaned, high level details of the analysis, and guidance on how the data has be presented to help you understand and interpret the information provided.

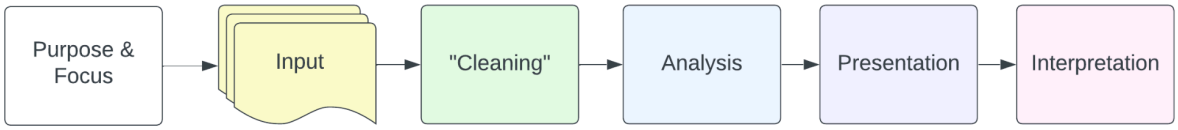
How is DNS Abuse Measured?

Projects to measure DNS Abuse typically follow a process of aggregating, cleaning, analyzing, and presenting data from multiple source lists.[3] Each step presents several decisions that can influence the outcome of the project. How these decisions are made will be influenced by the purpose and focus of the research as well as any priorities that have been identified. Finally, care must be taken in how to interpret data.

[1] These measurement challenges are subsequent to the challenge of definition DNS Abuse. We use the following definition: DNS Abuse is composed of five broad categories of harmful activity insofar as they intersect with the DNS: malware, botnets, phishing, pharming, and spam (when it serves as a delivery mechanism for the other forms of DNS Abuse). For more information see: https://dnsabuseinstitute.org/about-the-dns-abuse-institute/

[2] For Compass, our full methodology is available on our website: https://dnsabuseinstitute.org/wp-content/uploads/2022/10/DNSAI-Compass-Methodology.pdf

[3] The Statistical Analysis of DNS Abuse in gTLDs Final Report (SADAG) by SIDN Labs and Delft University of Technology commissioned by the Competition, Consumer Trust, and Consumer Choice Review Team with the support of ICANN was one of the first projects to utilize this method as a proof of concept. https://www.icann.org/en/system/files/files/sadag-final-09aug17-en.pdf ICANN's Domain Abuse Activity Reporting (DAAR) is also an early example of this approach. https://www.icann.org/octo-ssr/daar



Purpose & Focus

Most research projects will aim to clearly define their scope: what is being measured and what is not—and for what purpose. This will be impacted by their definition of DNS Abuse, and whether they have a particular research question or focus in mind. They may also identify priorities to guide their decisions. This can indicate which elements are most important and what compromises can be made.

Input

Most efforts to measure DNS Abuse typically ingest a number of source lists, sometimes called Reputation Block Lists (“RBLs”).

There are many sources lists to choose from, all with varying levels of quality and usually with limited or no evidence provided. Some are publicly available and free, others are provided under license agreements.

These lists are typically imperfect for the ultimate goal of measuring DNS Abuse at the domain name level because they are created for a different purpose: network blocking. This purpose means the lists are less sensitive to false positives than a registry or registrar needs to be when contemplating action at the DNS level because the impact of a false positive being blocked from a network is less severe compared to disabling a domain name.

Nevertheless, these lists are currently the best available starting point for understanding DNS Abuse. A research project will need to choose which combination of lists is best suited to their goals. After combining the best available options it is typically necessary to “clean” the dataset.

Cleaning

Because these lists are not bespoke to the project, the researcher will need to undergo a process of removing information that is not relevant to their purpose. If the project focuses on counting unique domain names for example, the researcher will need to isolate these by

removing duplicate domain names in different URLs,[4] and any URLs containing IP addresses rather than domain names. For DNS focused projects, it's usually important to remove domains for which mitigation at the DNS level would typically be inappropriate.[5]

Analysis

Once the researcher has refined their list in line with their purposes, they may conduct some analysis of the data. This can vary greatly and result in a very different presentation of the results. When analyzing registrar and TLD level data, it is important to remember the size of the Domains Under Management (DUM) varies considerably between different credentials and zones. For this reason, results are often normalized to reflect size (e.g. per 100,000 DUM).

Presentation

Once analyzed, the project will still need to make editorial decisions on how to present the information to the public. Some presentation choices may be made to reduce the impact of some of the challenges faced in measurement projects. For example, quality of the input data, false positives, small numbers, and skewed data.

When viewing a measurement project, it can be useful to ask questions to clarify what you're looking at. For example:

- Over what time period is the data displayed?
- Is the counting cumulative, and if so, can the domain leave the list if the harm has stopped? How does this happen?
- Are there any exclusions?
- Is there an attempt to manage small numbers or skewed data?
- Are the numbers raw or a relative figure (e.g., abuse per 100,000 domains under management (DUM)?

