

Trends in Abuse: New and Legacy gTLDs

Maciej Korczyński, TU Delft, Grenoble INP

Maarten Wullink, SIDN Labs

Brian Aitchison, ICANN

Drew Bagley, Secure Domain Foundation

Thursday, September 5, 2017

Toronto, Canada

Attendees Reminders:



What occurs in a M³AAWG meeting cannot be shared outside the membership

- Attendees can blog, tweet and post on either your personal or business social media account about the **selected, pre-approved sessions where we show a slide indicating that social media posting is allowed.** Please reference @maawg or #m3aawg41 where we are also tweeting.
- In all cases, respect M³AAWG anonymity: No publishing people or company names, except as cited on the official M³AAWG channels: @maawg, facebook.com/maawg, plus.google.com/+MAAWG
- No use of Wireshark or similar products on the M³AAWG network
- No photography - No video - No audio recording
- Any exception requires written permission from the Executive Director and may require permission from the session members
- All meeting attendees must wear and have their M³AAWG badge visible at all times during the meeting
- Please silence all electronic devices; be courteous to those listening to the presentations
- DO NOT LEAVE YOUR BELONGINGS UNATTENDED. Be aware and cautious at all times

Treat all attendees respectfully in and out of sessions. No less will be tolerated. Please review our meeting Conduct Policy at <https://www.m3aawg.org/conduct-policy>

For questions, please contact Jerry Upton at: jerry.upton@m3aawg.org

Reminders for Our Worldwide Friends

*All meeting content is confidential: No photos, no video, no recording.
See staff with questions.*



L'ensemble du contenu de la réunion est confidentiel : les photos, vidéos et enregistrements sont interdits. Pour toute question, demandez conseil au personnel.



Todo el contenido de la reunión es confidencial: No está permitido sacar fotografías ni grabar vídeo o audio. Consulte con el personal si tiene alguna pregunta.



Der gesamte Inhalt des Meetings ist vertraulich: Keine Fotos, kein Video, keine Tonaufzeichnung. Bei Fragen wenden Sie sich an die Mitarbeiter.



会議の内容はすべて機密扱いです。写真やビデオの撮影、録音は禁止されています。質問がある方は、スタッフまでご連絡ください。



所有会议内容均为保密信息：禁止拍照、录像、录音。如有疑问，请咨询职员。



회의에서 다루는 모든 내용은 기밀입니다. 사진 및 동영상 촬영과 녹음은 금지됩니다. 질문이 있으시면 직원에게 문의해 주십시오.



Материалы совещаний конфиденциальны. Фотографирование, видео- и звукозапись запрещены. В случае возникновения вопросов обращайтесь к сотрудникам.

– Social Media Posting Allowed –

Tweeter, Facebook, LinkedIn, other social media posts are welcomed in this session if you:

- Only post comments made by the speakers or panelists
- Do not post comments or questions from the audience (but you can share the speakers' responses to questions)
- Do not post the name, position or company of other meeting attendees
- Do not post conversations with attendees
- M³AAWG is not a deliverability conference; we are:
 - An industry working group meeting
 - An anti-abuse conference, or
 - A gathering of security experts
- All of the M³AAWG Membership, Trademarks and Logo guidelines apply (<https://www.m3aawg.org/members/how-promote-m3aawg#TrademarkGuidelines>)
- Appreciate a shout out to @maawg and #m3aawg41

Code of Conduct



M³AAWG is dedicated to making our meetings and business open to all members and guests and to making it a safe place for all. We do not tolerate harassment of any kind.

We insist that all participants, attendees and meeting staff adhere to a civil demeanor at all times. This includes refraining from inappropriate language, comments and behavior, in person or by electronic communications and/or public or semi-public social media. In accordance with applicable law, M³AAWG prohibits sexual harassment and harassment because of race, color, gender, age, religion, disability, sexual orientation or any other basis protected by federal, state or local law.

Participants, attendees and meeting staff who are being harassed, intimidated, or are dealing with otherwise improper behavior are encouraged to report it immediately to the Executive Director or a Board member without fear of repercussion.

Alternate methods of reporting issues include: contacts listed on the back of your badge, email to the Executive Director, jerry.upton@m3aawg.org, or if needed, calling the local police department.

Anyone who is found to be in violation of this policy may be handled in any one or more of these methods, depending on the offense: Warning, Expulsion, Contacting of employer, or Contacting the police or other legal authorities. Actions stronger than a warning will be taken at the discretion of the M³AAWG Board of Directors.

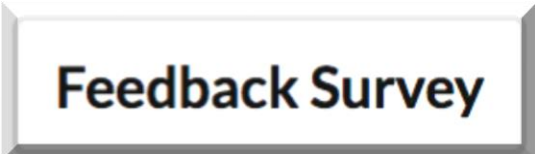
M³AAWG reserves the right to remove any participant or attendee at any time for any reason.

The policy also extends outside of the meeting rooms to include all areas of the meeting hotel and social gatherings sponsored by M³AAWG or M³AAWG member organizations.

Note: You can download this file at <https://www.m3aawg.org/conduct-policy>

Session Feedback

Please share your comments on this session with M³AAWG
– good, could-be-better or new ideas –
to help improve our meetings

Click on the session title in SCHED then
use the  button above the description

Thanks! Your comments are appreciated.

Agenda



- ⦿ Introduction from the ICANN organization: Background of Study
- ⦿ Presentation from SIDN and Delft University of Technology
- ⦿ Q & A

Study Background

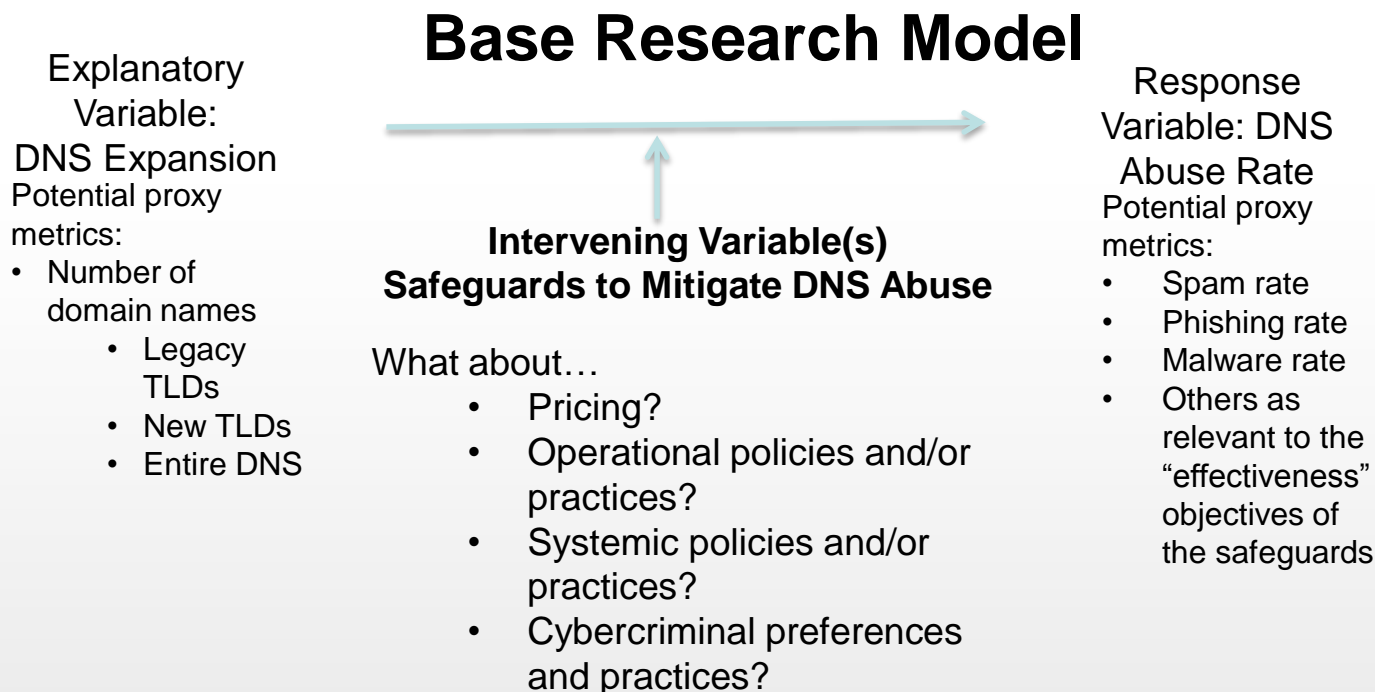
◎ 2009: Mitigating Malicious Conduct: New gTLD Program Explanatory Memorandum

Question	Recommendation(s)
1) How do we ensure that bad actors do not run registries?	1. Vet registry operators
2) How do we ensure integrity and utility of registry information?	2. Require DNSSEC Deployment 3. Prohibit “wildcarding” 4. Encourage removal of “orphan glue” records
3) How do we ensure more focused efforts on combating identified abuse?	5. Require “Thick” WHOIS records 6. Centralize Zone File access 7. Document registry- and registrar-level abuse contacts and policies 8. Provide an expedited registry security request process
4) How do we provide an enhanced control framework for TLDs with intrinsic potential for malicious conduct?	9. Create a draft framework for a high security zone verification program

Study Background (cont'd)

◎ 2016: New gTLD Program Safeguards Against DNS Abuse: Revised Report

- ◎ Research aid to Competition, Consumer Trust, and Consumer Choice Review Team
- ◎ How to measure effectiveness of safeguards?



Study Background (cont'd)



◎ 2016-2017: Competition, Consumer Trust, and Consumer Choice Review Team

- ◎ Affirmation of Commitments (AoC) specified that “malicious abuse issues” be addressed in expansion of top-level domain space
- ◎ CCT-RT mandated by AoC to examine “effectiveness of...safeguards put in place to mitigate issues involved in...the expansion [of the top-level domain space]”
- ◎ Required comprehensive descriptive statistics as **baseline measure** of abuse rates in new compared to legacy gTLDs in order to gauge safeguard effectiveness
- ◎ Also serves as proxy for “Trust”, i.e. changes in abuse rate → changes in trust
- ◎ CCT-RT Draft Report recommends ongoing DNS abuse measurement

Study

Statistical Analysis of DNS Abuse in gTLDs (SADAG)

Consortium: SIDN and TU Delft

Requested by: Competition, Consumer Trust, and
Consumer Choice Review Team

Goal

- Comprehensive statistical comparison of rates of DNS abuse in new and legacy gTLDs
 - Spam
 - Phishing
 - Malware
- Statistical analysis of potential abuse drivers

Motivation

- New Generic Top-Level Domain (gTLD) Program enabled hundreds of new generic top-level domains

Blacklists

- Anti Phishing Working Group
 - Phishing URLs
- StopBadware
 - Malware URLs
- SURBL (4 blacklists)
 - Phishing domains
 - Spam domains
 - Malware domains

Blacklists

- Spamhaus
 - Spam domains
- CleanMX (3 feeds)
 - Phishing URLs
 - Malware URLs
 - Defaced URLs
- Secure Domain Foundation
 - Phishing URLs
 - Malware URLs

WHOIS data

- WHOIS XML API
 - All new gTLDs
 - Subset of legacy gTLDs
- DomainTools
 - Providing missing domains

Domain data

- Zone files
 - Per gTLD
 - Per day
 - 3-year period

Active Web & DNS Scan

- Scanned
 - All new gTLDs
 - Sample of legacy gTLDs

Registry (ICANN)

- Sunrise periods
- Registry operators (parent companies of registry operators)

Security Metrics

- Distribution of malicious content: *
 - Number of unique domains
E.g. **malicious.com**

* **“Reputation Metrics Design to Improve Intermediary Incentives for Security of TLDs”**,
Maciej Korczyński, Samaneh Tajalizadehkhoob, Arman Noroozian, Maarten Wullink, Cristian Hesselman,
and Michel van Eeten, in the *IEEE European Symposium on Security and Privacy (Euro S&P)*

Security Metrics

- Distribution of malicious content:
 - Number of unique domains
E.g. malicious.com
 - Number of FQDNs
E.g. **connect.secure.wellsfargo.malicious.com,**
bankofamerica.com.malicious.com, (...)

