REMEDI3S-TLD: Reputation Metrics Design to Improve Intermediary Incentives for Security of TLDs

A project in collaboration with SIDN and NCSC

Maciej Korczyński
Delft University of Technology
Contact: maciej.korczynski@tudelft.nl

DHPA Techday

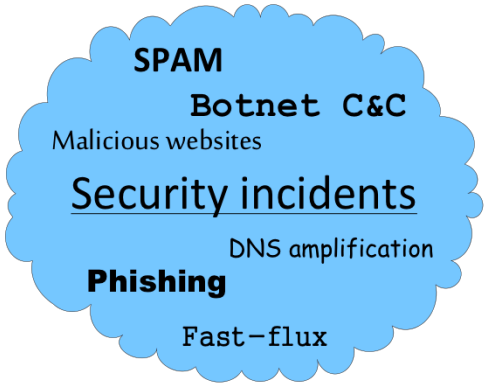
21 May 2015, The Hague

TU Delft SIDN SIDN labs

REMEDI3S-TLD

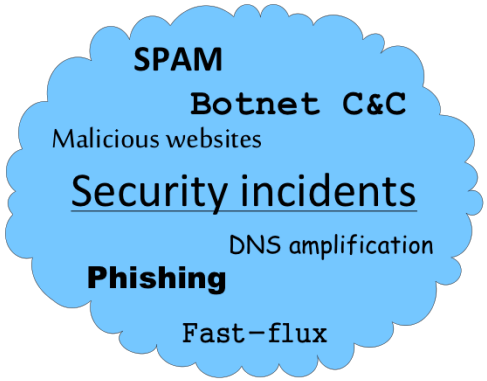
TU Delft SIDN SIDN labs

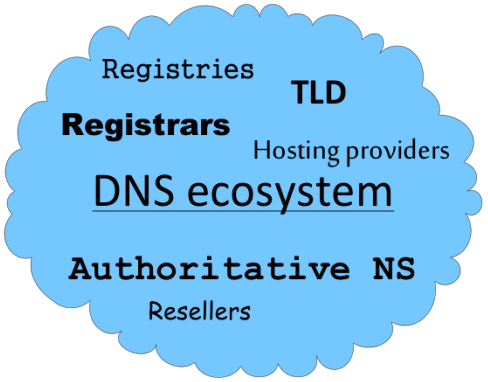
REMEDI3S-TLD



TU Delft SIDN SIDN labs

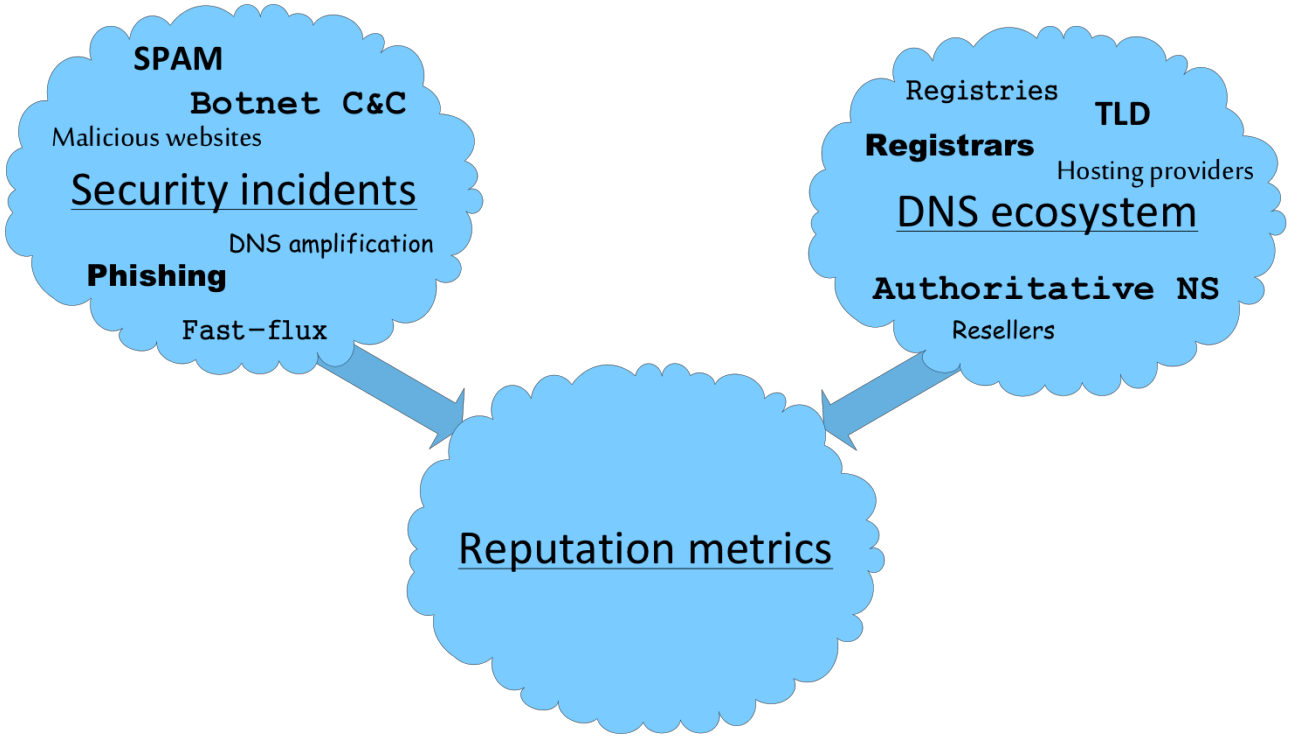
REMEDI3S-TLD





TU Delft SIDN SIDN labs

REMEDI3S-TLD



TU Delft SIDN SIDN labs

Agenda

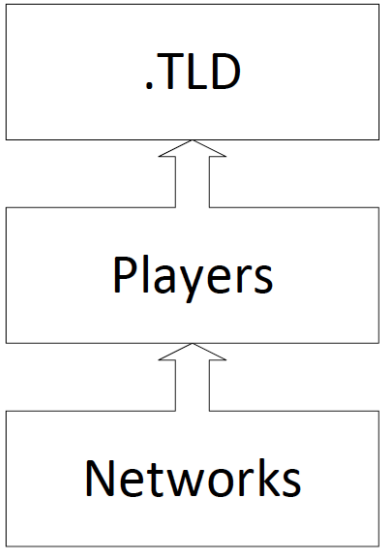
- REMEDI3S-TLD
- Security incidents
- Types of security metrics
- Security metrics for TLDs
- Security metrics for hosting providers
- Practical application
- Summary

Security incidents

- Blacklists
  - APWG
  - Shadowserver (botnet C&C, Sandbox URLs, etc.)
