Domain Abuse Activity Reporting (System)

Carlos Álvarez del Pino
Director of SSR Engagement
Office of the CTO, SSR Team

18 October 2018

DAAR

Introductory

- What is DAAR?
- Objectives

Reputation Data

- Criteria
- Uses
- RBLs
- Data Visualization

Current Status

- Where we're at
- Community participation

What's DAAR?

DAAR

What is the Domain Abuse Activity Reporting Project?

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

How does DAAR differ from other reporting projects?

- Studies all TLD registries and registrars for which we can collect zone and registration data
- Employs a very large set of reputation feeds
- Historical studies
- Studies multiple threats: phishing, botnet, malware, spam
- Scientific approach: unbiased, transparent, reproducible

DAAR & the Open Data Initiative

- Goal of Open Data Initiative is to facilitate access to data that ICANN organization or community creates or curates
- DAAR project uses data from public, open, and commercial sources
  - DNS zone data
  - WHOIS data
  - Certain open source reputation data
  - Certain commercial feeds requiring a license or subscription
- In cases where there are no limitations on redistribution of DAAR-related data, these DAAR project data or reports will be published periodically and included in the Open Data Initiative

Goals

Provide ICANN community with data to support the policy development process

- DAAR project data can be used to
  - Identify threats reported at TLD or registrar level for all TLDs for which we can obtain data
  - Historically track security threats, domain registration activity (adds, deletes) at a TLD or registrar level
  - Help operators understand or consider how to manage their reputations, their anti-abuse programs or their terms of service
  - Study malicious registration behaviors
  - Assist the operational security community by sharing open data or data analyzed by the reporting tools

DAAR Uses TLD Zone Data

- Collects zones for TLDs for registry analytics
  - Any {new, legacy, cc} from which we can get a zone
  - Currently gTLDs. Some ccTLD expressed interest in being added during ICANN 58, Copenhagen
- Currently, system collects zones from 1241 TLDs
  - Approximately 195 million domains
  - DAAR project uses publicly available methods to collect zone data (Centralized Zone Data Service, zone transfer)

Reputation Data

DAAR Is Not A Blocklist Service

- DAAR counts unique abusive domains
  - A domain in an RBL is counted only once
- DAAR uses multiple RBLs to:
  - Generate daily counts of domains associated with phishing, malware distribution, C&D and spam
  - Calculate daily and accumulated total counts of malicious domains
  - Calculate recently listed domains (monthly, up to 1yr))
  - Create histograms, graphics, day in the life of, etc.

DAAR reflects how external organizations see the domain ecosystem

DAAR – Criteria to choose RBLs

- Must classify threats as DAAR does
- There must be evidence that OpSec community trusts RBLs due to accuracy and clarity of processes
- Positive reputation in academic literature
- Widely adopted among security community:
  - RBLs must be incorporated by commercial security systems
  - Must be used by network operators for user and device protection
  - Must be used by email and messaging providers to protect users

RBLs: Protecting Users

- RBL use is almost omnipresent
- The block much more than just unwanted email
- RBLs in browsers:
  - Google Chrome uses APWG and Safe Browsing URL Data
- RBLs in the Cloud and in content delivery systems
  - Akamai uses SURBL, Symantec, ThreatSTOP, and specific RBLs
  - AWS Web Application Firewall uses RBLs to block abuse and volumetric attacks
  - Google Safe Browsing blocks malicious URLs and fraud via AdWords

RBLs: Private Network Operators

- RBLs in commercial firewalls, UTM (Unified Threat Management) devices
  - Palo Alto Networks, Barracuda Networks, SonicWall, Check Point, Fortigate, Cisco IronPort, and WatchGuard
  - TitanHQ SpamTitan, Sophos UTM, and Proofpoint
  - RBLs mencionadas: Spamhaus, SURBL, SpamCop, Invaluement, abuse.ch, Open ORDBL, Spam and Open Relay Blocking System (SORBS), Squidblacklist.org,
- RBLs in corporate email/messaging systems:
  - Spam Solutions de GFI MailEssentials, SpamAssassin, y Vamsoft ORF incluye Spamhaus o SpamCop RBLs disponibles para Microsoft Exchange
- RBLs and email providers (ESPs):
  - Amazon Simple Email Service RBL or DNS block lists
  - Observar ESPMail Exchange (MX) and Sender Policy Framework (SPF) resource records

Current RBLs

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

SpamAssassin: malware URLs list
Carbon Black Malicious Domains
Postfix MTA
Squid Web proxy blocklist
Symantec Email Security for SMTP
Symantec Web Security
Firekeeper
DansGuardian
ClamAV Virus blocklist
Mozilla Firefox Adblock
Smoothwall
MailWasher

RBLs in the Academia: Trust in RBLs

Partial list of academic studies and RBL citations (RBLs that feed DAAR) :

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

Does DAAR Identify All Abuse?