* **“Reputation Metrics Design to Improve Intermediary Incentives for Security of TLDs”**,
Maciej Korczyński, Samaneh Tajalizadehkhoob, Arman Noroozian, Maarten Wullink, Cristian Hesselman,
and Michel van Eeten, in the *IEEE European Symposium on Security and Privacy (Euro S&P)*

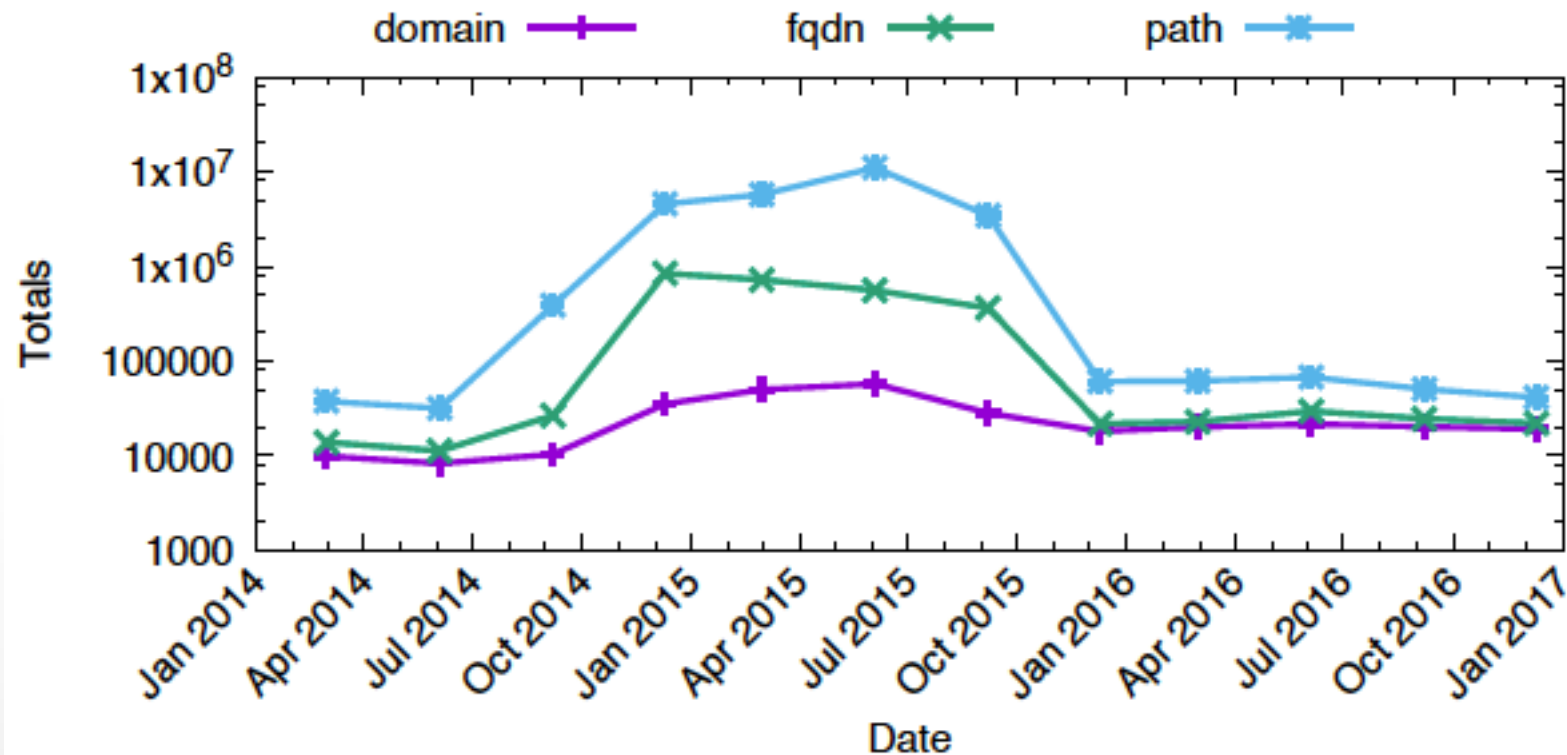
Security Metrics

- Distribution of malicious content:
 - Number of unique domains
E.g. malicious.com
 - Number of FQDNs
E.g. connect.secure.wellsfargo.malicious.com,
bankofamerica.com.malicious.com, (...)
 - Number of URLs
E.g. **malicious.com/wp-content/file.php,**
malicious.com/wp-content/gate.php, (...)

* **“Reputation Metrics Design to Improve Intermediary Incentives for Security of TLDs”**,
Maciej Korczyński, Samaneh Tajalizadehkhoob, Arman Noroozian, Maarten Wullink, Cristian Hesselman,
and Michel van Eeten, in the *IEEE European Symposium on Security and Privacy (Euro S&P)*

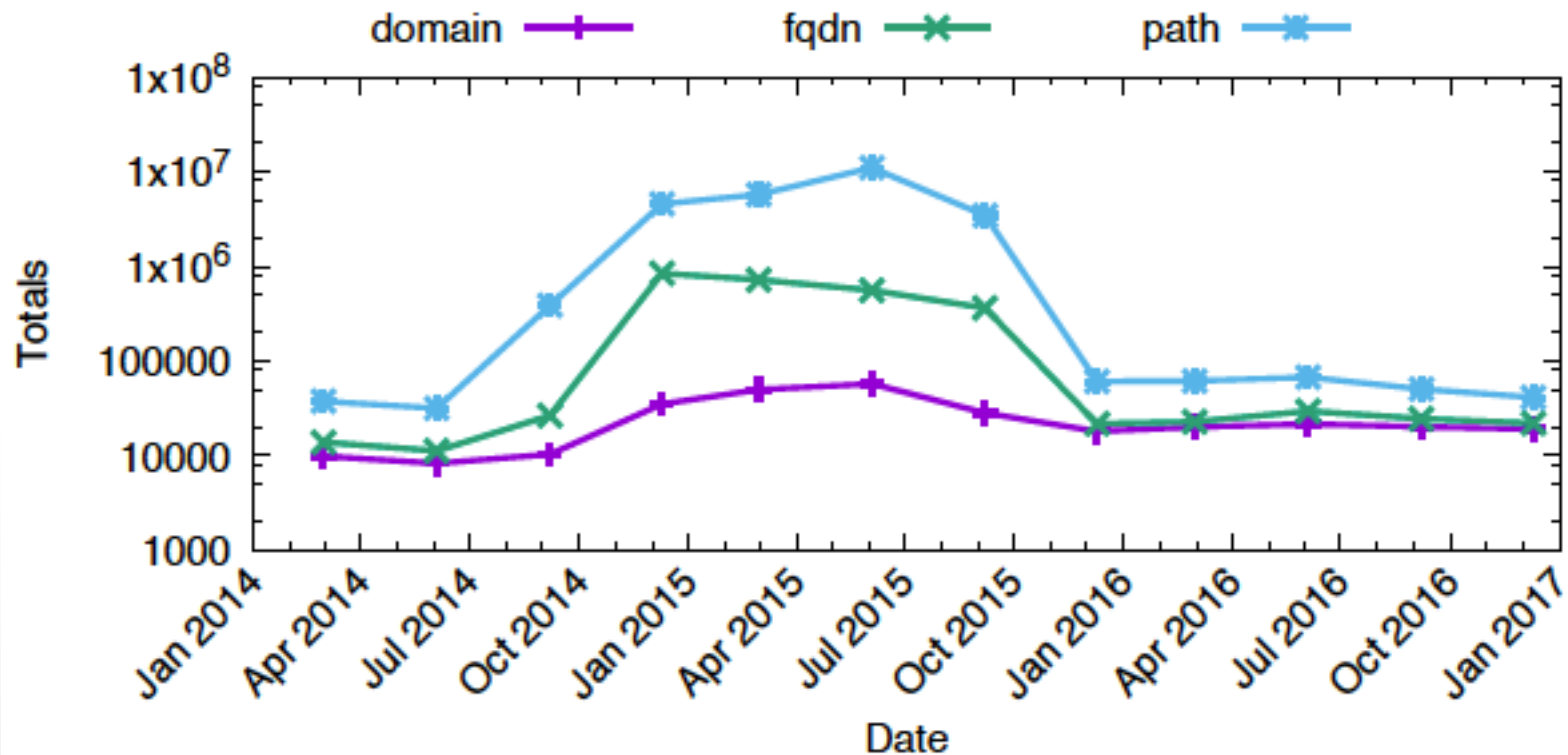
Security Metrics for gTLDs

Phishing domains, FQDNs, and URLs (APWG) per legacy gTLDs



Security Metrics for gTLDs

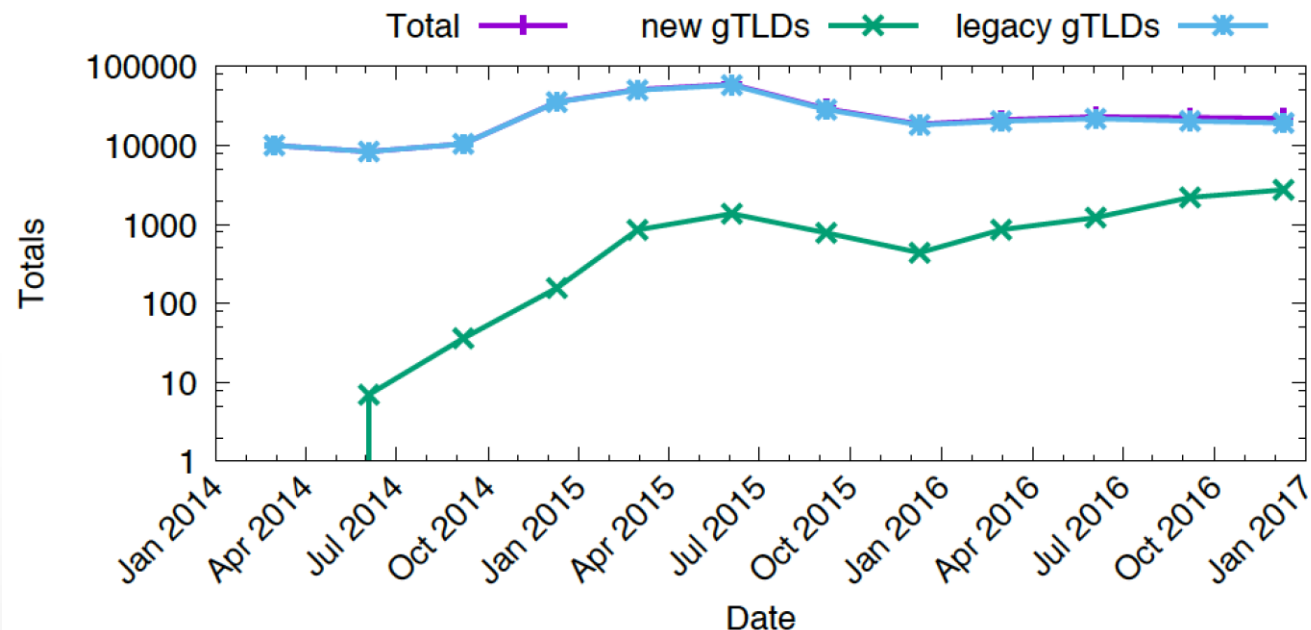
Phishing domains, FQDNs, and URLs (APWG) per legacy gTLDs



Three measures reflect attackers' profit-maximizing behavior. They abuse free legitimate services and affect the reputations of such associated services.