  - ESET, Sophos, Fortinet
  - Google's Safe browsing appeals
  - Malware Must Die
  - Phishtank
  - Zeus tracker
  - Dutch child pornography hotline
  - Etc.
- Farsight security (dns-db)

Types of security metrics

- Different layers of security metrics:
  - Top Level Domains (TLDs)
  - Market players related to the TLD (infrastructure providers): registrars, hosting providers, DNS service providers
  - Network resources managed by each of the players, such as resolvers, name servers



Security metrics for TLDs

- Size estimate for different market players, e.g. TLDs
  - Problem: access to zone files of all TLDs
  - Solution: zone files, APWG reports, DNS-DB

Security metrics for TLDs

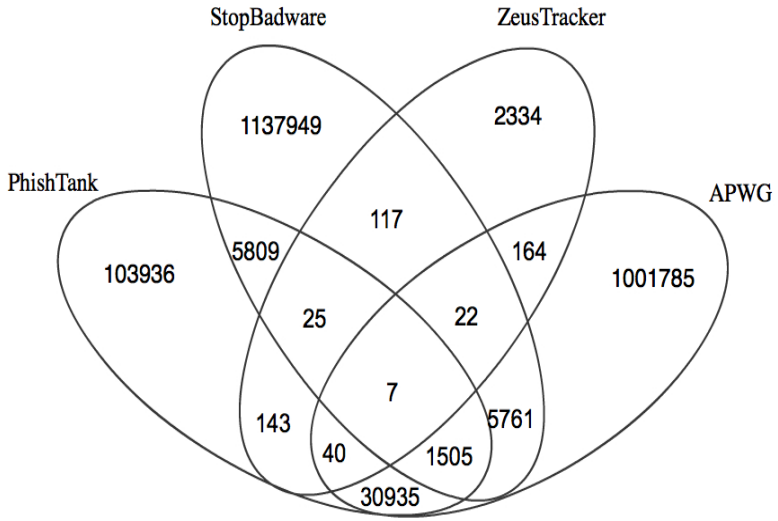
- Type of reputation metrics
  - Problem: estimation of the amount of badness
  - Solutions (TLDs):
    a) Number of unique domains
    b) Number of FQDN
    c) Number of URLs