- No reputation provider can see all the abuse
  - Each is catching only some (what they see)
- Providers look for different types of abuse, use different methods or infrastructures
- Some lists are big and some are small.
  - The smaller the list, the less % overlap it might have with a larger list

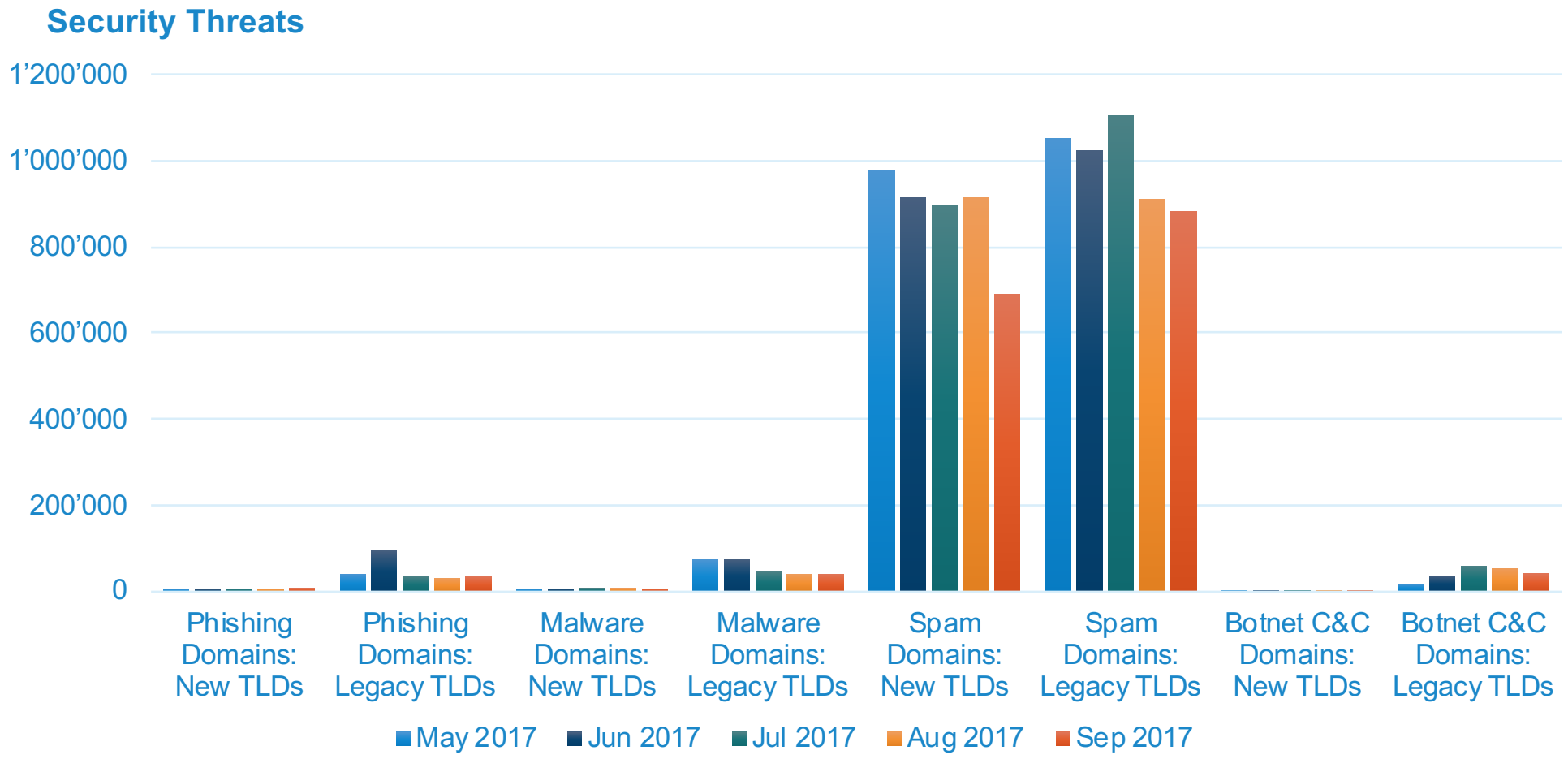
Why does DAAR include spam data?