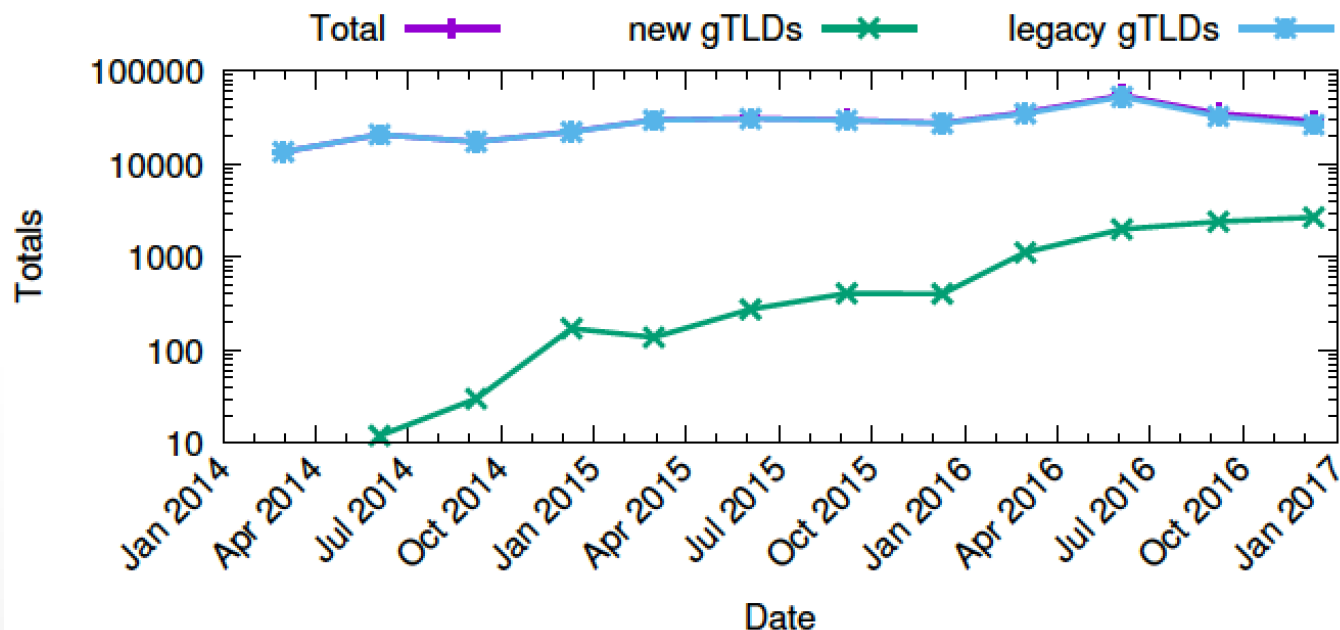
Security Metrics for gTLDs

Phishing domains (APWG) per new and legacy gTLDs



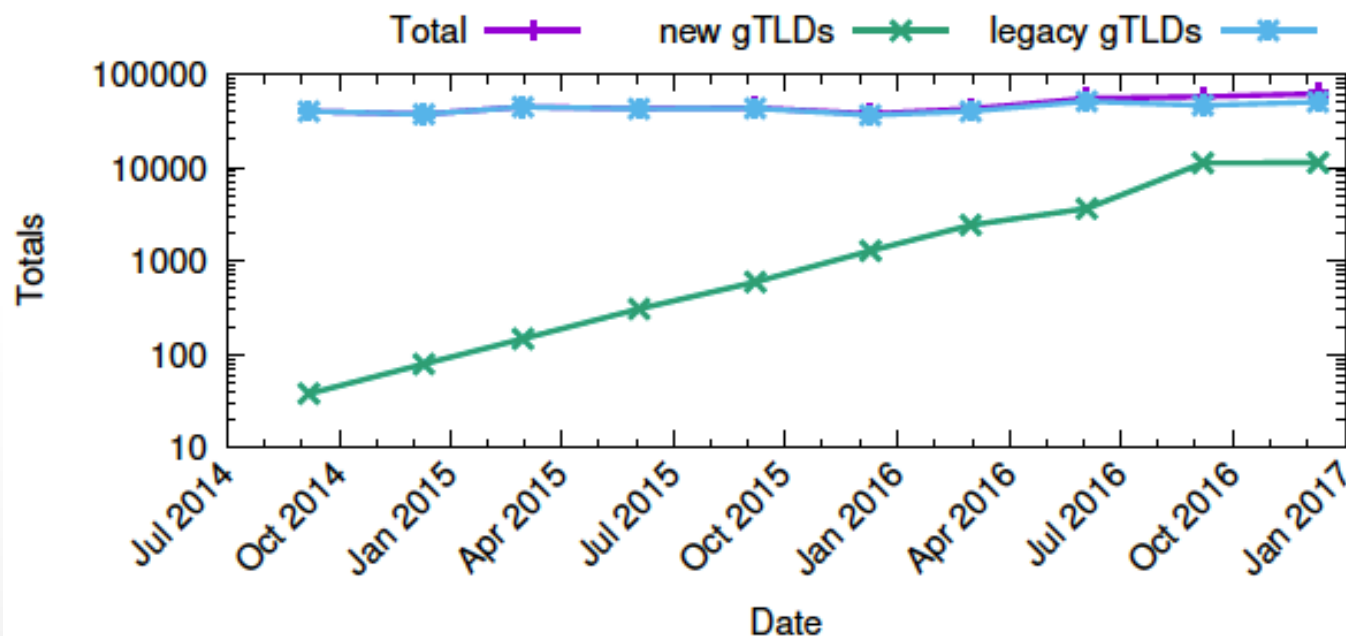
Security Metrics for gTLDs

Phishing domains (CleanMX ph) per new and legacy gTLDs



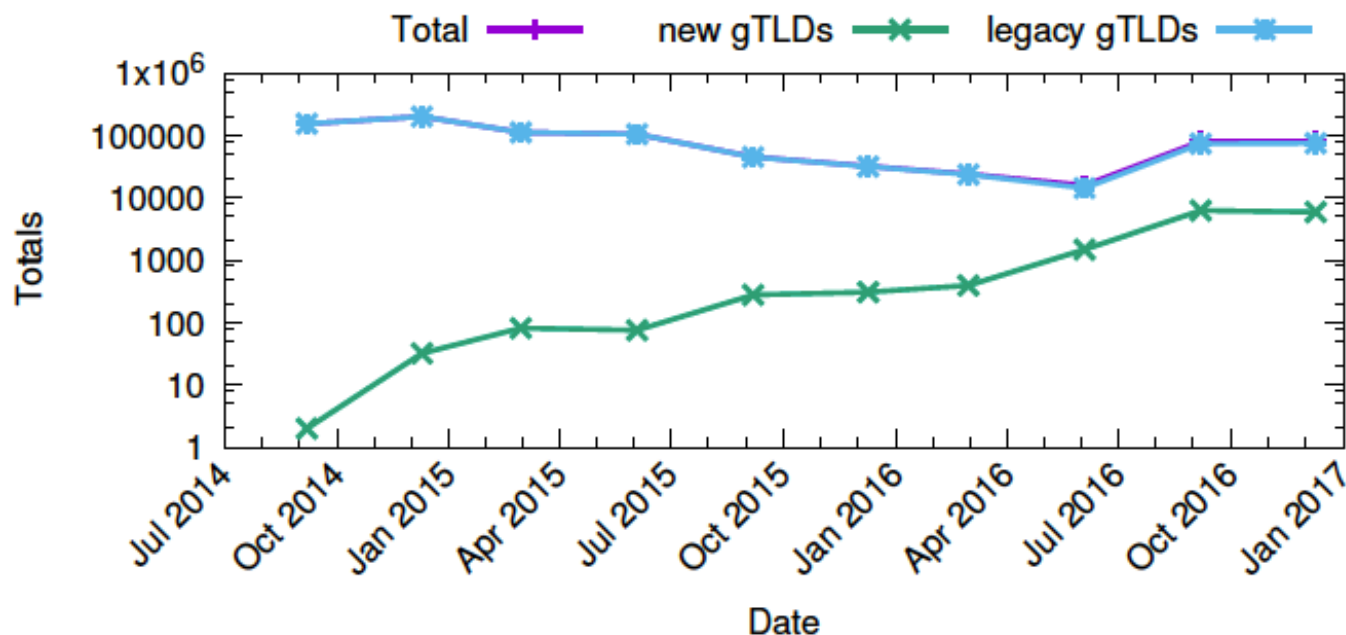
Security Metrics for gTLDs

Phishing domains (SURBL ph) per new and legacy gTLDs



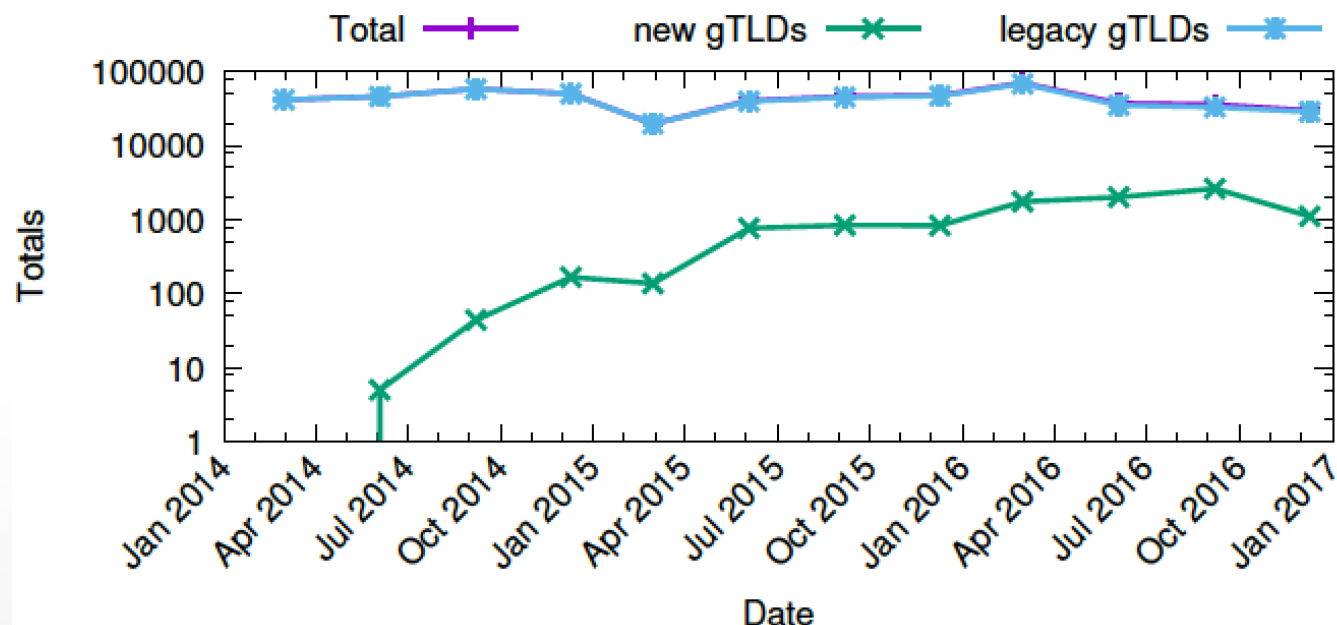
Security Metrics for gTLDs

Malware domains (SURBL mw) per new and legacy gTLDs



Security Metrics for gTLDs

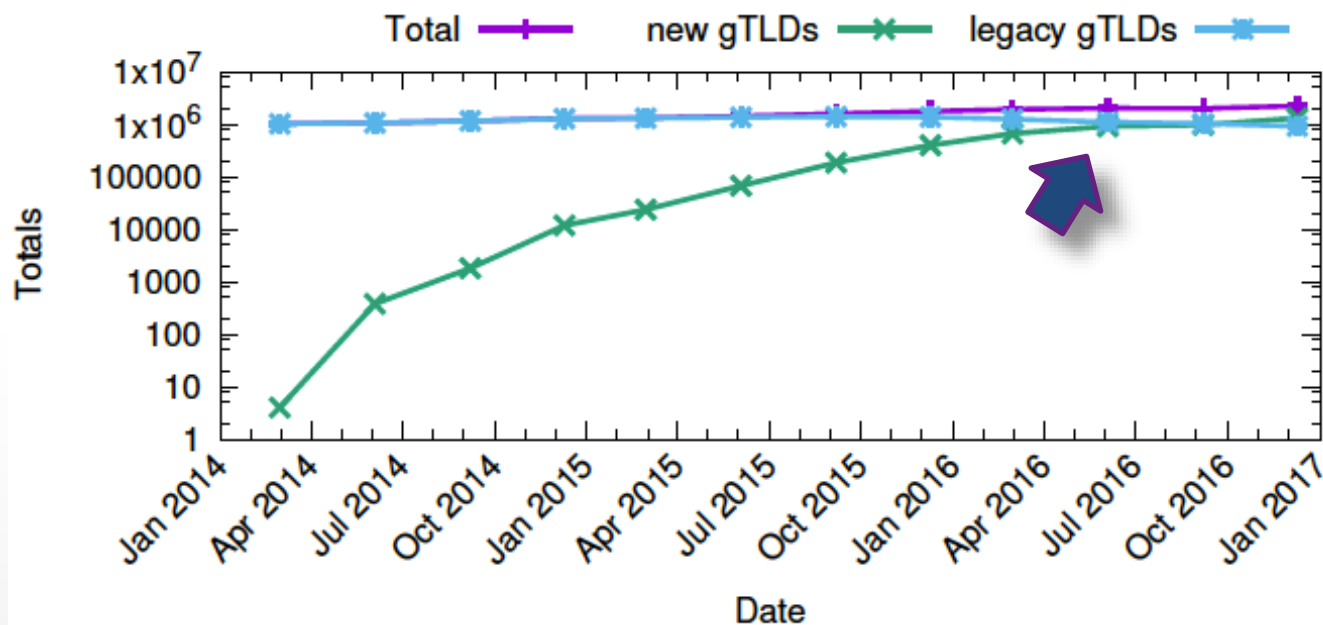
Malware domains (CleanMX mw) per new and legacy gTLDs



While the number of abused domains remains approximately constant in legacy gTLDs, we observe a clear upward trend in the absolute number of **phishing** and **malware** domains in new gTLDs.