Security metrics for TLDs

- Type of reputation metrics
  - Problem: up-times of maliciously registered/compromised domains
  - Solutions:
    a) DNS-based scanner
    b) Content-based scanner

Results

- Estimation of the amount of badness for TLD
  - Datasets: suitability, coverage



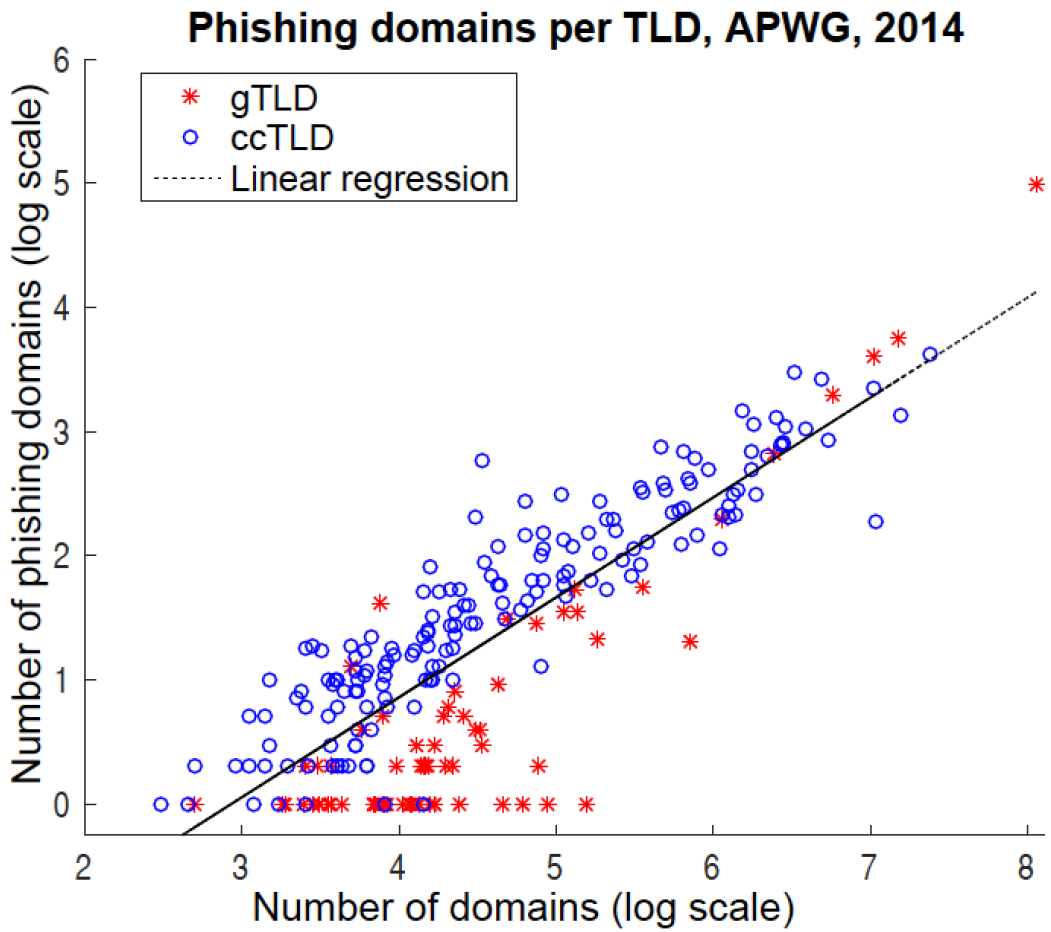
Dataset	# domains	# FQDNs	# URLs
StopBadware	760035	1137949	2357260
APWG	150525	1001785	10474045
Phishtank	83206	103936	—
ZeusTracker	2220	2334	—

Table 1: Dataset statistics: unique 2nd– and 3rd–level domains, FQDNs, and URLs for the 4 datasets from 2014.