- ICANN's GAC expressed its interest in domains associated with spam as a security threat (Hyderabad Communiqué).
- Why?
  - Most spam is sent via illegal means (i.e., botnets).
  - Spam is, long ago, not only associated with email sending. Link spam, spamdexing, tweet spam, messaging spam (text/SMS)
  - Spam is a very relevant distribution means of other threats: Avalance.
- DAAR counts domains found in the body of spam email

- IMPORTANT: The reputation of spam domains influences how security administrators or email services filter traffic

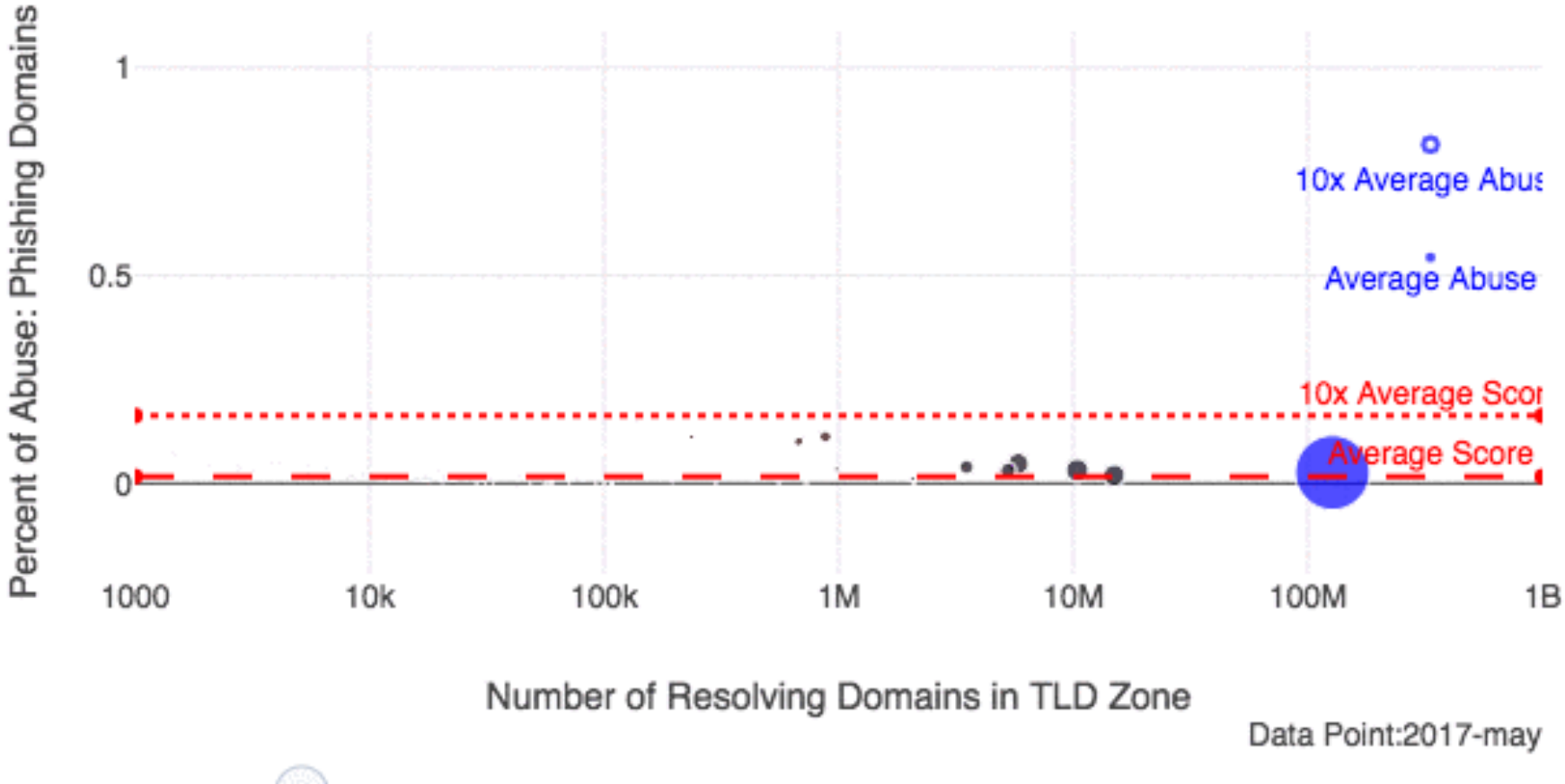
Data Visualization

Data Set: All gTLDs with at least one reported domain

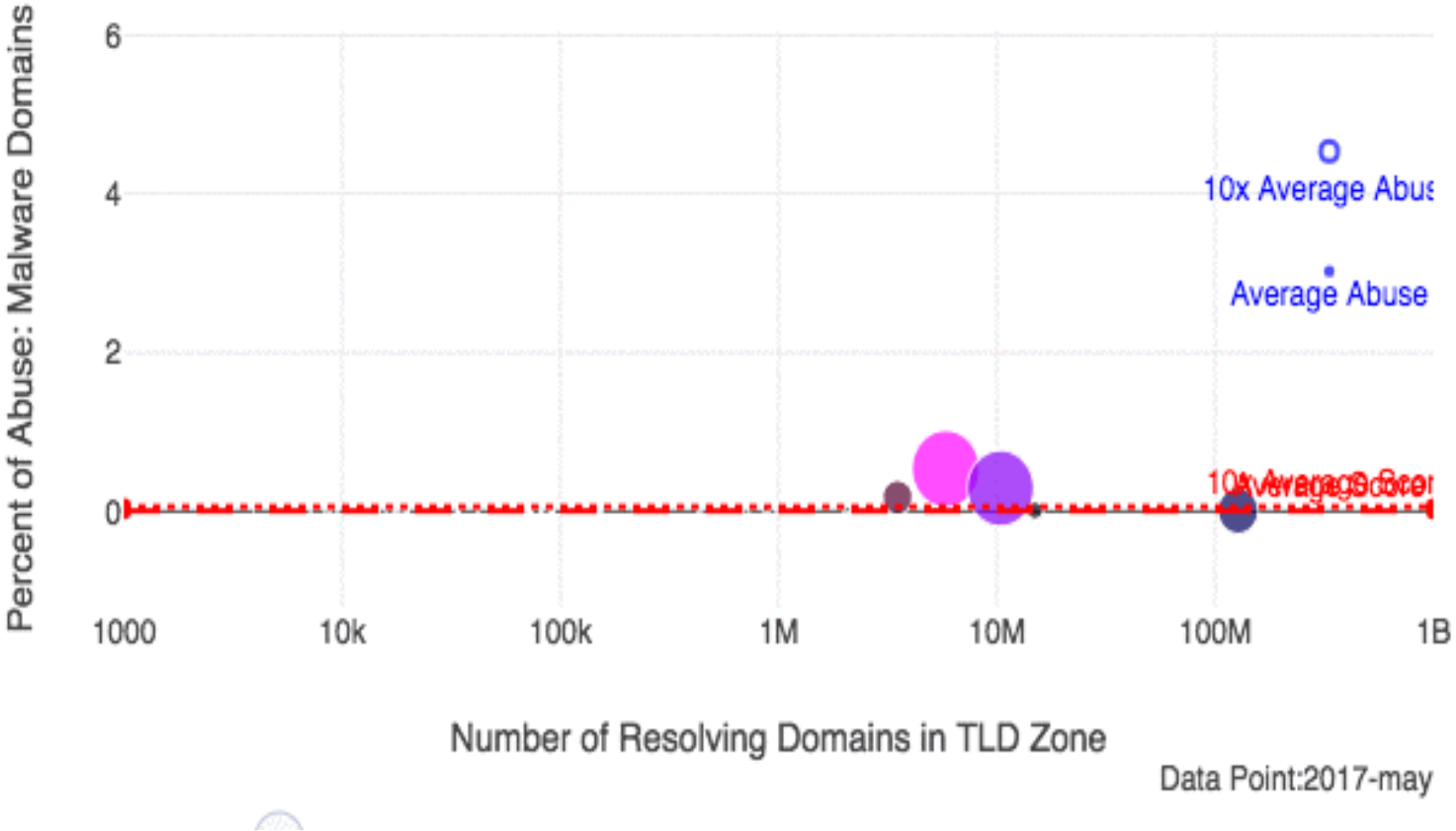


(Sep/17)