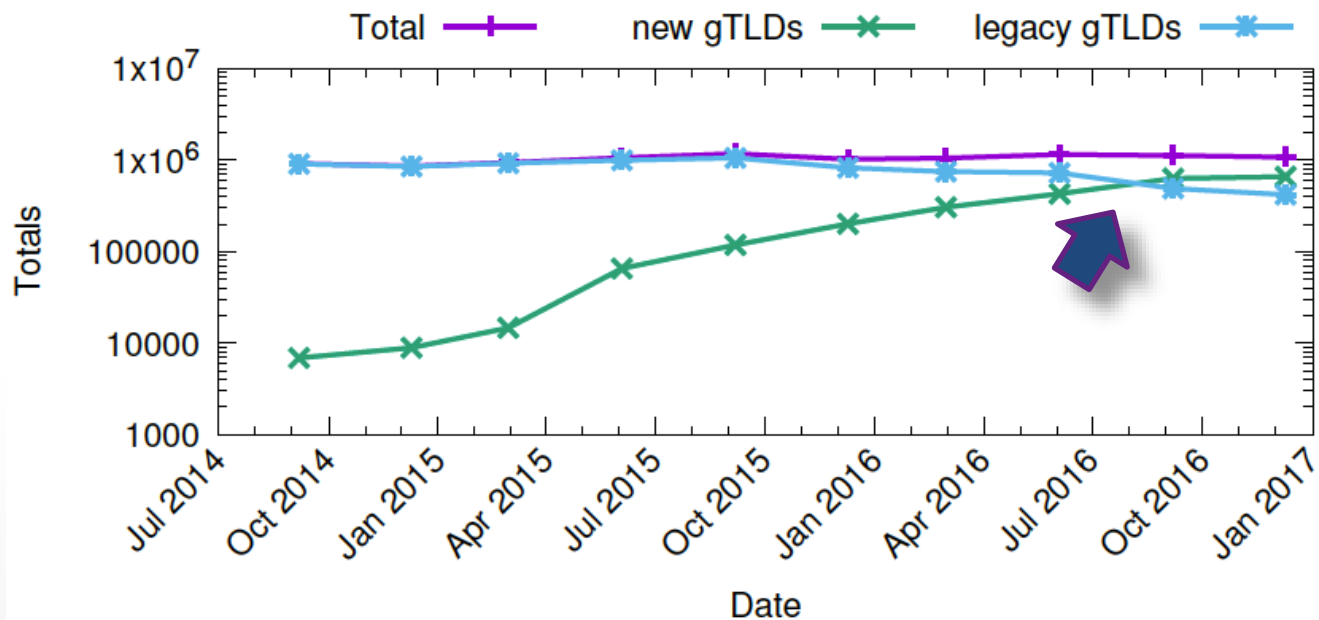
Security Metrics for gTLDs

Spam domains (Spamhaus) per new and legacy gTLDs



Security Metrics for gTLDs

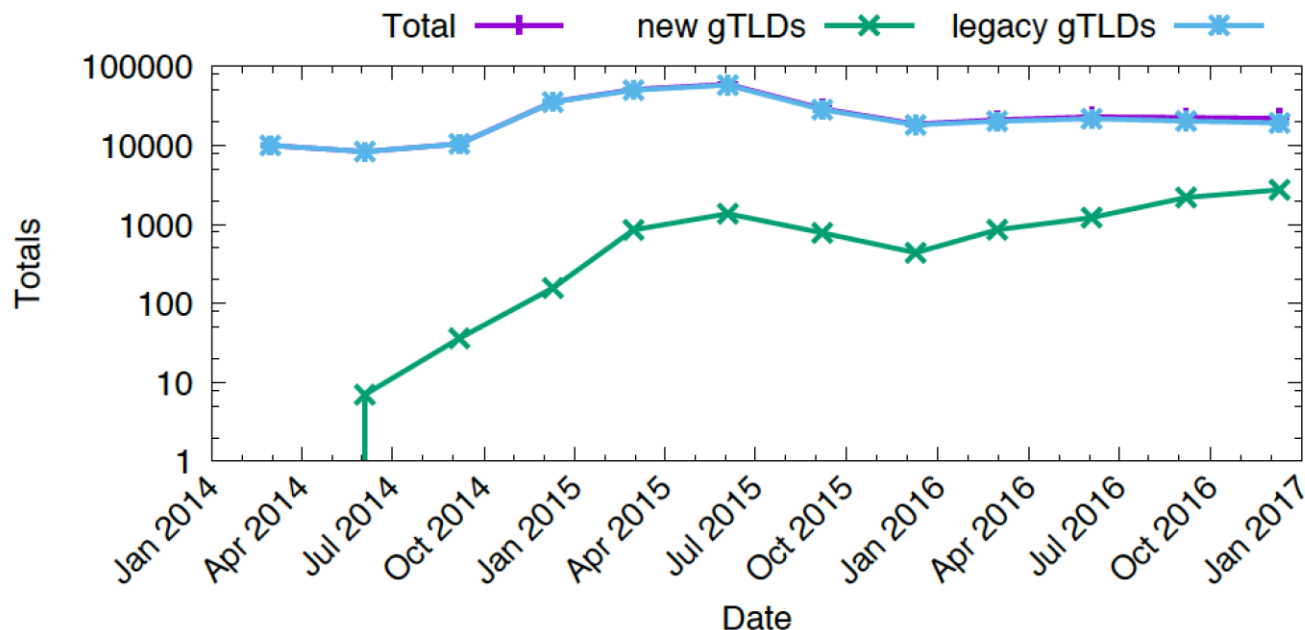
Spam domains (SURBL ws) per new and legacy gTLDs



The **absolute** number of **spam** domains in new gTLDs higher than in legacy gTLDs at the end of 2016

Security Metrics for gTLDs

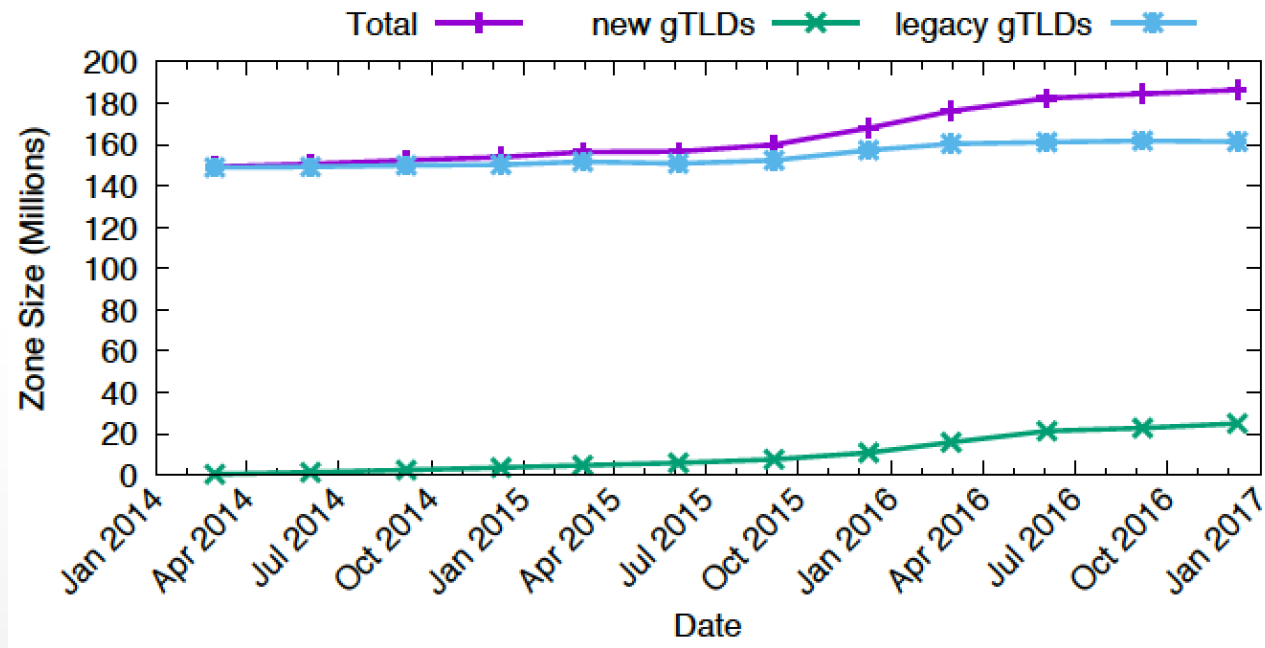
Phishing domains (APWG) per new and legacy gTLDs



Size matters!

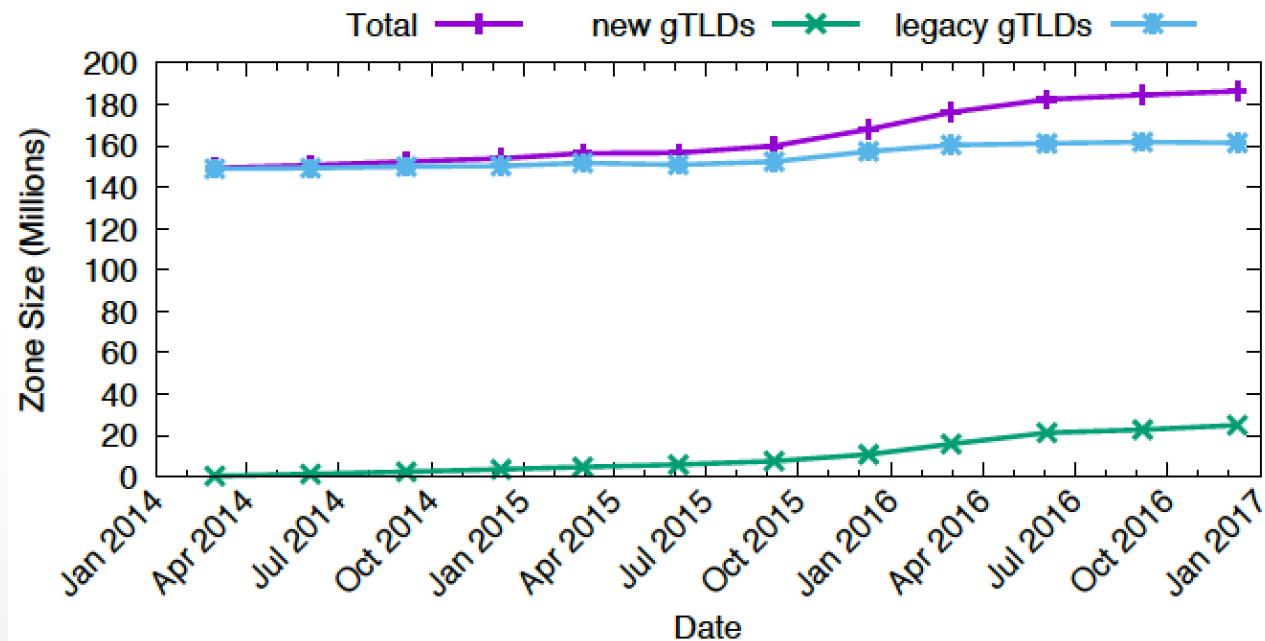
Size

- Size estimate: Number of domains in each gTLD zone file



Size

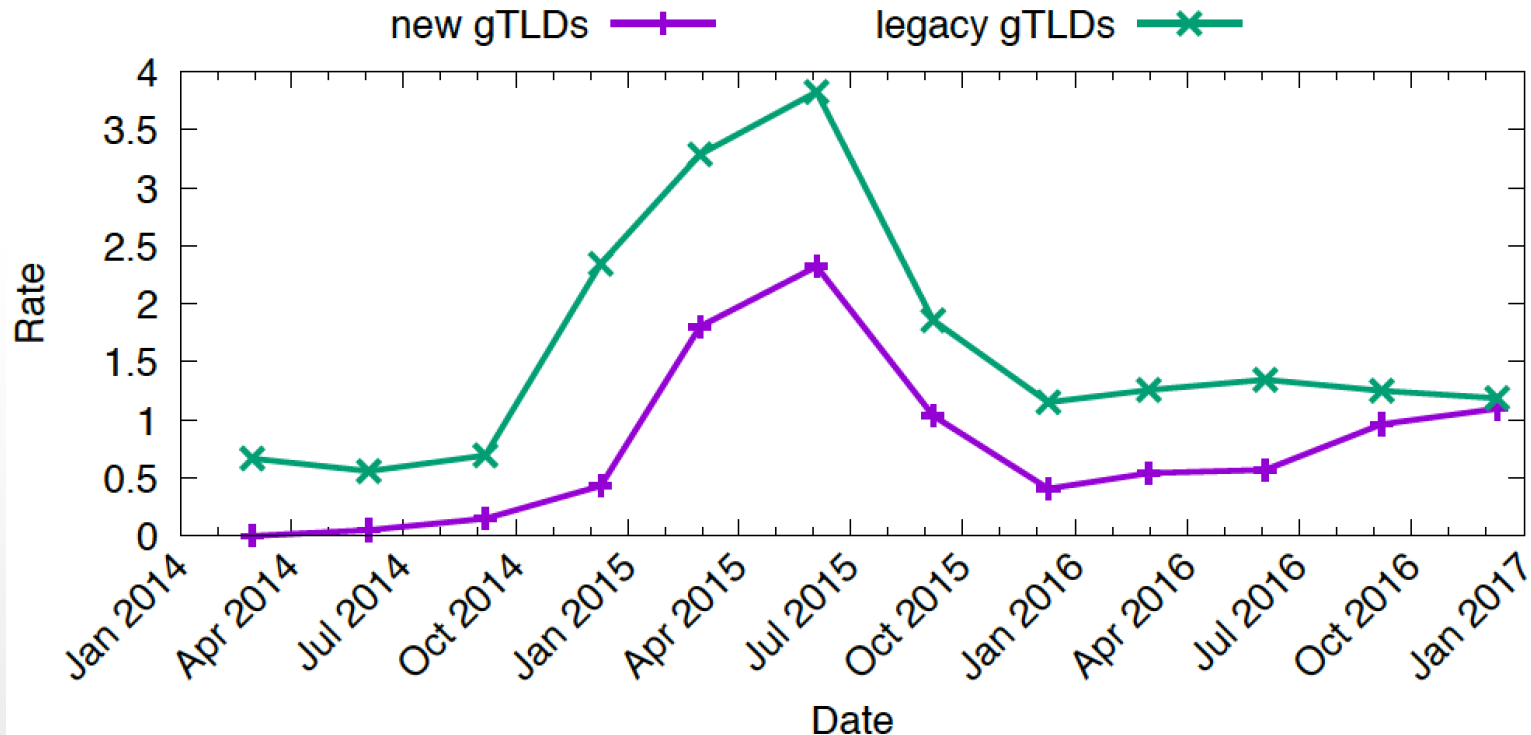
- Size estimate: Number of domains in each gTLD zone file