TUDelft SIDN SIDN labs

Results

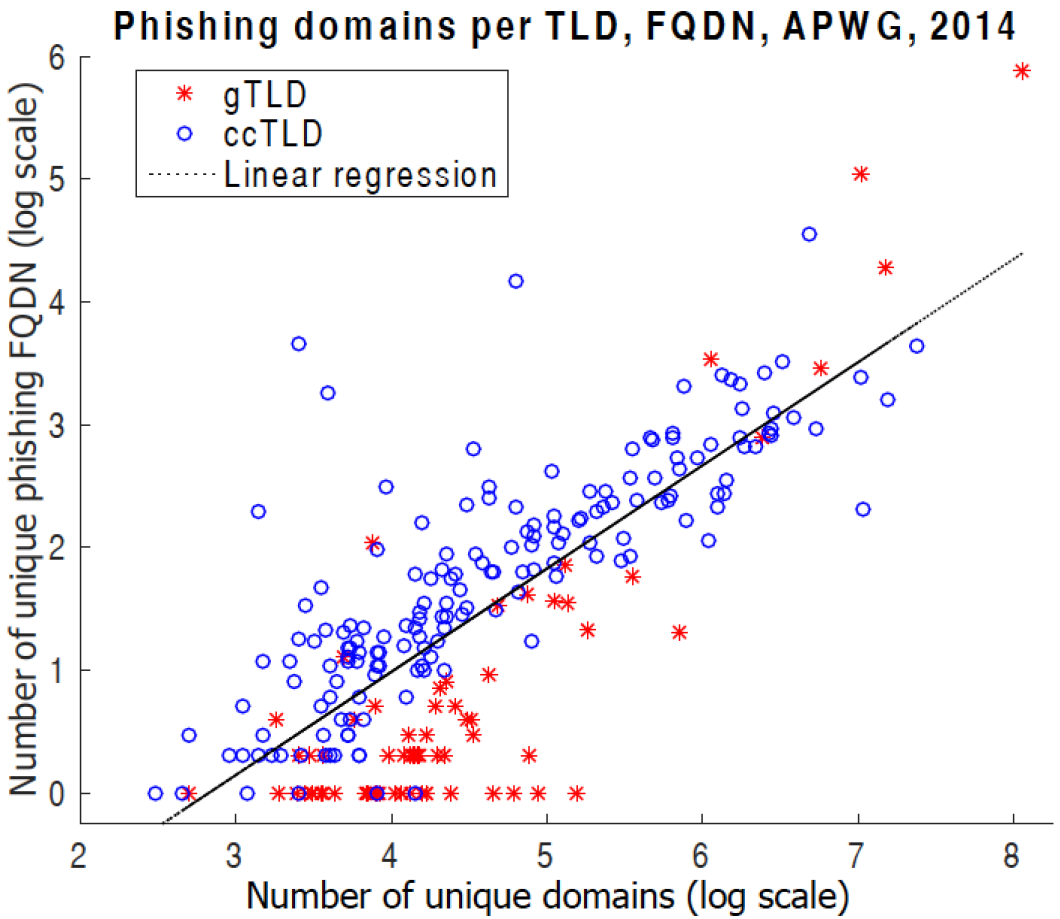
- Estimation of the amount of badness



TU Delft SIDN SIDN labs

Results

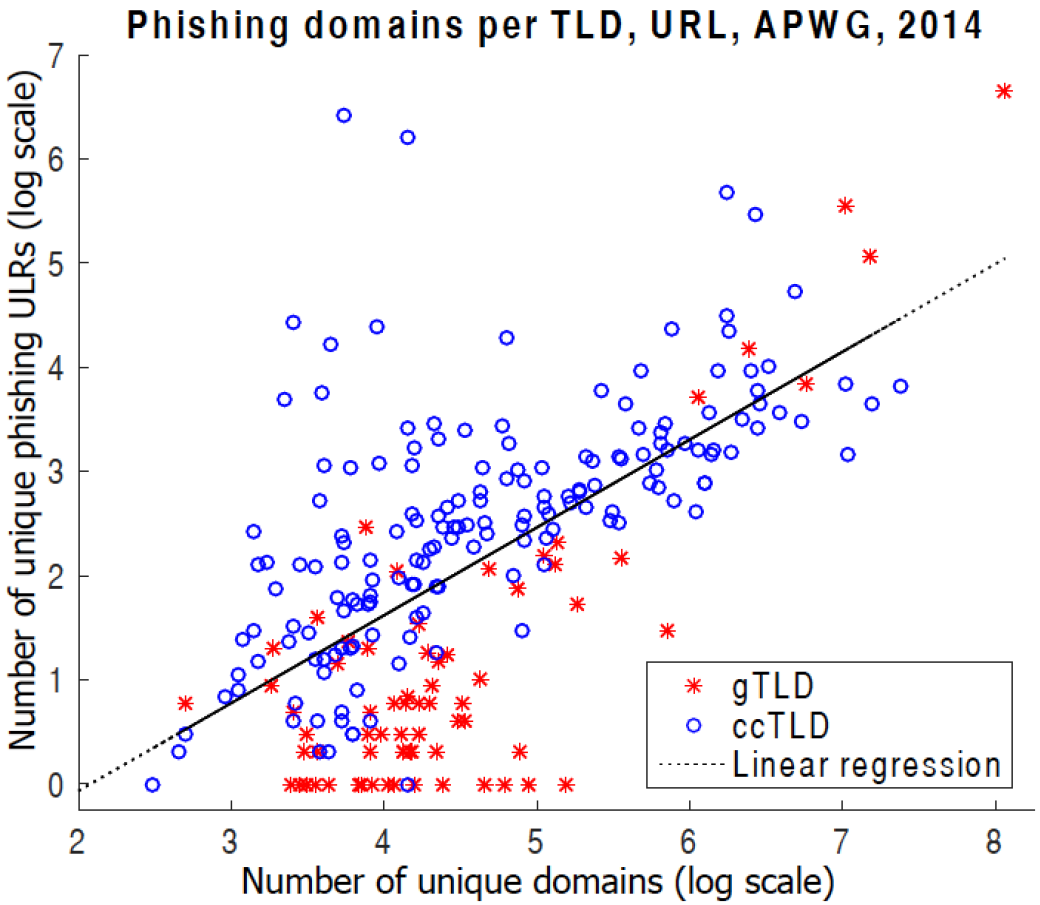