Interpretation

Finally, care must be taken in interpreting the results. While it's tempting to make definitive sweeping statements (e.g. "DNS Abuse is going up/down"), it's important to look for consistent

[4] As some phishing or malware campaigns generate a new URL every time the domain is visited (either by a human or a bot) it is not uncommon to see 70,000+ instances of one unique domain name on an RBL.

[5] The Institute and KOR Labs refer to these as "special domains". KOR Labs manually maintains a list of special domains, which is publicly available in our methodology and often utilized by other researchers as part of their measurement initiatives. Additional information is included below.

and sustained trends over time and consider all possible reasons for variations. There may be multiple explanations for changes in data, including seasonality, process or staffing variations, and the potential completeness/biases of source lists (e.g. language or geographical).

Compass

Compass was created with the intention of informing the mission and activities of the Institute and to empower the industry with relevant data to understand the prevalence and persistence of phishing and malware across the DNS. We hoped Compass would enable focused conversations, and identify opportunities for reducing abuse across the DNS ecosystem.

We use the following definition: DNS Abuse is composed of five broad categories of harmful activity insofar as they intersect with the DNS: malware, botnets, phishing, pharming, and spam (when it serves as a delivery mechanism for the other forms of DNS Abuse).

It was important for us to have an academically robust and independent approach, separate from commercial interests. We partnered with an external academic research organization: KOR Labs led by Dr Maciej Korczynski, a professor at Grenoble Alpes University in France.

One of our priorities was to optimize for accuracy and reliability, rather than attempt to measure all harm online. We believe it is more important to create high quality data that is measured consistently across the ecosystem and over time. This data serves as a foundation to understand changes and compare across different zones rather than comprehensively capturing all DNS Abuse. Our scope is therefore narrow and involves the collection of evidence.

As a result, KOR Labs chose to focus on phishing and malware insofar as they intersect with the DNS. This is because phishing and malware are the types of DNS Abuse believed to be most accurately evidenced.

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

KOR Labs chose source lists that were well reputed, had minimal incidence of false positives and are typically publicly accessible. Our input sources lists for Compass are:

- Anti-Phishing Working Group

- PhishTank
- OpenPhish
- URLHaus

Our focus for Compass is registries and registrars. This means our cleaning process is targeted at refining the source lists to unique domain names. As a result, our cleaning process de-duplicates by focusing on unique domain names, not URLs and removes IP addresses

We are also aiming to capture cases that are typically actionable for mitigation at the DNS level. This means a substantial step in our methodology is the removal of "special domain names," which are domain names that provide subdomains or redirections 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).

KOR Labs maintains and manually updates a list of "special domains" which are available to the research community. Maintaining this list is a substantial amount of manual work and it is a challenge to ensure we are aware of all current special domains. This list is utilized by other researchers and we encourage the research community to contribute to the manual maintenance.[6]

In terms of the analysis, it is important for our purposes that we understand the abuse we measure at a granular detail. This means there is significant analysis completed by KOR Labs including:

- Observing registry and registrar level data, not just aggregate data.
- A process to collect evidence; a mere domain name on a list is not sufficient.
- An observation of whether mitigation action has occurred and if it has, how quickly.
- An understanding of the type of registration: malicious or compromised.[7]

Finally, when summarizing the data we made several decisions about presentation with the intention to show a clear depiction of the data and how phishing and malware was relatively concentrated. These decisions are detailed in our monthly PDF publications on our website. For example, they include:

[6] See page 2 of our Methodology for more detail: https://dnsabuseinstitute.org/wp-content/uploads/2022/10/DNSAI-Compass-Methodology.pdf

