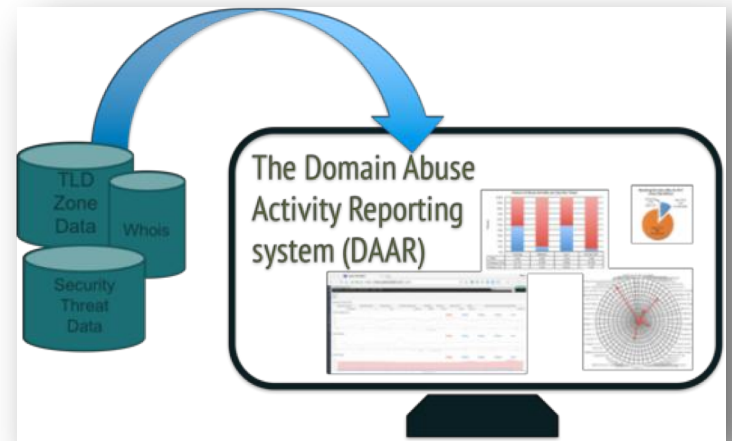


The Domain Abuse Activity Reporting System (DAAR)

Dr. Samaneh Tajalizladehkhoob

January 2019



The Domain Abuse Activity Reporting System

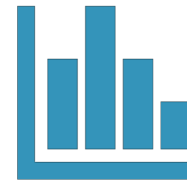
What is the Domain Abuse Activity Reporting system?

A system for reporting on domain name registration and abuse data across TLD registries and registrars

The Domain Abuse Activity Reporting System

How does DAAR differ from other reporting systems?

- Studies all gTLD registries and registrars for which we can collect zone and registration data
- Employs a large set of abuse feeds (e.g., blocklists)
- Accommodates historical studies
- Studies multiple threats: phishing, botnet, malware, and spam
- Takes a scientific approach: transparent, reproducible

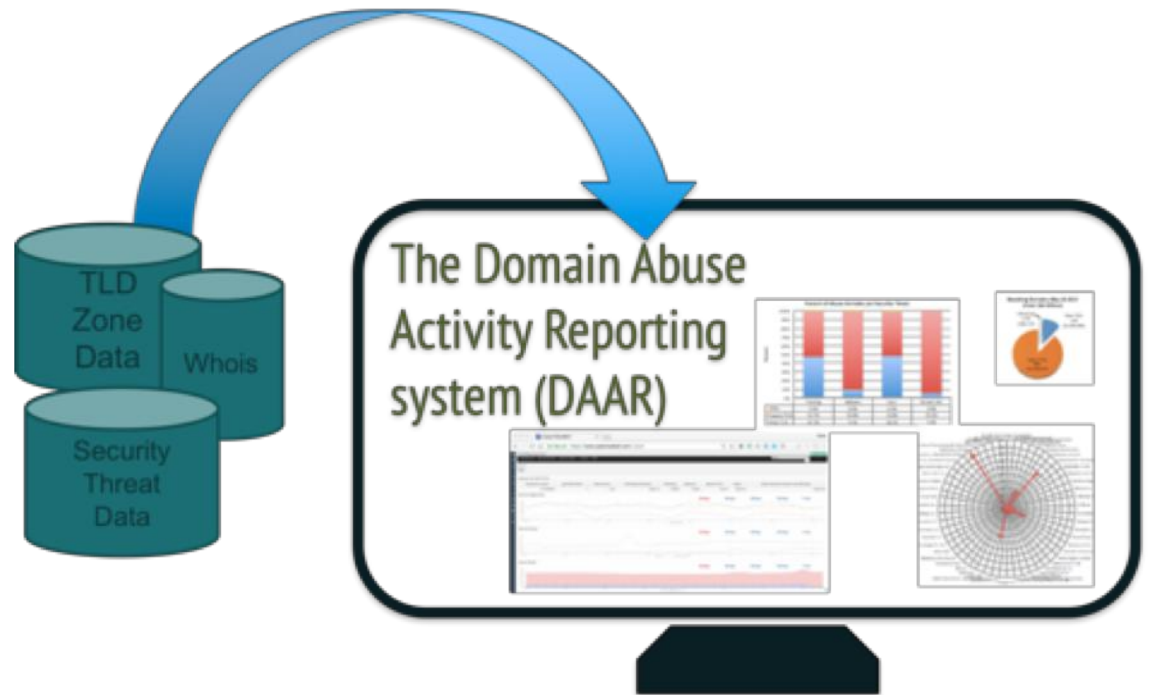


DAAR data can be used to

- Report on threat activity at TLD or registrar level
- Study historical security threats or domain registration activity
- Help operators understand or consider how to manage their reputations, their anti-abuse programs, or terms of service
- Study malicious registration behaviors
- Assist operational security communities