- Rates: $(\text{\#blacklisted domains} / \text{\#all domains}) * 10,000$

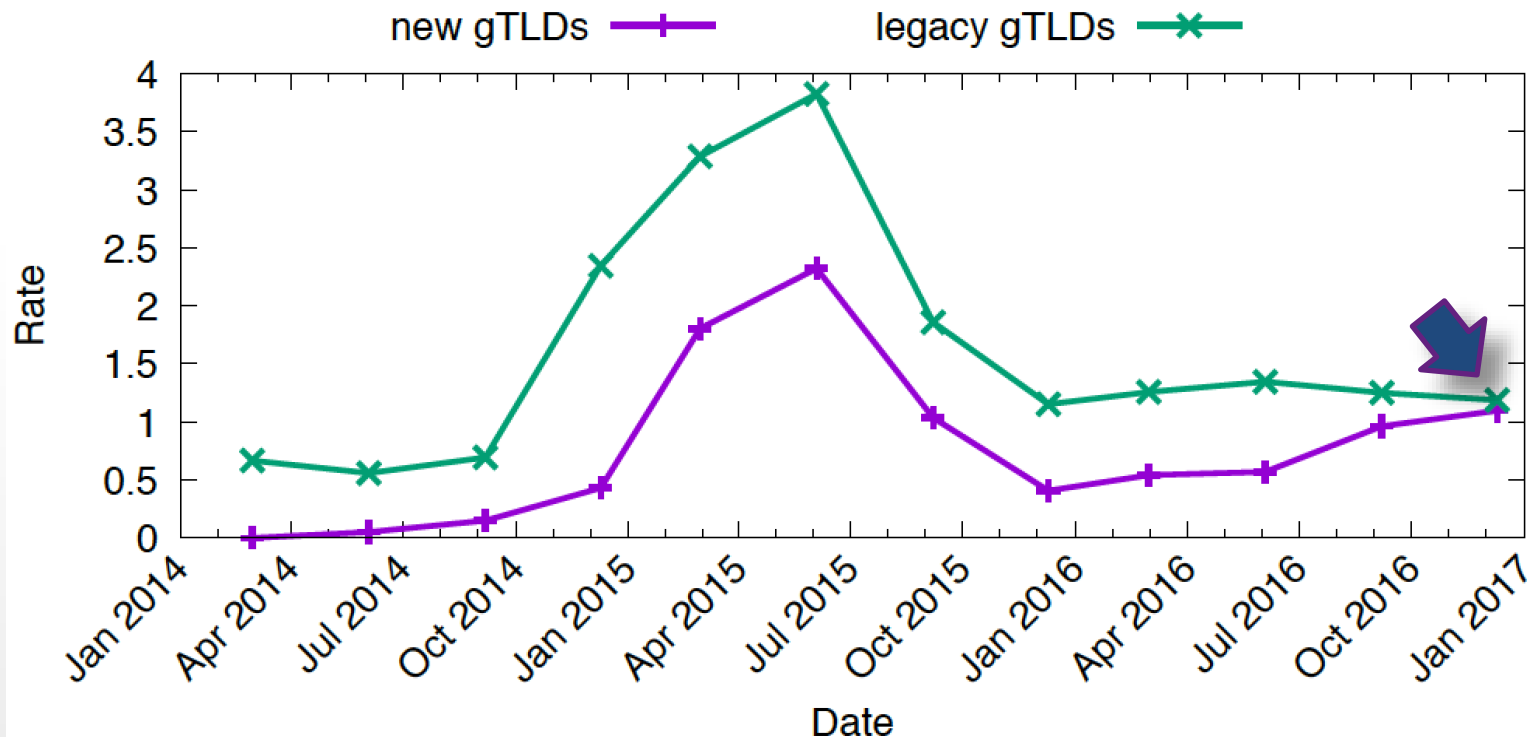
Abuse Rates

- Time series of abuse rates of **phishing** domains in legacy gTLDs and new gTLDs based on the APWG feed



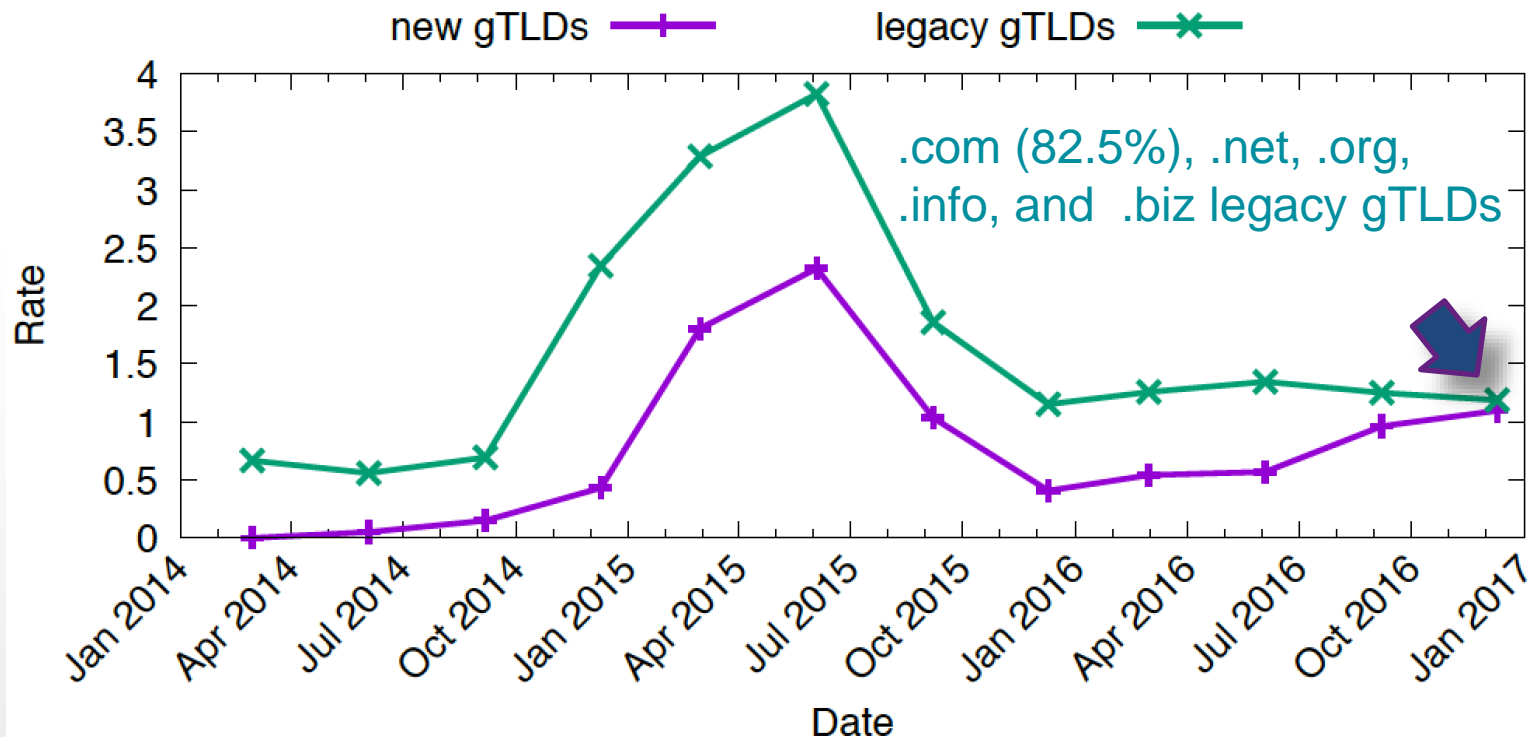
Abuse Rates

- Time series of abuse rates of **phishing** domains in legacy gTLDs and new gTLDs based on the APWG feed



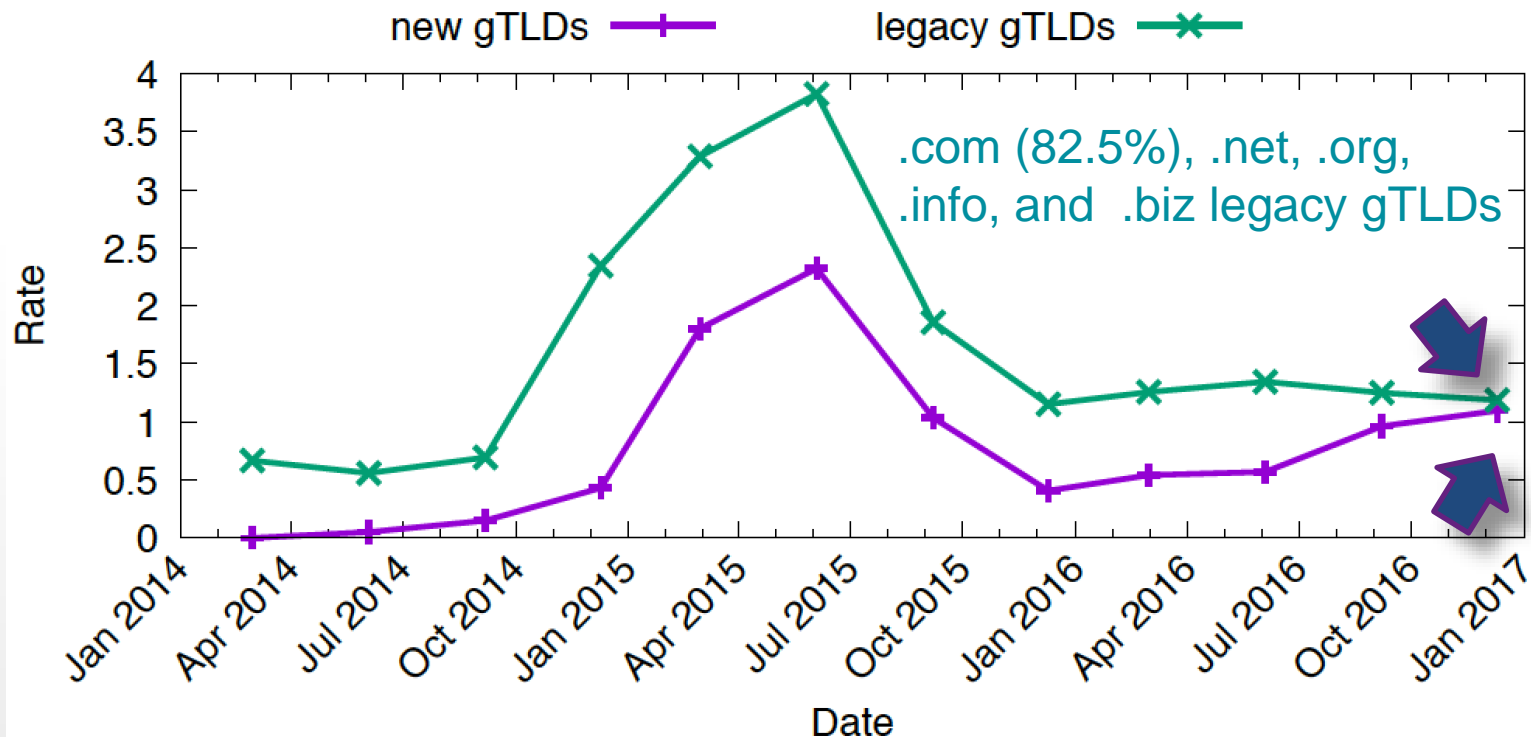
Abuse Rates

- Time series of abuse rates of **phishing** domains in legacy gTLDs and new gTLDs based on the APWG feed



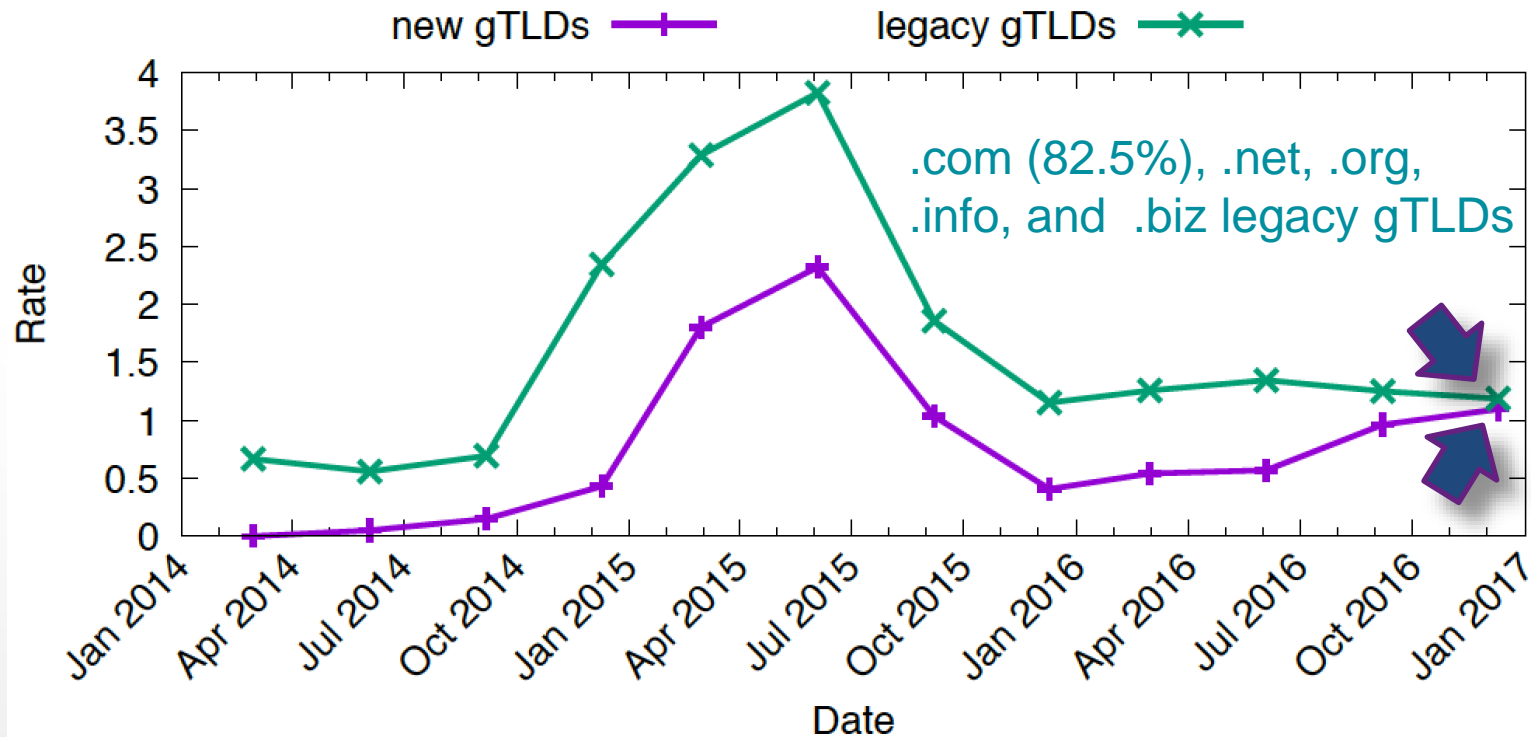
Abuse Rates

- Time series of abuse rates of **phishing** domains in legacy gTLDs and new gTLDs based on the APWG feed