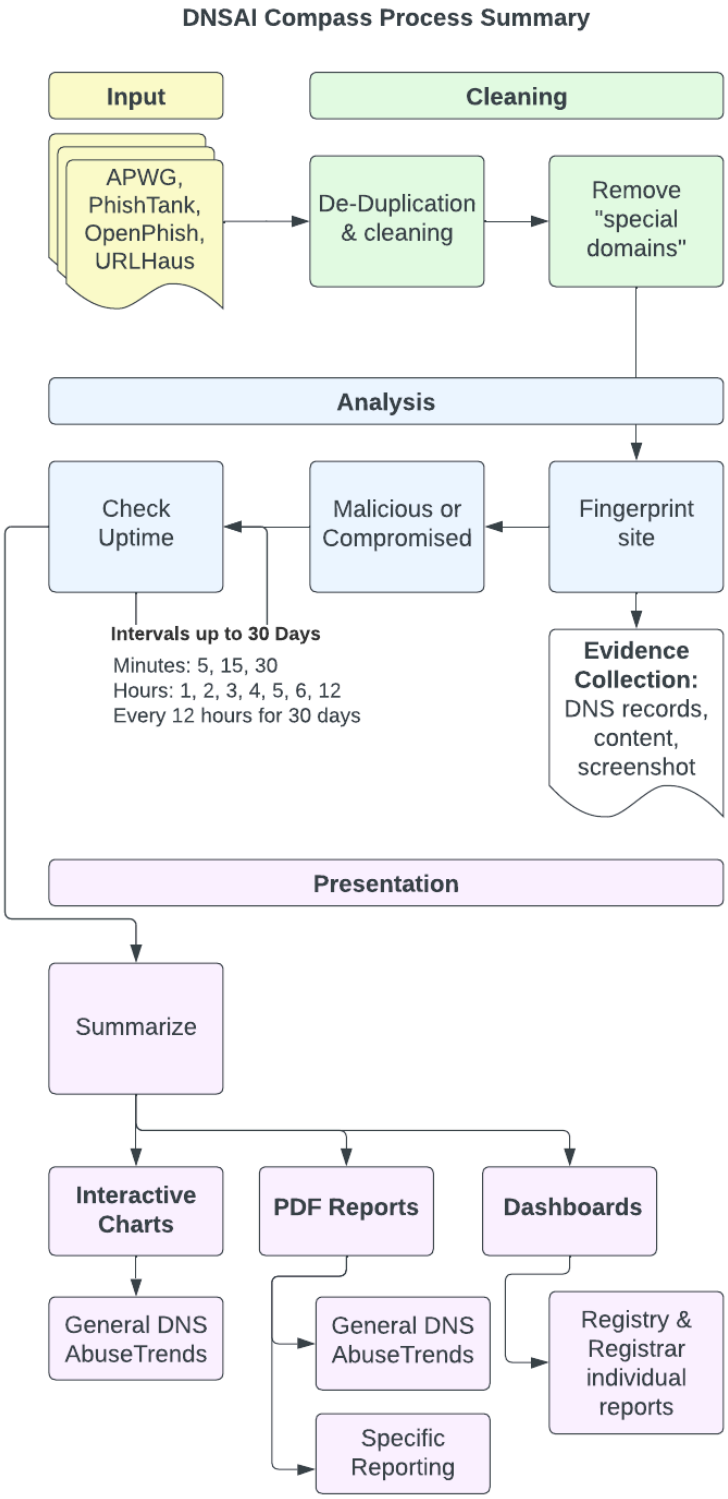
[7] We use the following definitions: Malicious: a domain registered for malicious purposes (i.e., to carry out DNS Abuse). Compromised: A benign domain name that has been compromised at the website, hosting, or DNS level.

- Minimum requirements for identified phishing and malware per month.
- A focus on malicious registrations as opposed to compromised.
- Consistency requirements for relatively higher rates of abuse each month, and a redaction process for those who do not consistently appear in the tables over a 6 month period. We chose a consistency requirement instead of an average due to challenges with skewed data and small numbers.[8]

In terms of interpretation, our Compass reports have typically only presented the observed results of the methodology. If we do include interpretations, we will clearly identify these as our interpretations and highlight any relevant limitations and caveats.

The following graphic shows an overview of our process.

[8] See slide 10 for an example. Any registry or registrar can be the target of a malicious campaign which could temporarily elevate their observed malicious abuse per 100,000 DUM.


DNSAI Compass Process Summary
Input
Cleaning
APWG, PhishTank, OpenPhish, URLHaus
De-Duplication & cleaning
Remove "special domains"
Analysis
Check Uptime
Malicious or Compromised
Fingerprint site
Intervals up to 30 Days
Minutes: 5, 15, 30
Hours: 1, 2, 3, 4, 5, 6, 12
Every 12 hours for 30 days
Evidence Collection: DNS records, content, screenshot
Presentation
Summarize
Interactive Charts
PDF Reports
Dashboards
General DNS AbuseTrends
General DNS AbuseTrends
Specific Reporting
Registry & Registrar individual reports

Summary

For more information on the Compass methodology, see our published methodology and our PDF reports.

In October, I had the pleasure of participating in a ccNSO DNS Abuse Standing Committee (DASC) Session at ICANN78 on this very topic, the session is available to watch online. Slides are available here. This blog post is a summary of my contributions to that session.

Compass provides free Dashboards for registries and registrars to understand the prevalence and persistence of phishing and malware in the domains they manage. If you'd like to access your own data, or meet with the Institute in person at ICANN79 or virtually, please email: support@DNSAbuseInstitute.org

About the DNS Abuse Institute:

The Domain Name System (DNS) Abuse Institute is tasked with creating outcomes-based initiatives that will create recommended practices, foster collaboration, and develop industry-shared solutions to combat the five areas of DNS Abuse: malware, botnets, phishing, pharming, and related spam. The Institute was created by Public Interest Registry, the registry operator for the .ORG top-level domain.