The purpose of DAAR is to provide data to support community, academic, or sponsored research and analysis for informed policy consideration

DAAR
Methodology
& Domain
Data



DAAR system uses data from public, open, and commercial sources

- I. DNS zone data
- II. WHOIS data
- III. Open source or commercial abuse threat (RBL) data*

* Certain data feeds require a license or subscription

The background of the slide features several thin, curved lines in shades of gray, some solid and some dashed, creating a sense of motion or data flow. On the left side, there is a large blue square with a smaller blue rectangle on top of it. The text 'i. DNS Zone Data' is centered within the blue square.

i. DNS Zone Data

- gTLD zones for gTLD domain and registry analytics
 - Uses publicly available methods to collect zone data such as Centralized Zone Data Service, zone transfer
- Uses domain names that appear in zone files
- Collect zone files from
 - Approximately 1220 gTLDs
 - Approximately 192 million domains

ii. WHOIS

- DAAR uses published registration data (WHOIS)
 - Uses registrar name and IANA ID
- Reliable, accurate registrar reporting depends on WHOIS
 - Scaling data collection from WHOIS is a big challenge

```
Domain Name: GOOGLE.COM
Registry Domain ID: 2138514_DOMAIN_COM-VRSN
Registrar WHOIS Server: whois.markmonitor.com
Registrar URL: http://www.markmonitor.com
Updated Date: 2011-07-20T16:55:31Z
Creation Date: 1997-09-15T04:00:00Z
Registry Expiry Date: 2020-09-14T04:00:00Z
Registrar: MarkMonitor Inc.
Registrar IANA ID: 292
Registrar Abuse Contact Email: abusecomplaints@ma
Registrar Abuse Contact Phone: +1.2083895740
Domain Status: clientDeleteProhibited https://ica
Domain Status: clientTransferProhibited https://i
Domain Status: clientUpdateProhibited https://ica
Domain Status: serverDeleteProhibited https://ica
Domain Status: serverTransferProhibited https://i
Domain Status: serverUpdateProhibited https://ica
Name Server: NS1.GOOGLE.COM
Name Server: NS2.GOOGLE.COM
Name Server: NS3.GOOGLE.COM
Name Server: NS4.GOOGLE.COM
```


iii) Abuse Threat Data

DAAR uses multiple abuse Reputation Blocklist (RBL) datasets to

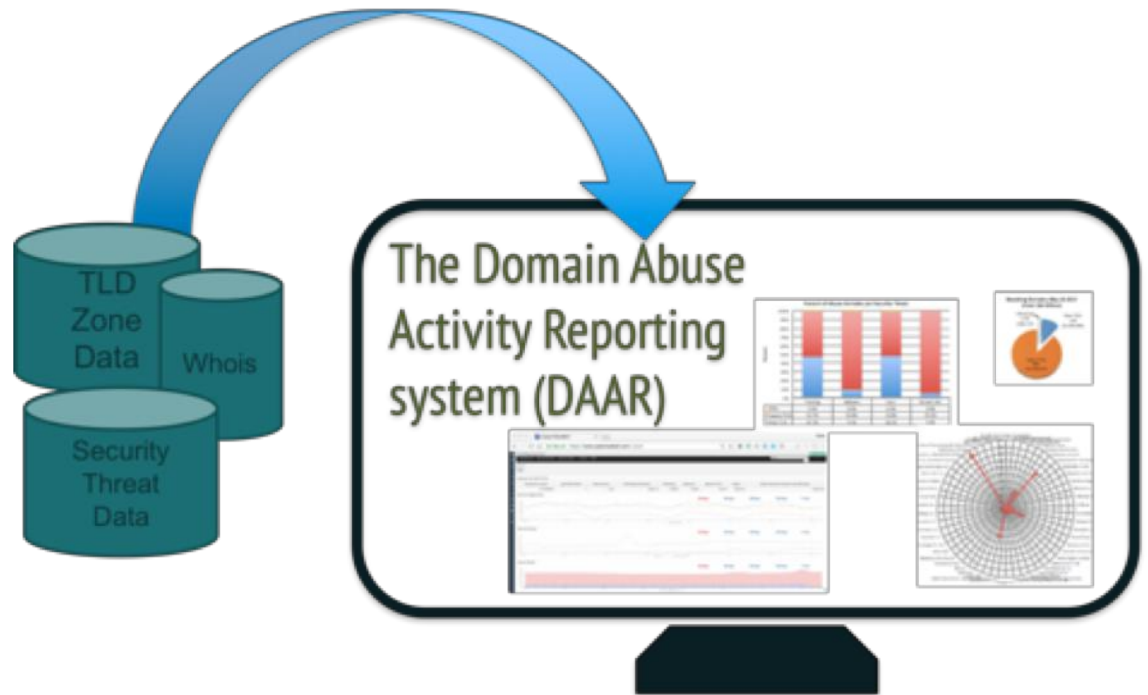
- Generate daily counts of domains associated with phishing, malware hosting, botnet C&C, and spam
- Calculate daily total and cumulative abuse domains
- Calculate monthly/yearly newly added abuse domains
- Create visual analytics regarding abuse trends in gTDLs