Abuse Rates

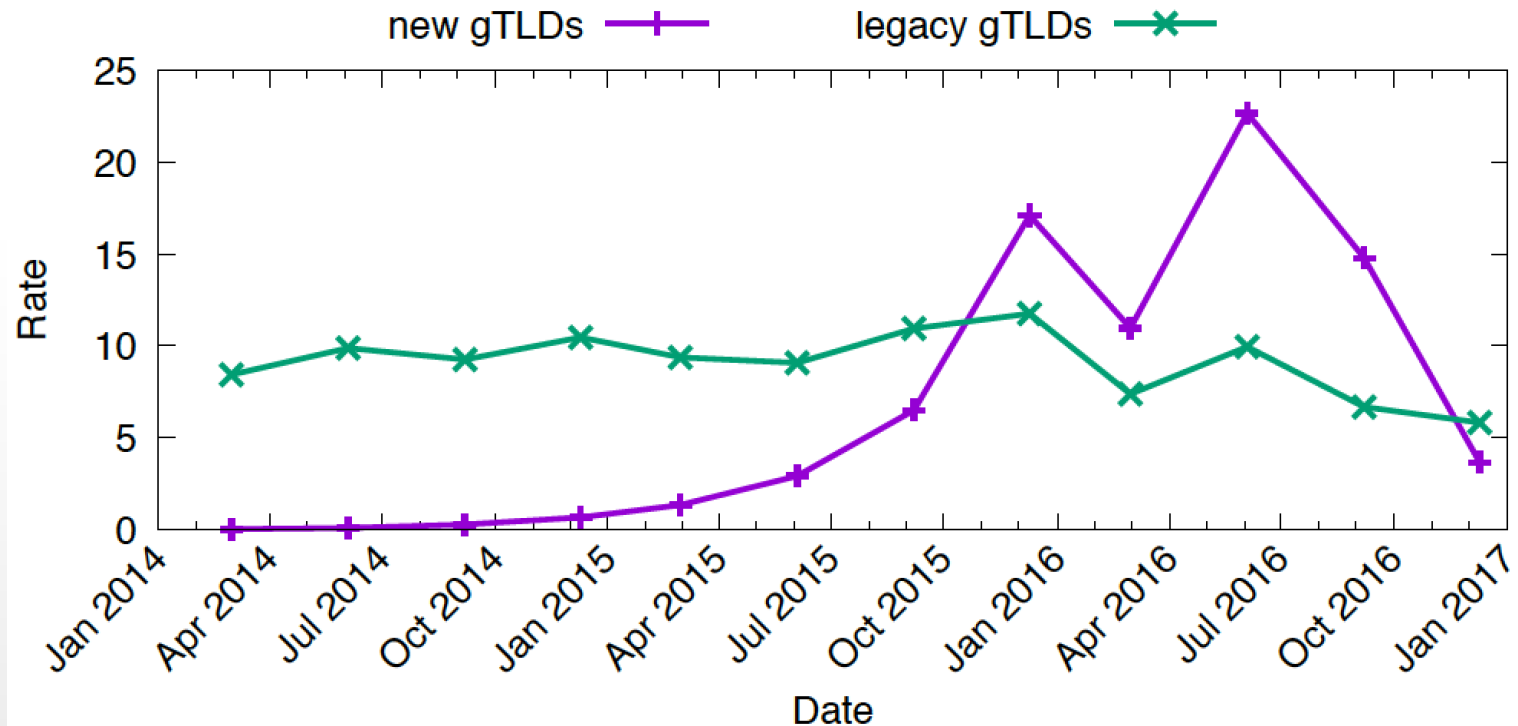
- Time series of abuse rates of **phishing** domains in legacy gTLDs and new gTLDs based on the APWG feed



Top 5 most abused new gTLDs collectively owned 58.7% of all blacklisted domains in all new gTLDs

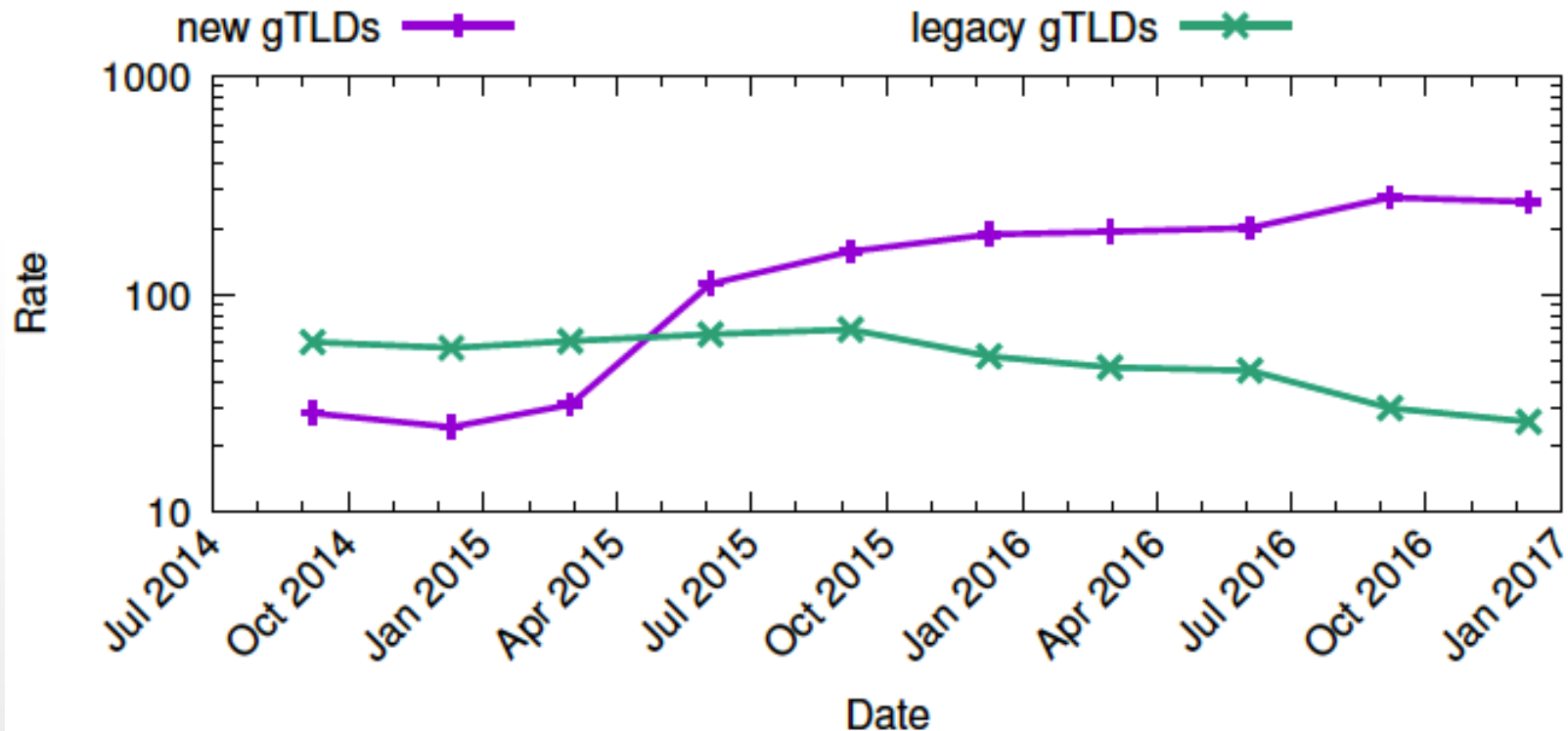
Abuse Rates

- Time series of abuse rates of **malware** domains in legacy gTLDs and new gTLDs based on the StopBadware feed



Abuse Rates

- Time series of abuse rates of **spam** domains in legacy gTLDs and new gTLDs based on the Spamhaus feed



Abuse Rates

- Top 10 new gTLDs with the highest relative concentrations of blacklisted domains for SURBL and Spamhaus datasets (4Q 2016)

Spamhaus

TLD	# Domains	Rate
SCIENCE	117,782	5,154
STREAM	18,543	4,756
STUDY	1,118	3,343
DOWNLOAD	16,399	2,016
CLICK	20,713	1,814
TOP	736,339	1,705
GDN	45,547	1,602
TRADE	23,581	1,521
REVIEW	9415	1,318
ACCOUNTANT	6,722	1,279

SURBL ws

TLD	# Domains	Rate
RACING	51,443	3,812
DOWNLOAD	21,515	2,645
ACCOUNTANT	10,543	2,007
REVIEW	12,615	1,766
GDN	49,427	1,739
FAITH	5,540	1,301
TRADE	19,330	1,247
CLICK	13,270	1,162
STREAM	4,406	1,130
DATE	1,3851	999

- Rates: $(\text{\#blacklisted domains} / \text{\#all domains}) * 10,000$

Abuse Rates

- Does the problem affect all new gTLDs?

Abuse Rates

- Does the problem affect all new gTLDs?
- No

Abuse Rates

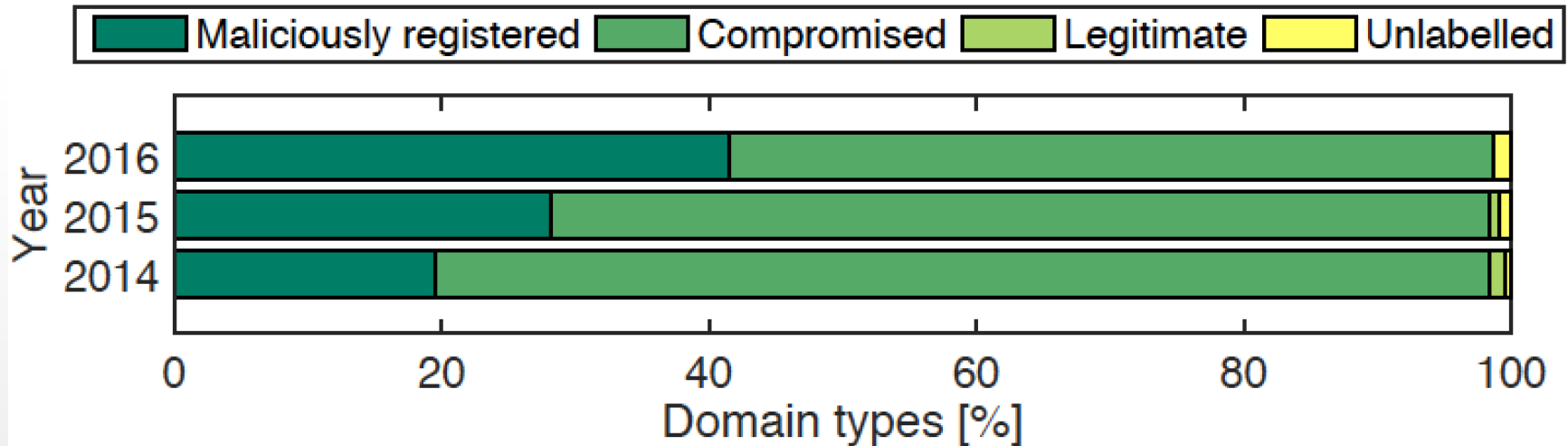
- Does the problem affect all new gTLDs?
- No
- Spamhaus and SURBL blacklists reveal that 32% and 36% of all new gTLDs available for registration did not experience a single incident in 4Q 2016.
- Spamhaus blacklisted at least 10% of all registered domains in as many as 15 new gTLDs in 4Q 2016.