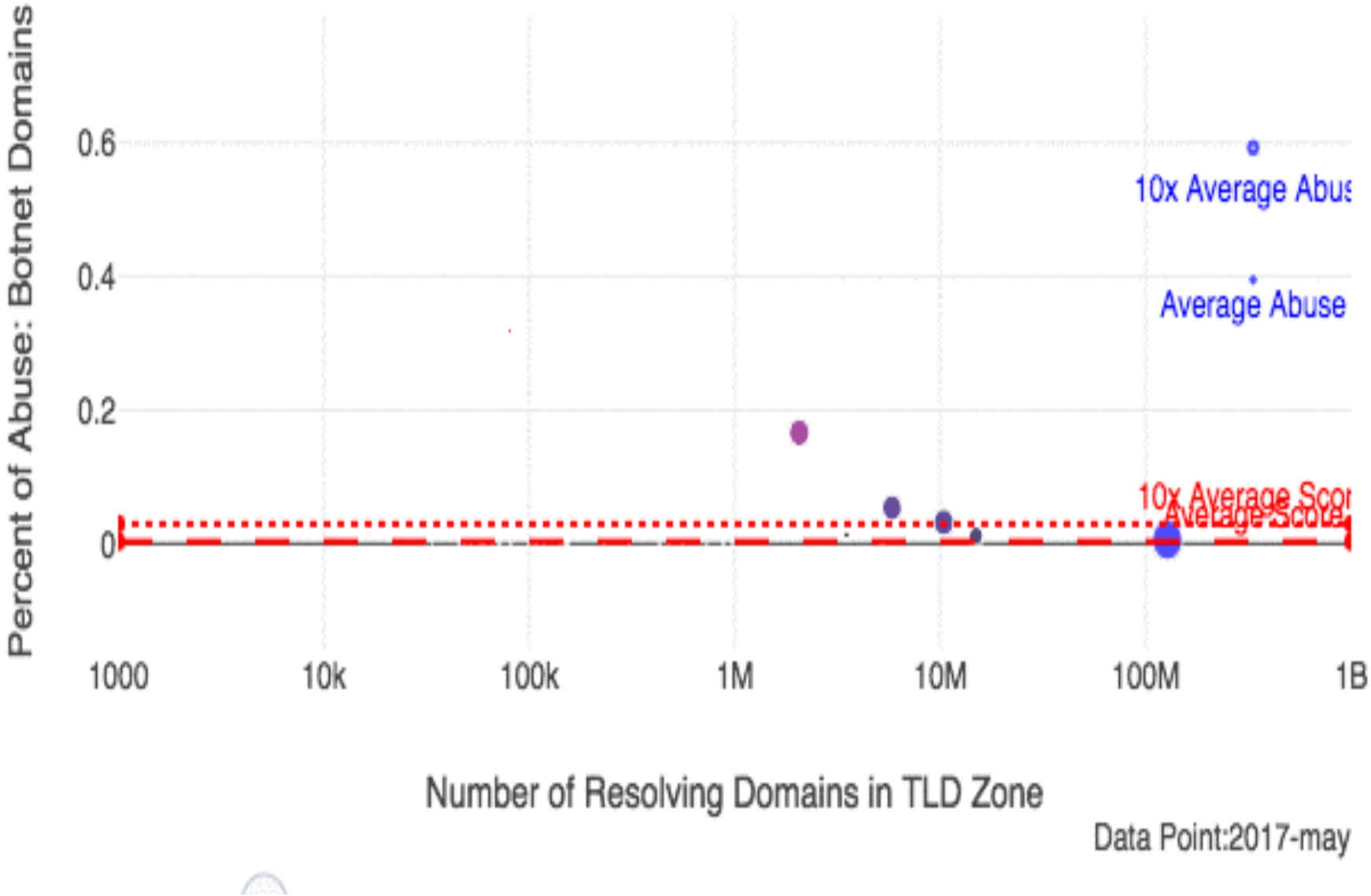
End of month : Phishing, abuse %



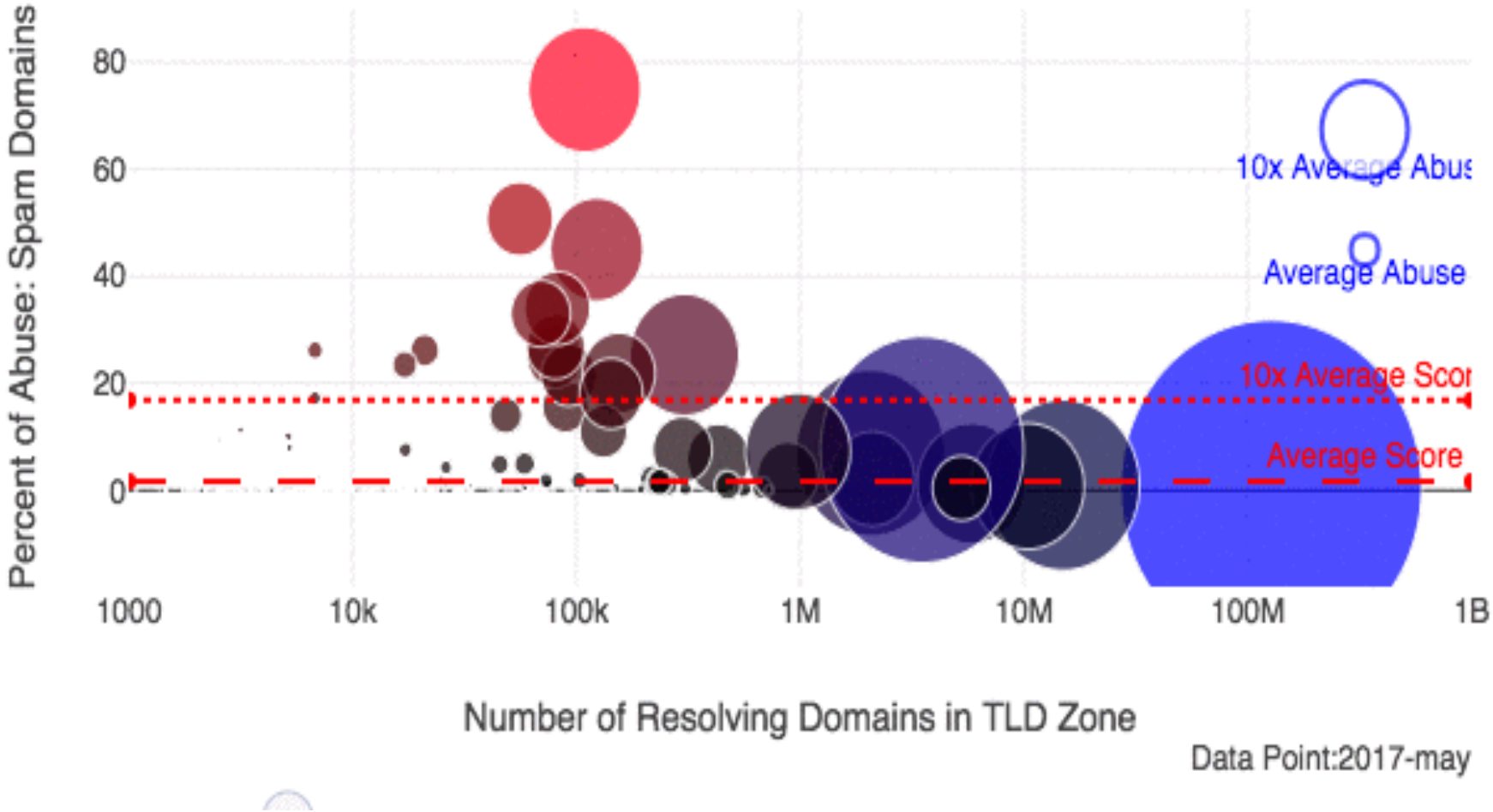
End of month snapshot: Malware, abuse %



End of month snapshot: Botnet (C2), abuse %



End of month snapshot: Spam, abuse %



Project Status

- https://www.icann.org/octo-ssr/daar

- Published Methodology:
  - https://www.icann.org/news/announcement-2018-07-20-en
  - Reviews: Marcus Ranum and John Bambenek

daar@icann.org

Thank You and Questions

Visit us at **icann.org**

 @icann

 facebook.com/icannorg

 youtube.com/icannnews

 flickr.com/icann

 linkedin/company/icann

 slideshare/icannpresentations

 soundcloud/icann