The background of the slide features several thin, curved lines in shades of gray, some solid and some dashed, creating a sense of motion or a globe-like pattern.

iii) Abuse Threat Data

- DAAR counts “unique” abuse domains
 - A domain that appears on **any** abuse datasets reporting to DAAR is included in the counts **once**
- DAAR reflects how entities external to ICANN community see the domain ecosystem

Reputation Block Lists : Identifying Threats



DAAR Criteria for RBL Data

- RBLs must provide threat classification that match our set of security threats
- RBLs have positive reputations in academic literature
- RBLs have positive reputations in operational and security communities for accuracy, clarity of process
- RBLs are broadly adopted across operational security community
 - Feeds are incorporated into commercial security systems
 - Used by network operators to protect users and devices
 - Used by email and messaging providers to protect users

Other Reputation Block List Uses

- RBLs in Browsers
- RBLs in the Cloud and Content-Serving Systems
- RBLs in Your Social Media Tools
- RBLs in the DNS
- RBLs in commercial firewalls, UTM devices
- RBLs in enterprise mail/messaging systems
- RBLs and Third-Party Email Service Providers (ESPs)


DAAR Is Not an Abuse List Service

- ICANN does **not** compose its own reputation blocklists
 - DAAR presents a composite of the data that external entities use to block threats
- DAAR collects **the same** abuse data that is reported to industry and Internet users and is used by
 - Commercial security systems
 - Academia and industry
- Academic studies and industry use validate these datasets exhibit accuracy, global coverage, reliability and low false positive rates

Does DAAR Identify All Abuse Data/Types?

- **No.** DAAR lists domain names associated with abuse identified by third parties.
- Only those names associated with generic TLDs are measured and only for specific abuse types.

- SURBL lists (domains only)
- Spamhaus Domain Block List
- Anti-Phishing Working Group
- Malware Patrol (Composite list)
- Phishtank
- Ransomware Tracker
- Feodotracker