Compromised and Maliciously Registered Domains

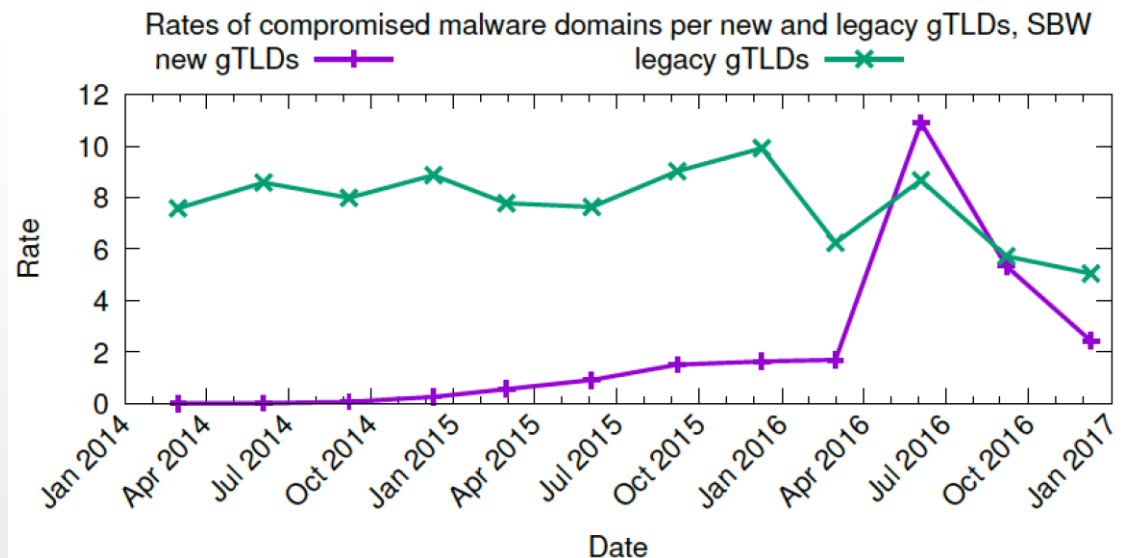
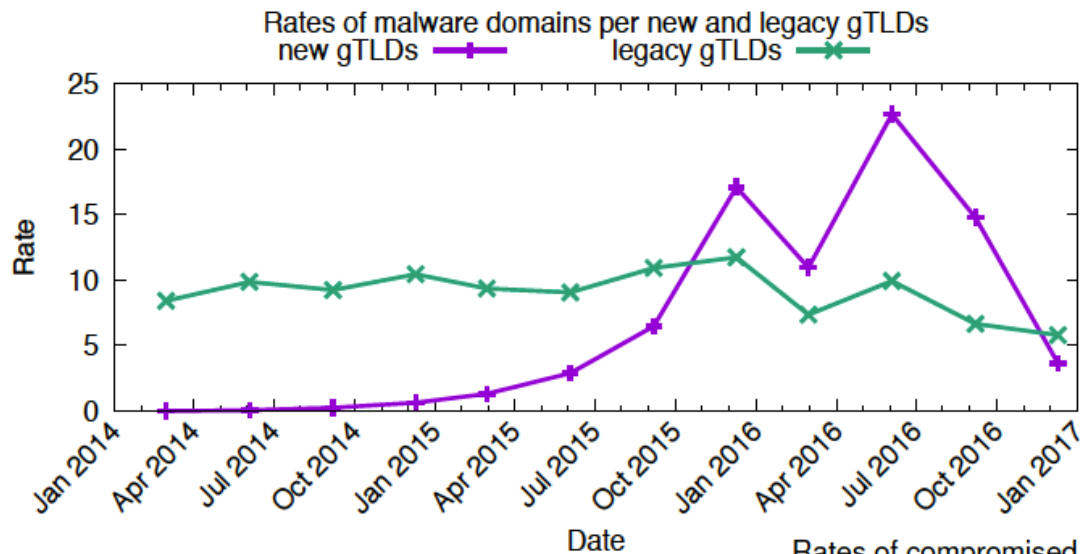
- Distinguishing between compromised and maliciously registered domains is critical because they require different mitigation actions by different intermediaries
- Three heuristics:
 - if a given domain name contains a string of a brand name, or
 - if its misspelled version, or
 - if it's involved in malicious activity within three months after creation.

Compromised and Maliciously Registered Domains

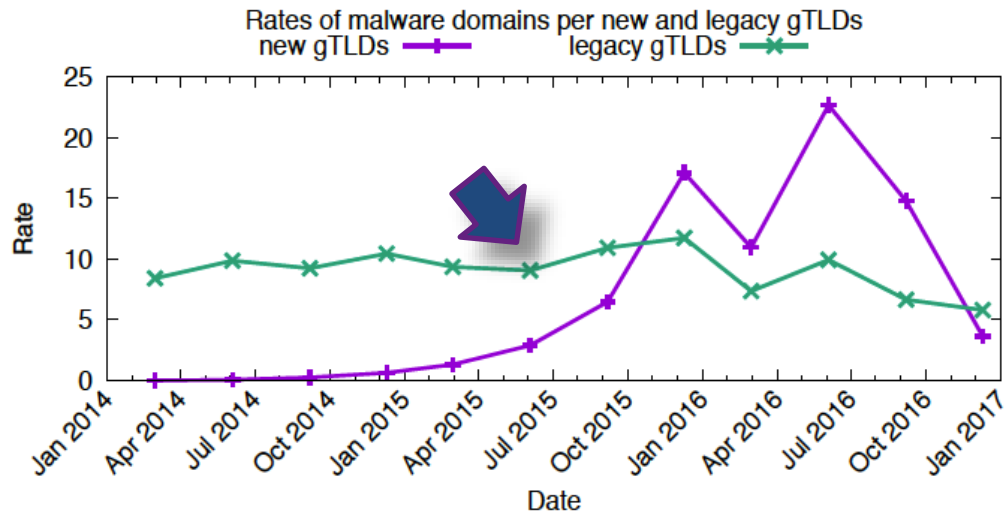
- Distinguishing between compromised and maliciously registered domains is critical because they require different mitigation actions by different intermediaries



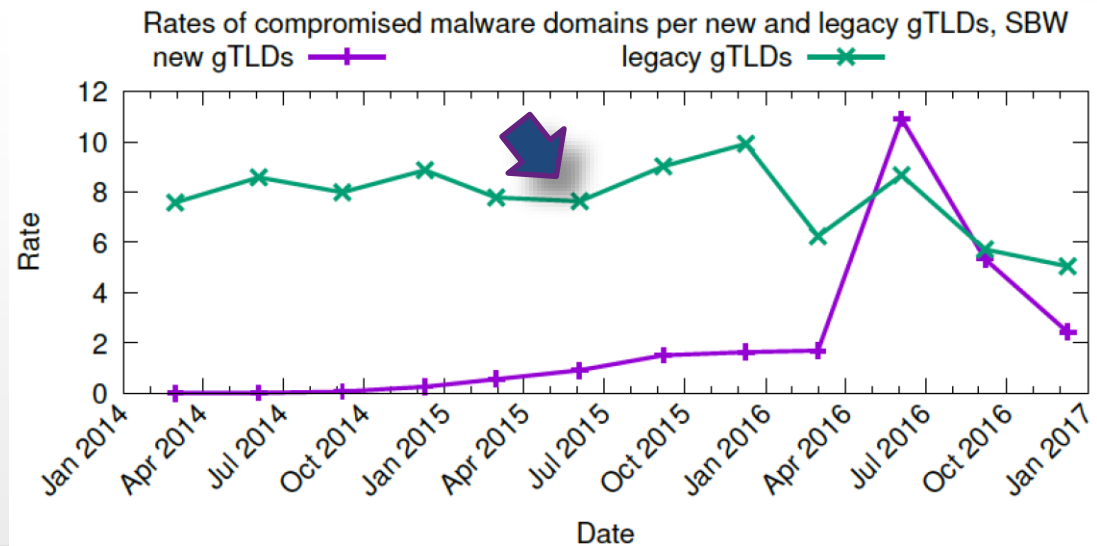
Compromised Domains



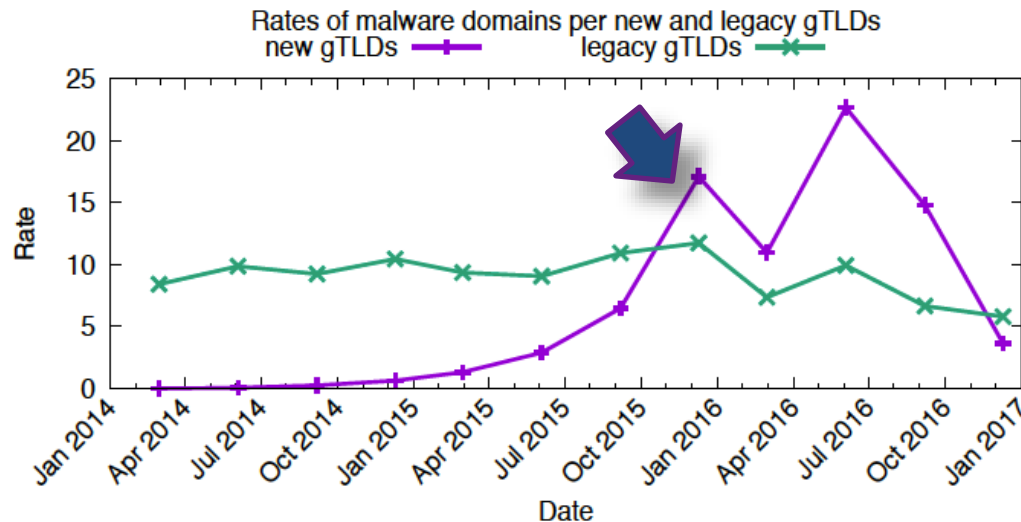
Compromised Domains



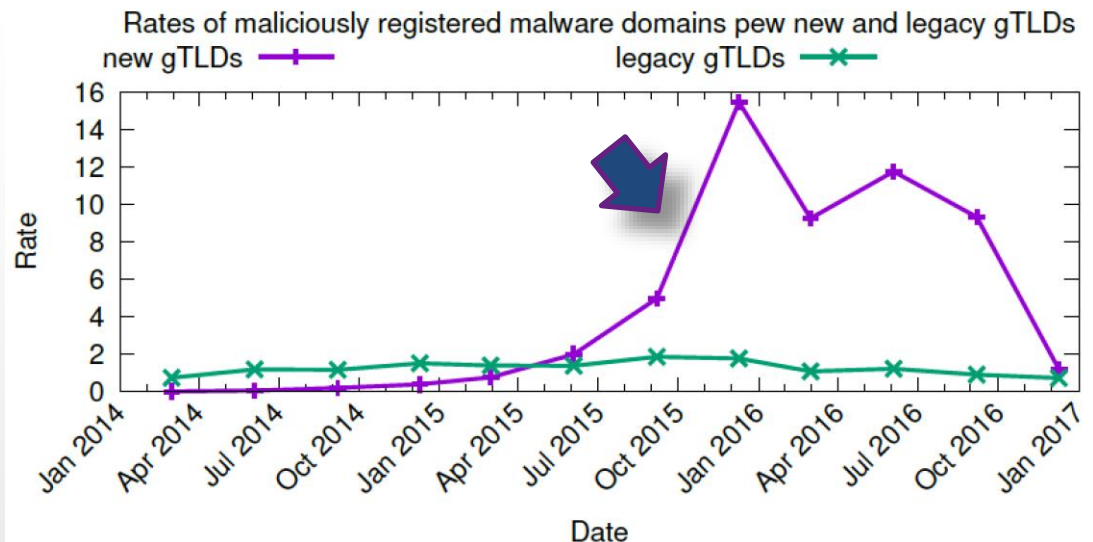
- Rates of abused domains in legacy gTLDs (StopBadware URL blacklists) are driven by compromised domains



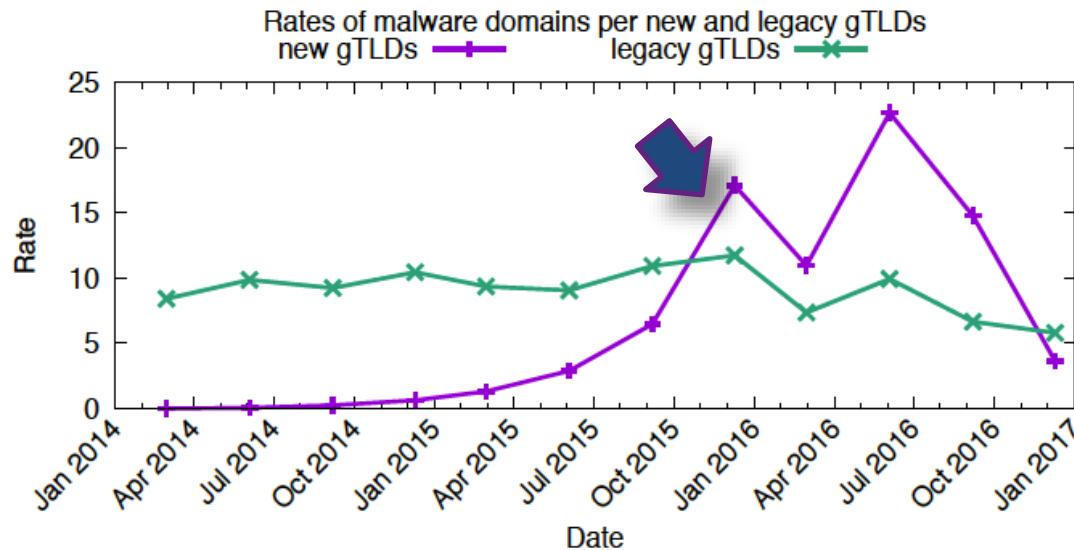
Maliciously Registered Domains



- Rates of abused domains in new gTLDs (StopBadware URL blacklist) are driven by maliciously registered domains

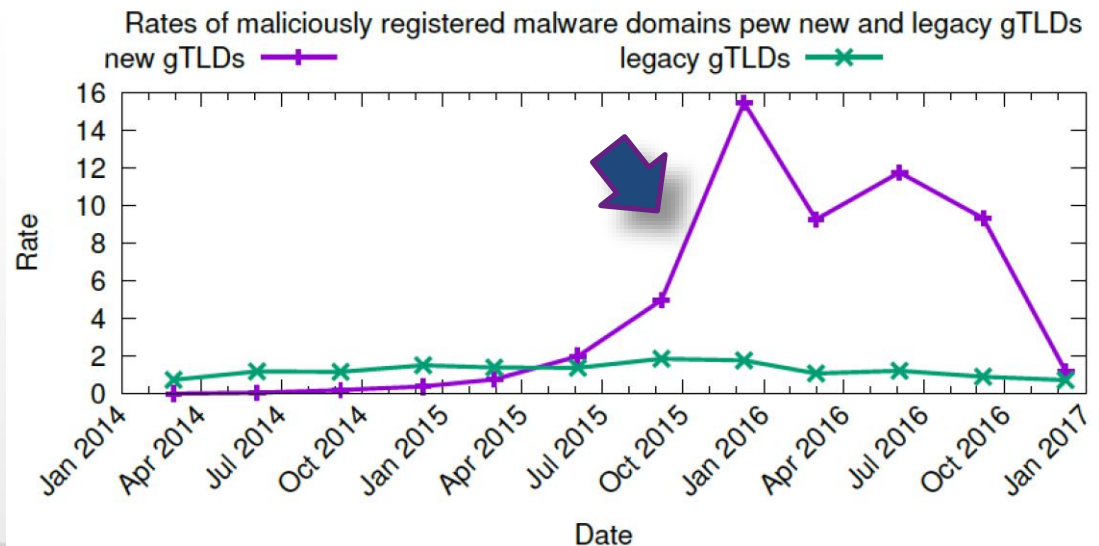


Maliciously Registered Domains



- Rates of abused domains in new gTLDs (StopBadware URL blacklist) are driven by maliciously registered domains