- Estimation of the amount of badness



TU Delft SIDN SIDN labs

Results

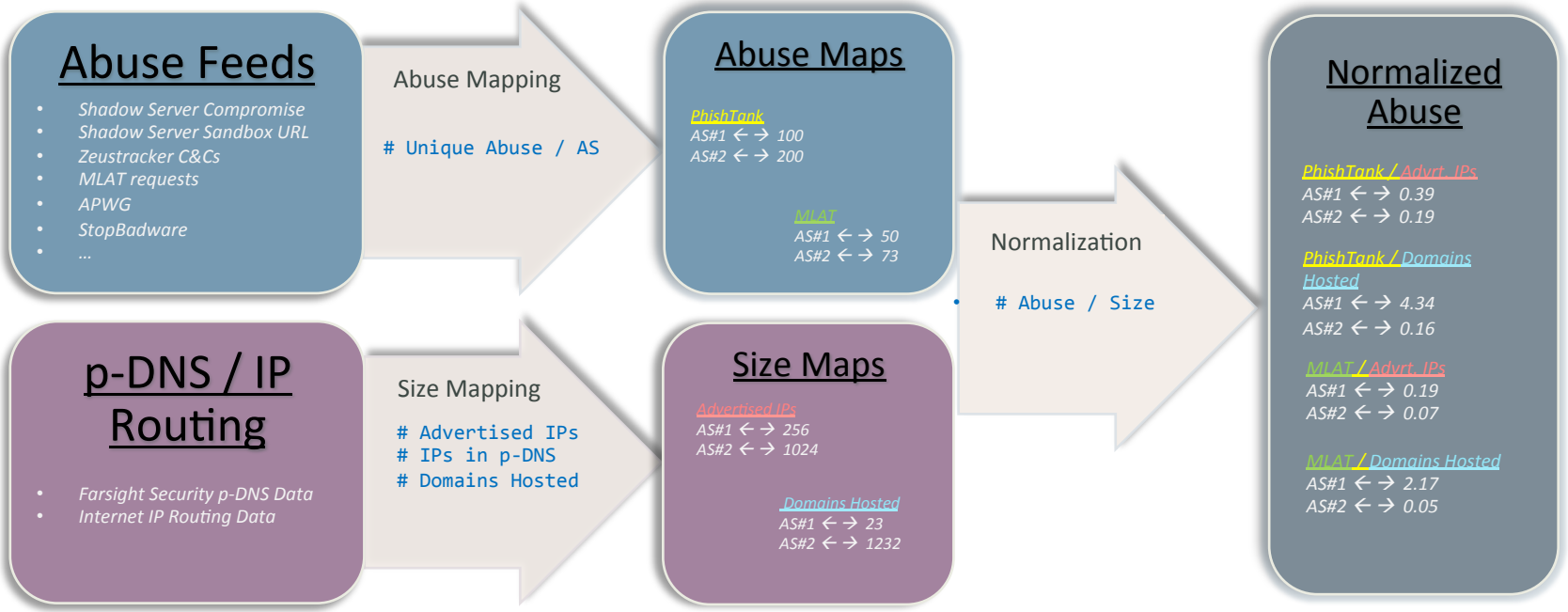
- Estimation of the amount of badness



TU Delft SIDN SIDN labs