RBLs in Academia: a Method to Assert RBL Confidence

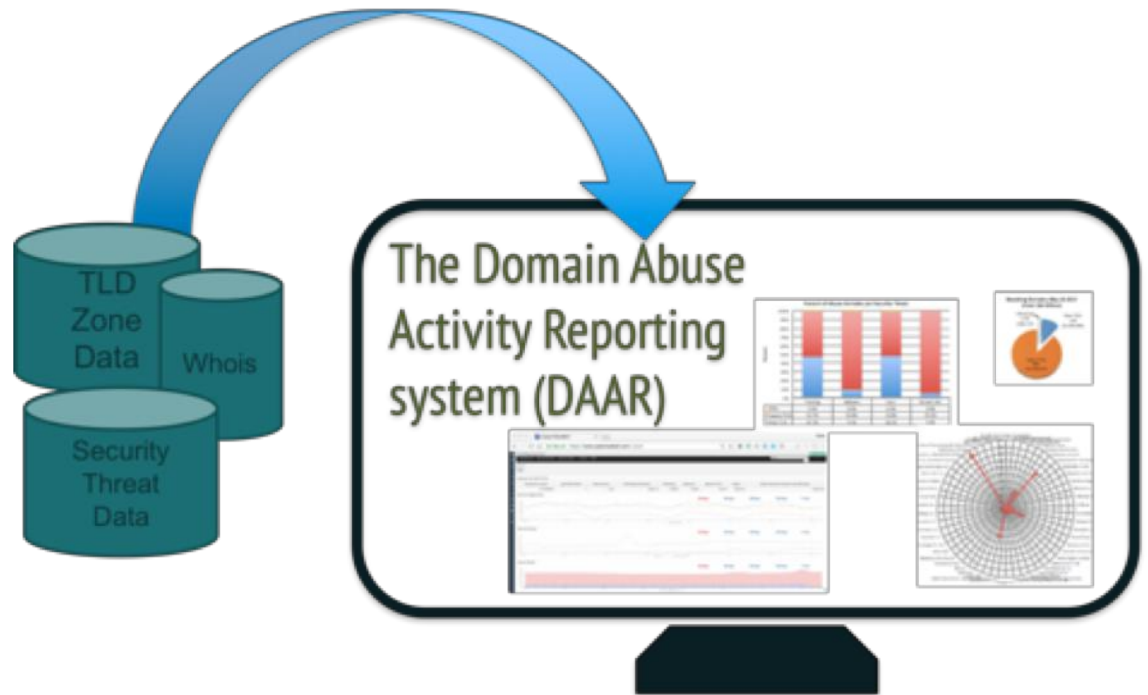
Partial list of academic studies and citations of RBLs that report to DAAR

- [Rotten Apples or Bad Harvest? What We Are Measuring When We Are Measuring Abuse](#)
- [Reputation Metrics Design to Improve Intermediary Incentives for Security of TLDs](#)
- [Empirically Characterizing Domain Abuse and the Revenue Impact of Blacklisting](#)
- [Blacklist Ecosystem Analysis: Spanning Jan 2012 to Jun 2014](#)
- [Taster's Choice: A Comparative Analysis of Spam Feeds](#)
- [Learning to Detect Malicious URLs](#)
- [Understanding the Domain Registration Behavior of Spammers](#)
- [The Statistical Analysis of DNS Abuse in gTLDs \(SADAG\) Report](#)
- [Shades of grey: On the effectiveness of reputation-based blacklists](#)
- [Click Trajectories: End-to-End Analysis of the Spam Value Chain](#)

Why Is DAAR Reporting Spam Domains?

- The ICANN Governmental Advisory Committee (GAC) expressed interest in spam domains as a security threat in its Hyderabad correspondence to the ICANN Board of Directors... Why? Because
- Most spam are sent via illegal or duplicitous means (e.g., via botnets).
- Spam is no longer singularly associated with email
 - Link spam, spamdexing, tweet spam, messaging spam (text/SMS)
- Spam is a major means of delivery for other security threats
 - Spam has evolved to a (cloud) service: Avalanche, for example, provided domain registrations to customers
- DAAR mainly measures domain names found in the bodies of spam messages
- MOST IMPORTANTLY, spam domain reputation influences how extensively or aggressively security or email administrators apply filtering

Project Status



The SSR team

- Reviewed all the reviews and comments received
- Published SSR [responses to DAAR comments](#) on February 1st, 2019

The SSR team

- Published the first series of the monthly reports
 - On Monday 4 February 2019 ICANN published the first monthly report from the DAAR system for January 2019.
 - The reports contain aggregated and anonymous descriptive statistics and trend analysis on abuse concentrations in gTLDs.
 - Monthly reports from previous months (January 2018 through Dec 2018) will be published before the end of February 2019 as well.
- The data has already enabled constructive and data driven discussions with industry members

\

DAAR & the Open Data Program

- Open Data Program aims to facilitate access to data that ICANN organization or community creates or curates
- In cases where licensing permits, DAAR data or reports will be published and included in the Open Data Program

Project Next Steps

- Investigating publication of data into the Open Data Program
- Improving the system based on comments and reviews
- Further developing new metrics and analytics based on current and future research based on DAAR

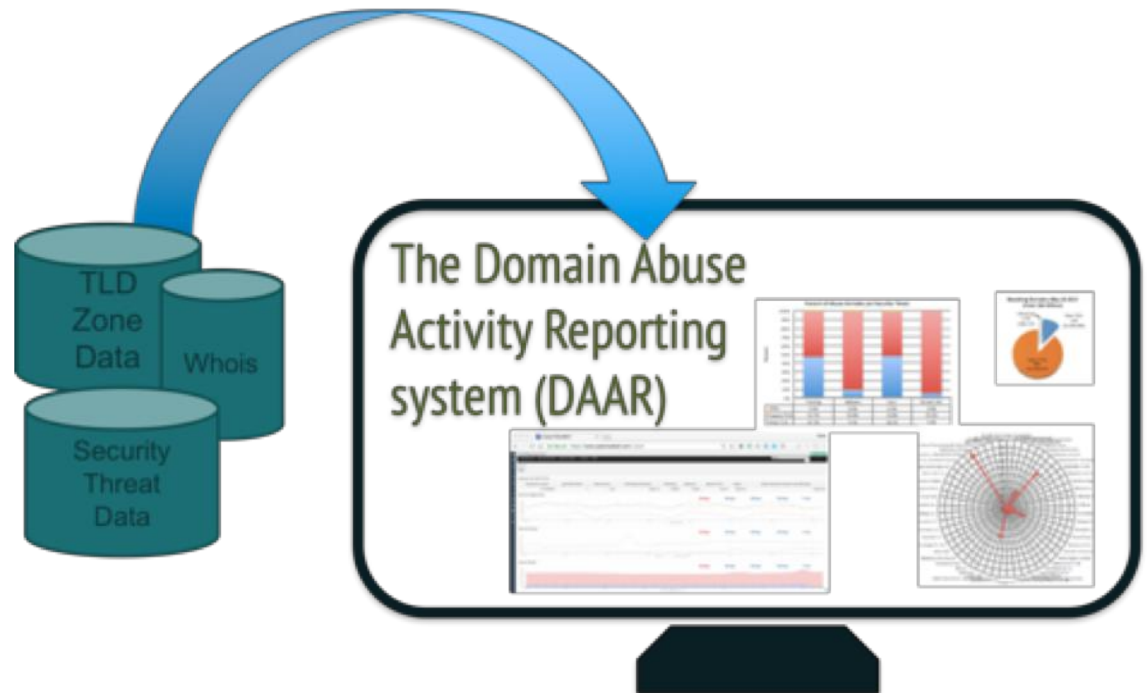
\

Project Next Steps

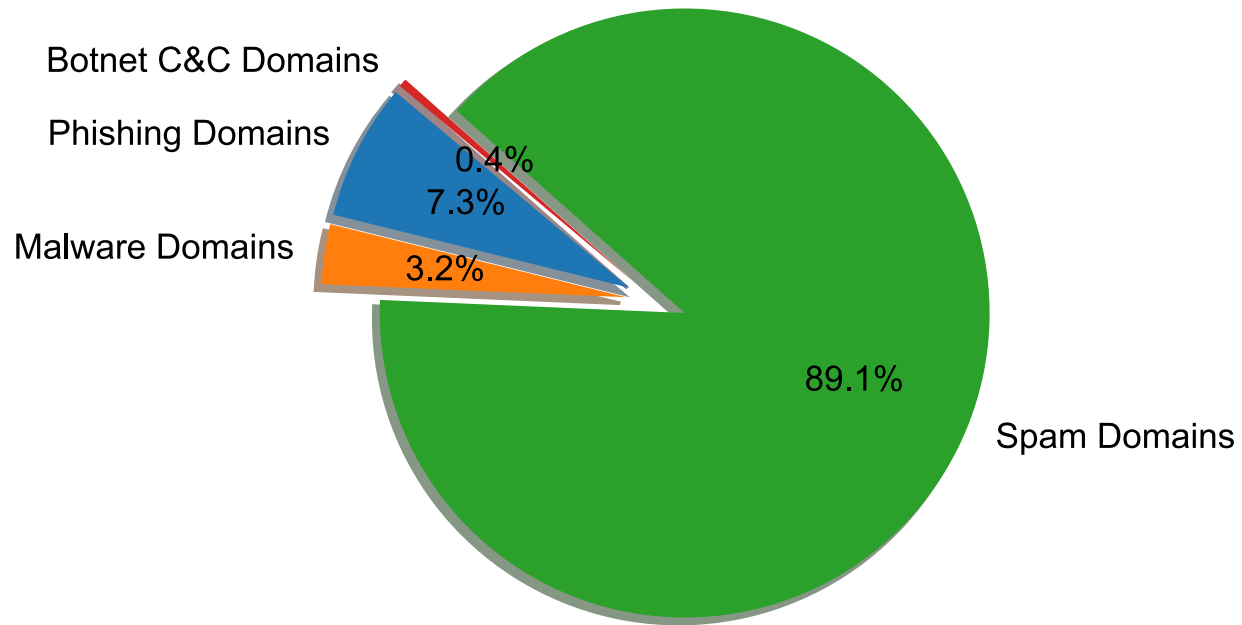
- Having discussions with registries who are interested in viewing their own data
- The SSR team does this in the context of sharing and learning from other security professionals in the industry