Security metrics for hosting providers

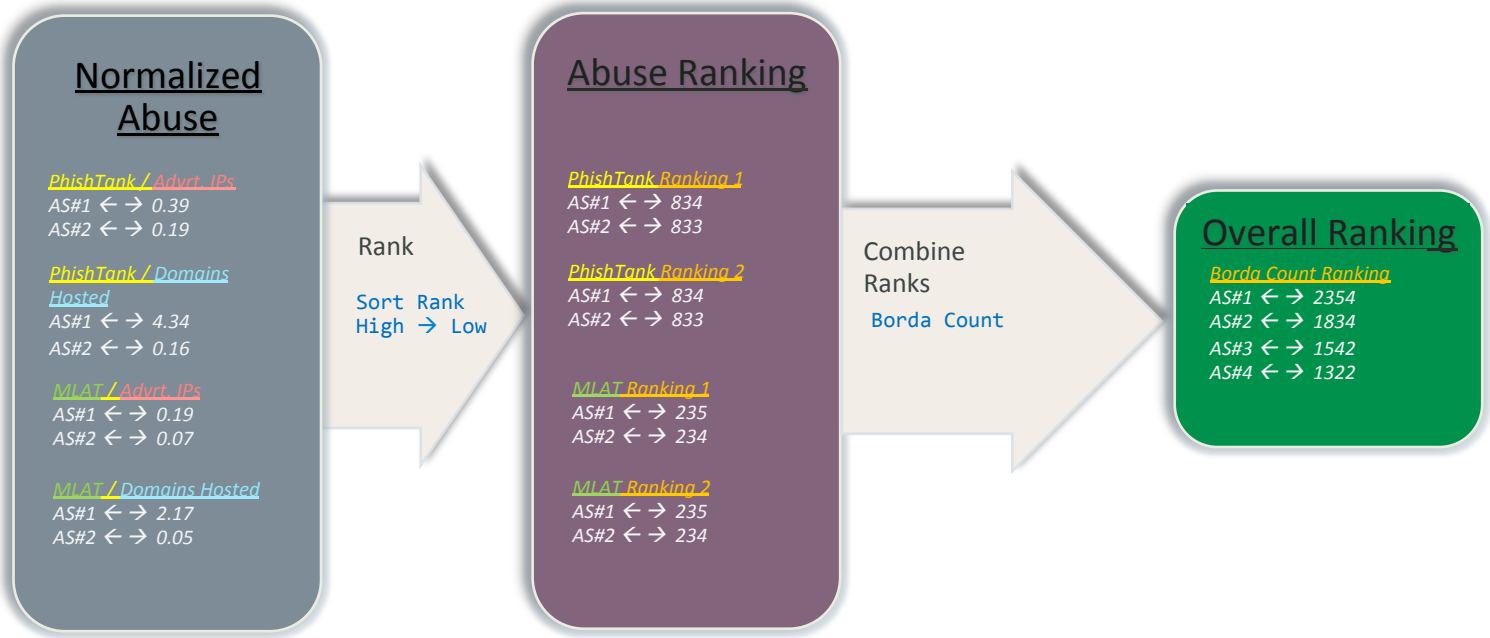
1. Count badness per AS across different data sources
2. Normalize for the size of the AS (in 3 ways)



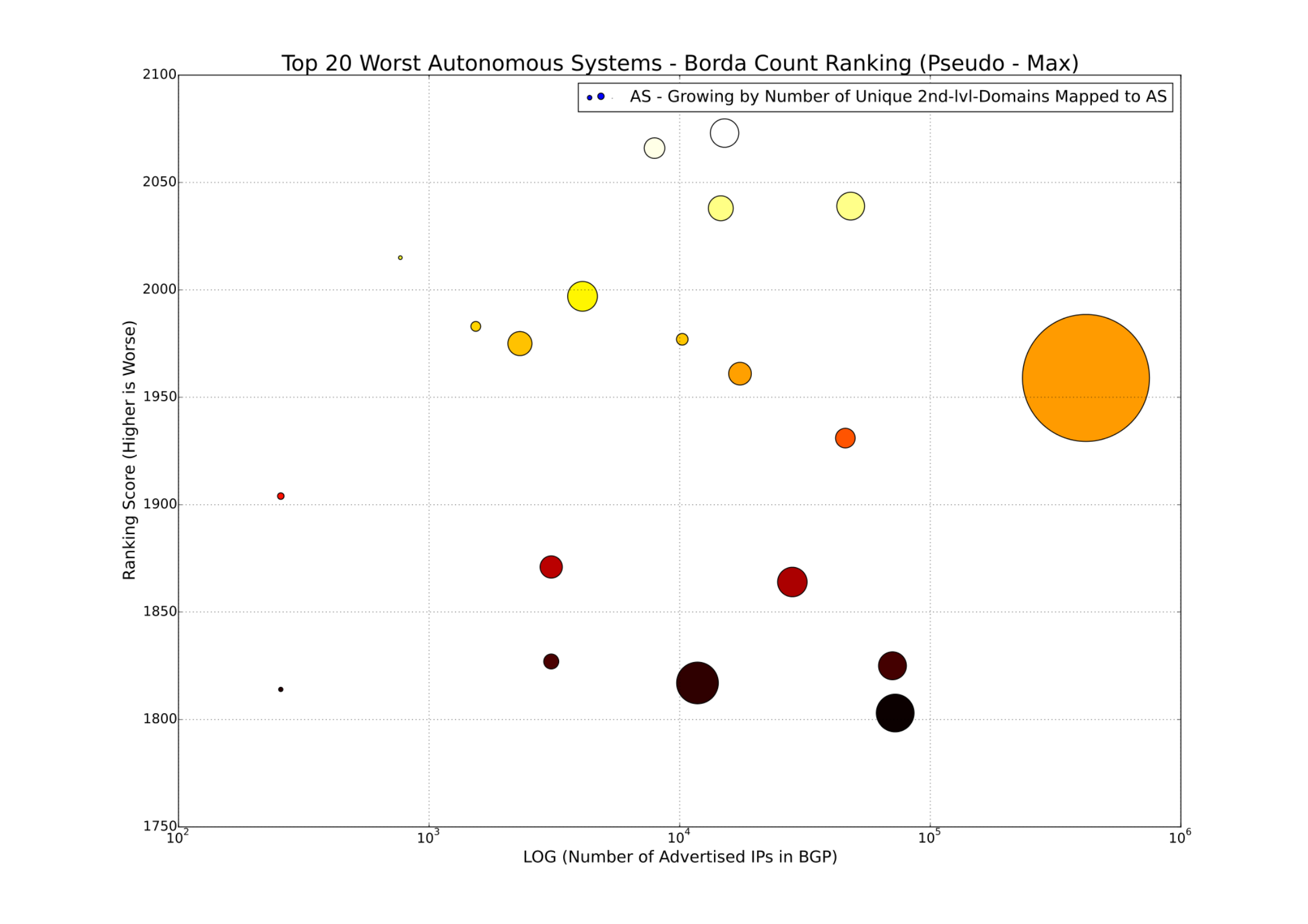
TU Delft SIDN SIDN labs

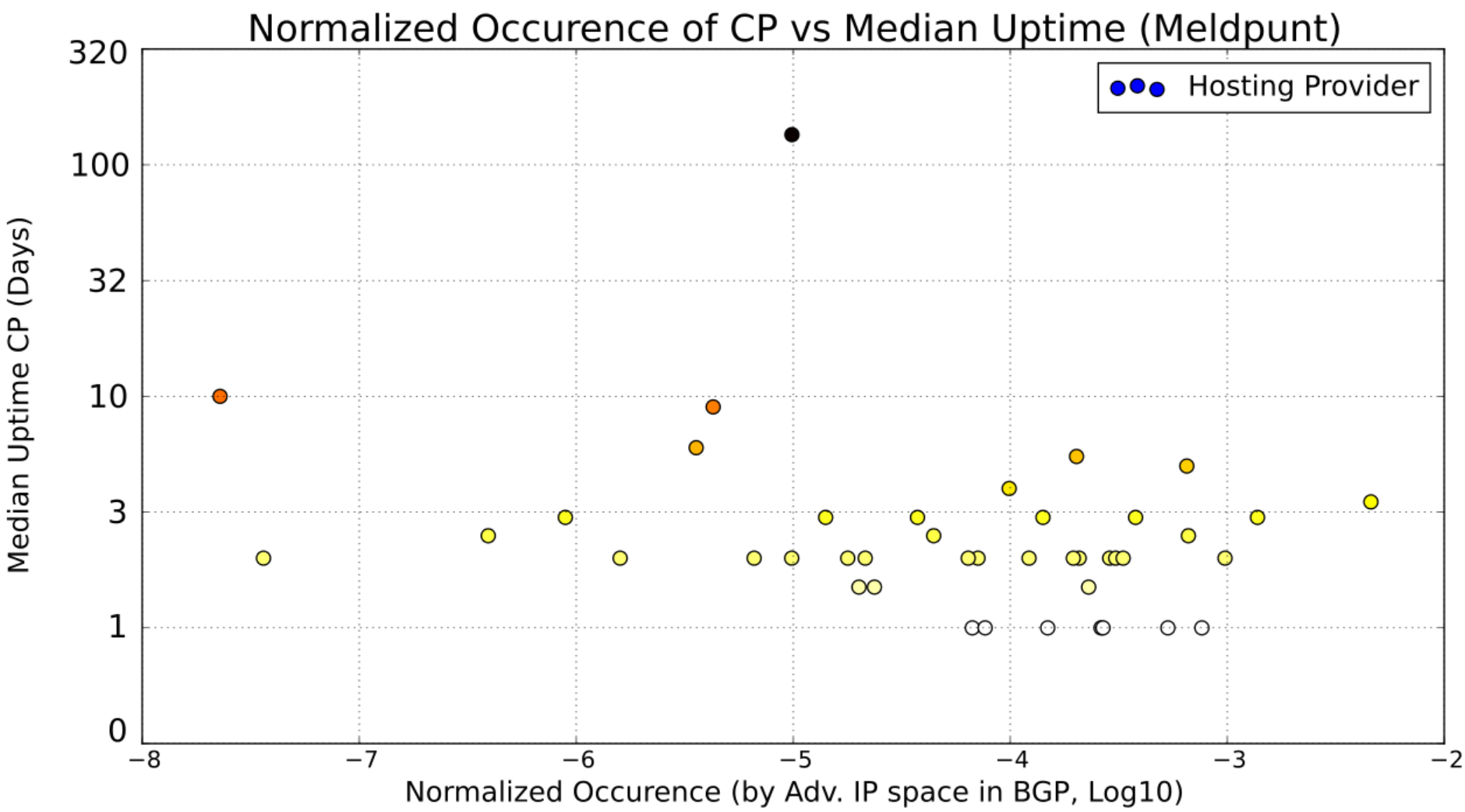
Security metrics for hosting providers

3. Rank ASes on amount of badness
4. Aggregate rankings (Borda count)
5. Identify ASes with consistently high concentrations of badness



TU Delft SIDN SIDN labs


Top 20 Worst Autonomous Systems - Borda Count Ranking (Pseudo - Max)
AS - Growing by Number of Unique 2nd-lvl-Domains Mapped to AS
Ranking Score (Higher is Worse)
LOG (Number of Advertised IPs in BGP)
2100
2050
2000
1950
1900
1850
1800
1750
10^2
10^3
10^4
10^5
10^6


Normalized Occurence of CP vs Median Uptime (Meldpunt)
Hosting Provider
Median Uptime CP (Days)
320
100
32
10
3
1
0
−8
−7
−6
−5
−4
−3
−2
Normalized Occurence (by Adv. IP space in BGP, Log10)

Practical application

- Incentive structures that drive the DNS ecosystem
- "Clean Netherlands": Enhance self cleansing ability of the Dutch hosting market by
    - promoting best practices and awareness
    - pressuring the rotten apples

Summary

- REMEDI3S-TLD
- Security metrics for TLDs
- Security metrics for hosting providers
- Practical application

ACKNOWLEDGEMENTS

The research leading to these results
was funded by SIDN (www.sidn.nl)

TU Delft SIDN SIDN labs