\

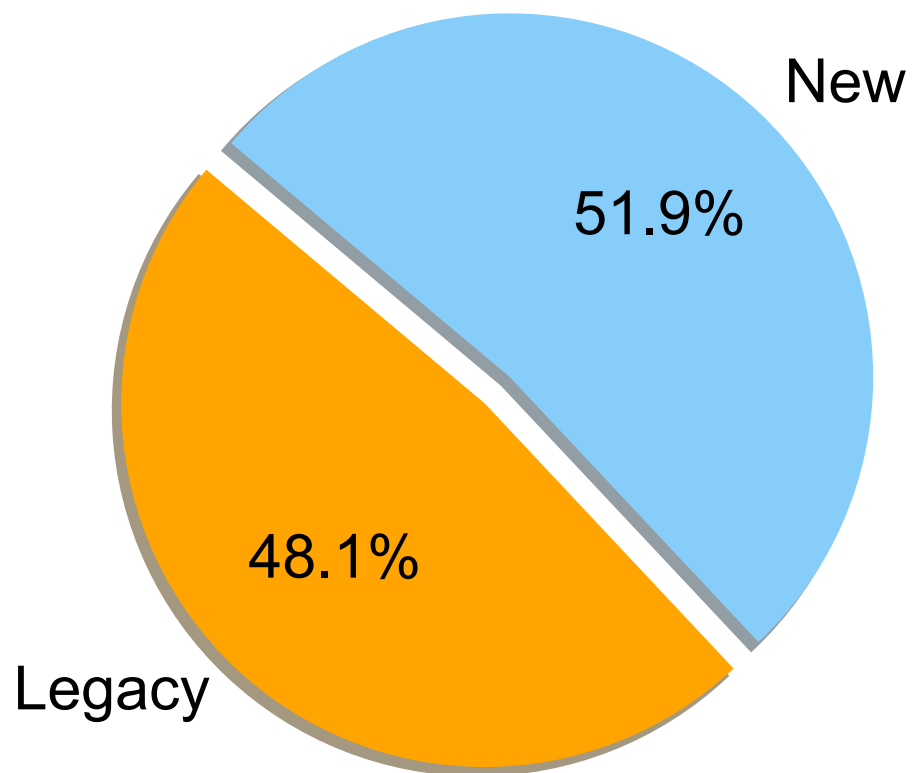
Visualizing DAAR Data



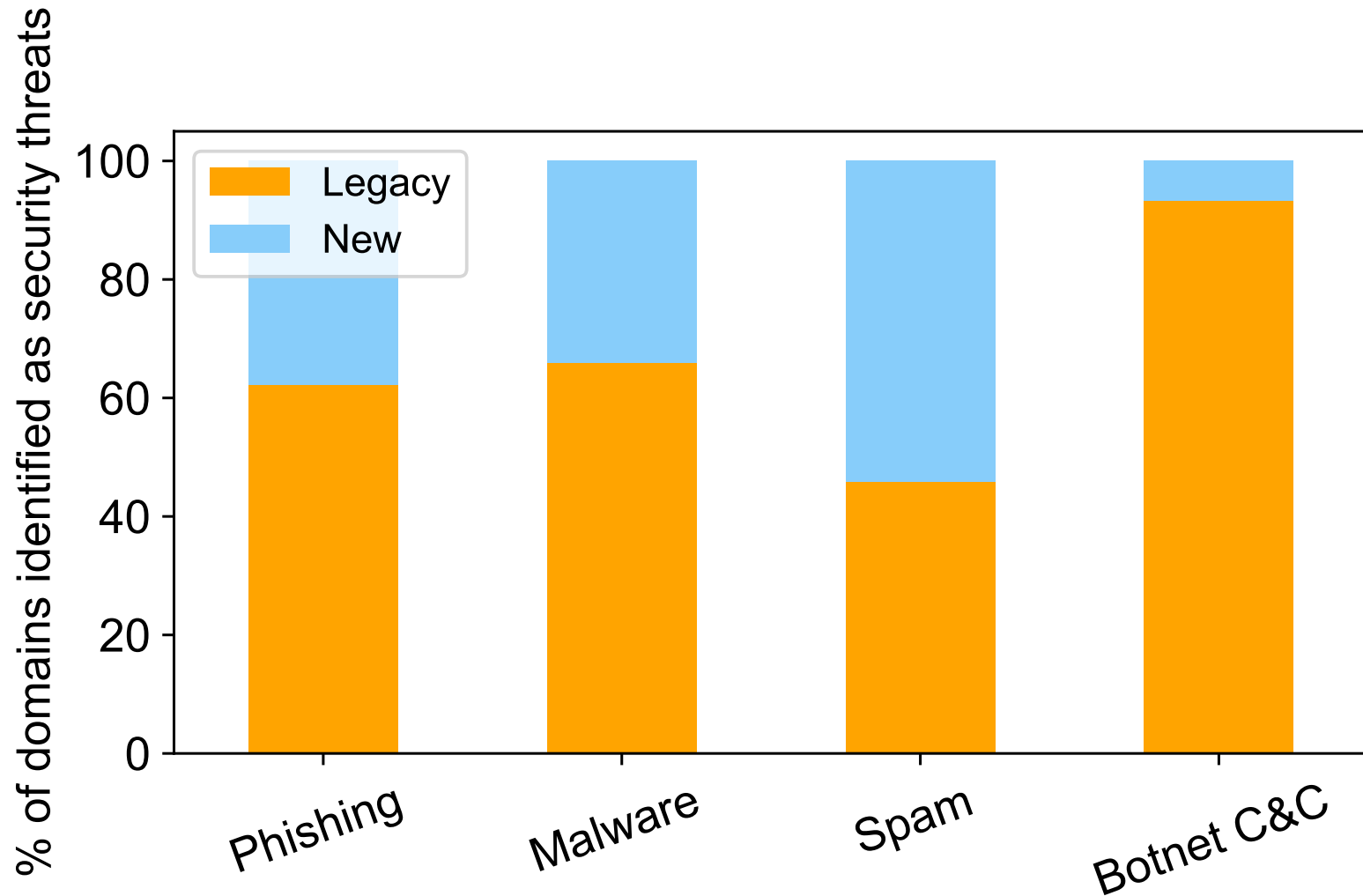
Overall Abuse Distribution in DAAR Data (Jan. 2019)



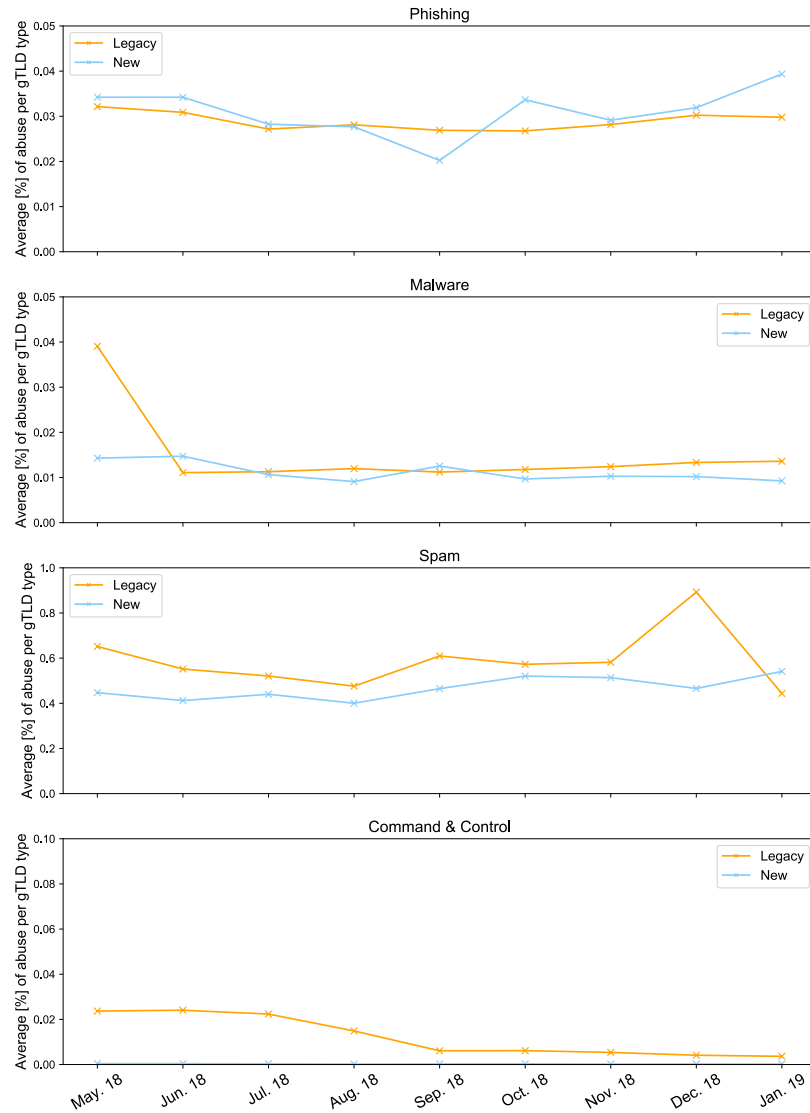
Distribution of Abused Domains in gTLDs



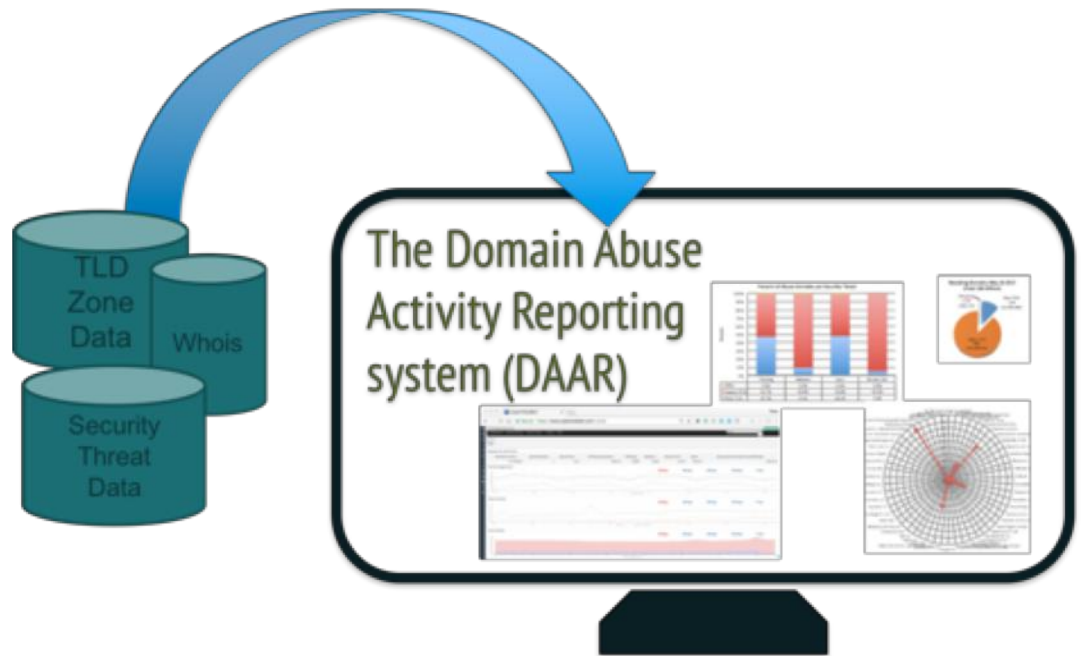
Distribution of Domains with Different Abuse Types in gTLDs



Average Abuse Percentage per gTLD Across All Abuse Types



Where do
We Want
to Go from
Here?



Measuring Abuse

- We are always open to discussion on improvements or other ways the data can be used to help inform discussions around DNS abuse
- Feel free to use daar@icann.org to contact us

Discussions on DNS Abuse at IDS (May 10-11)



DNS-OARC

Domain Name System Operations Analysis and Research Center

12-13 May 2019

Questions?



Thank You



[@icann](https://twitter.com/icann)



facebook.com/icannorg



youtube.com/icannnews



flickr.com/icann



linkedin/company/icann



slideshare/icannpresentations



soundcloud/icann

Contact Info:

DAAR@icann.org

Samaneh.tajali@icann.org

John.crain@icann.org