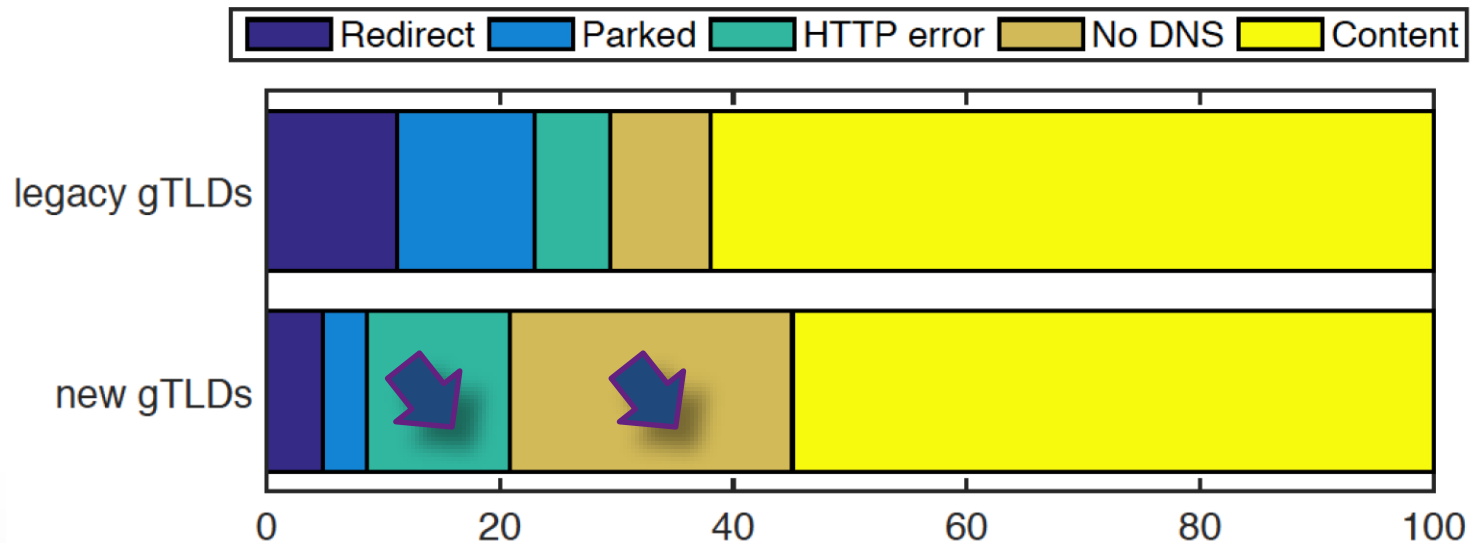
...and can be driven by single campaigns (domains registered in bulk, common patterns in domain names)



Inferential Analysis of Abuse in New gTLDs

Driver	Rationale
New gTLD size	Larger TLDs have a larger “attack surface” (compromised domains)
DNSSEC	Hypothesis: proxy for security efforts, however, miscreants could be interested in deploying DNSSEC and signing their maliciously registered domains
Parked	Domains serving content are exposed to certain types of vulnerabilities and can be hacked. However, parked domains may be used to scam users or to distribute malware
No DNS, HTTP error	Domains serving content are exposed to certain types of vulnerabilities and can be hacked
Type	Proxy for strict registration policies (registration “levels” to new gTLDs, from the least to most restricted groups: 1 generic, 2 geographic, 3 community, and 4 brand)
Registry operator (parent companies of registry operators)	Proxy for registration practices (e.g. pricing, registration in bulk, payment methods)

Inferential Analysis of Abuse in New gTLDs



“No DNS” domains account for 24.2% of all domains, whereas domains for which the websites serve an HTTP error account for another 12.2%.

Inferential Analysis of Abuse in New gTLDs

Driver	Correlation with abuse counts
New gTLD size	Very weak positive
DNSSEC	Very weak positive
Parked	Very weak positive
No DNS	Very weak negative
HTTP Error	Very weak negative
Type	Negative (statistically significant results for phishing)
Registry operator	No statistically significant results

Privacy or Proxy Services

- Why use Privacy and Proxy services
 - Protecting your personal data
 - Blocking Spam
 - Stopping unwanted solicitations
- Analyzing use of Privacy and Proxy
 - Extract list of registrants
 - keyword search using “privacy”, “proxy”, “protect” etc.
 - Manual inspection
- How many?
 - We found 570

Privacy or Proxy Services

⚠️ Unprotected

yourdomain.com

Your Real Name
Your Business Name
123 Real Home Address, Apt 213
Your Hometown, VA 22201
Phone: (703) 555-5555
Email: yourname@yourdomain.com

🔒 Protected

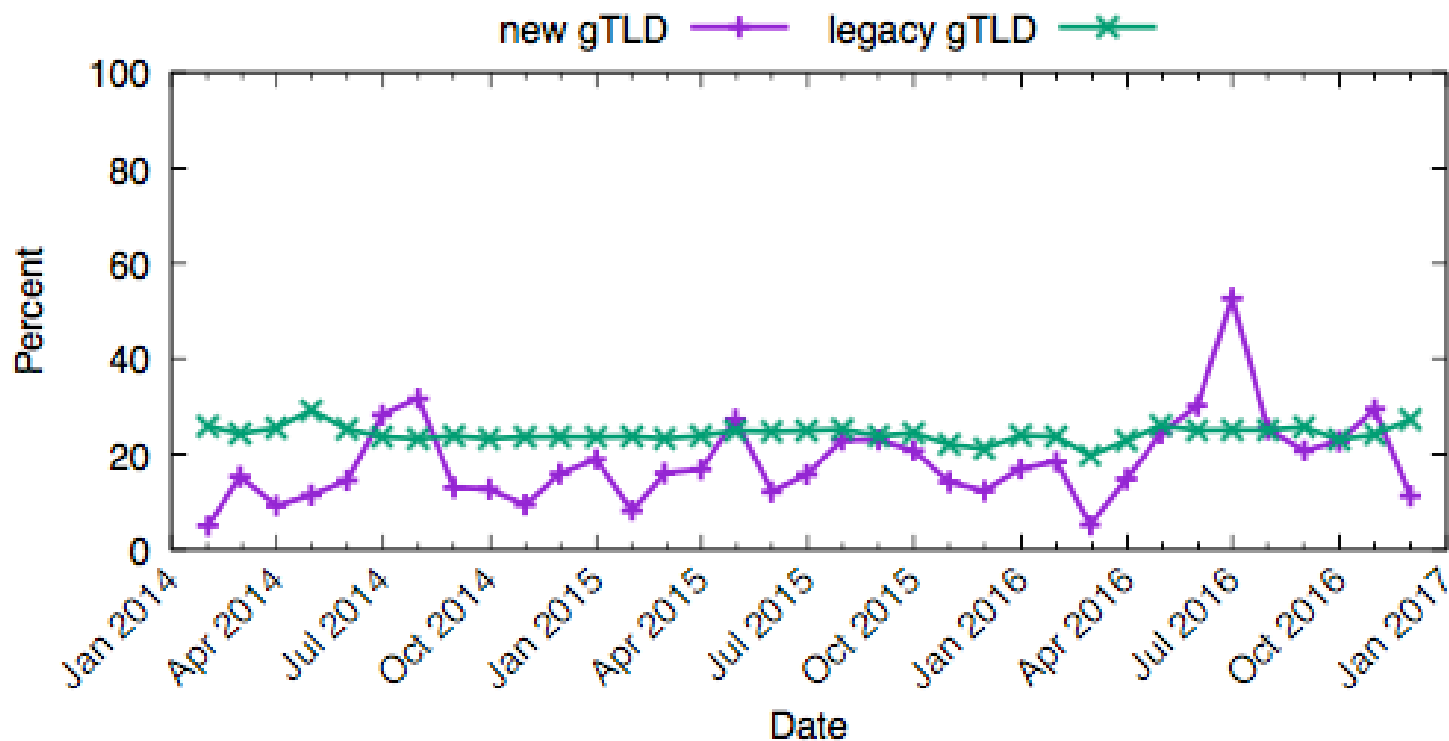
domain.example

Whois Agent
Whois Privacy Protection Service, Inc.
PO Box 639
Kirkland, WA 98083
+1 425.274.0657
domain@protecteddomainservices.com

Image source: <https://www.name.com/whois-privacy>

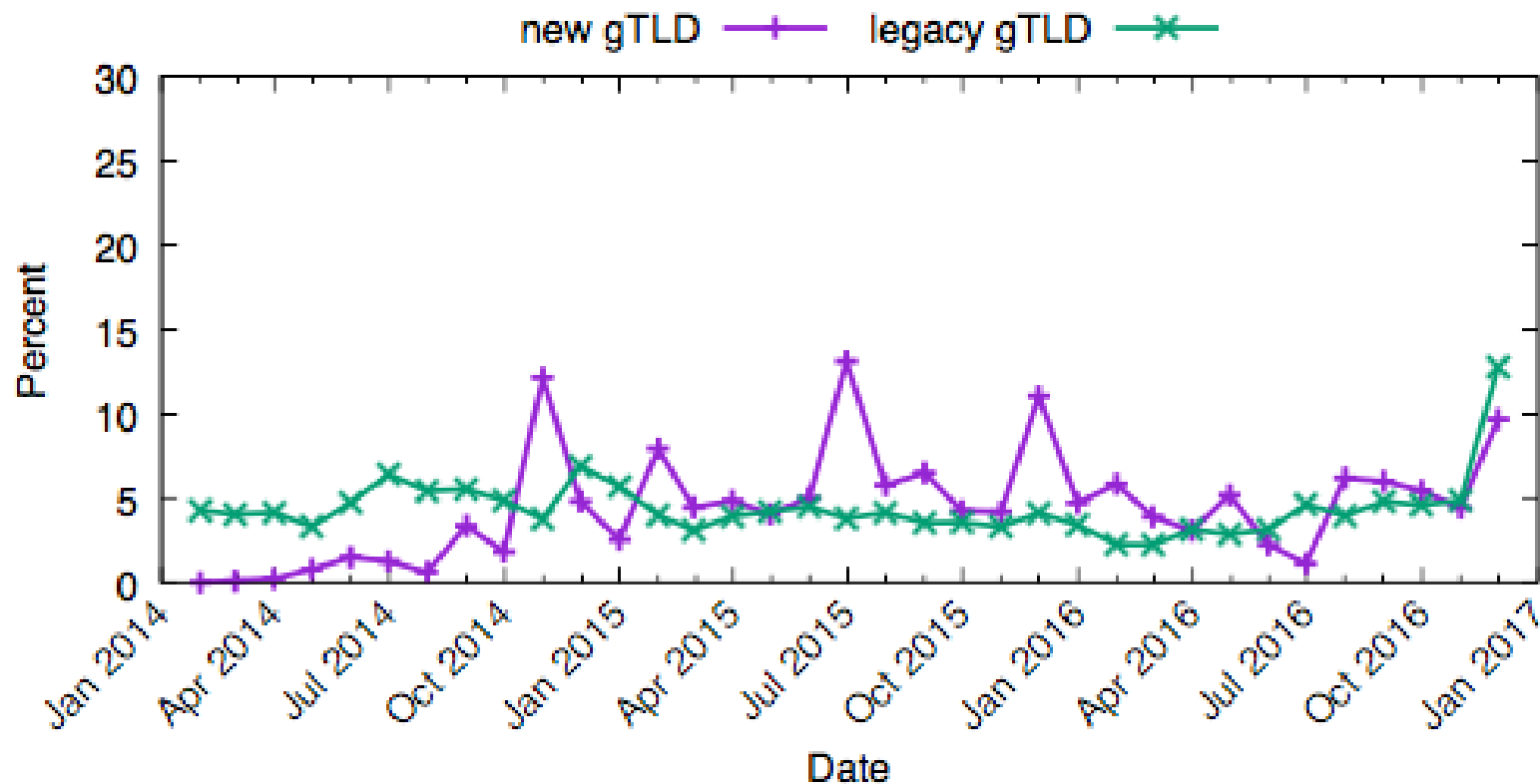
Privacy or Proxy Services

Usage for Newly Created Domains



Privacy or Proxy Services

Usage for Abusive Newly Registered Domains



Privacy or Proxy Services

- The usage of Privacy or Proxy Services by itself is not a reliable indicator of abuse.
- Usage of Privacy or Proxy Services remains higher for legacy gTLDs.

Geographical Location

- Using domain registrar location from WHOIS
 - Registrant details not reliable
- Method
 - Extract unique "registrar name" from WHOIS data.
 - Combine the registrar name with the country information for ICANN-Accredited Registrars.
 - Match remaining name variants
 - Manually lookup the country information for missing registrars
- Result
 - 5,985 registrars
 - 99.99% of domains

Geographical Location

Registrar Distribution

Country	#Registrars	share
United States	2,682	53.88
China	281	5.64
Germany	201	4.04
Canada	177	3.56
United Kingdom	160	3.21
India	144	2.89
France	116	2.33
Australia	111	2.23
Spain	105	2.11
Japan	95	1.91

Geographical Location

Domain Distribution

New	#Domains	Share	Legacy	#Domains	Share
China	8,076,776	27.92	US	152,527,872	56.72
US	6,283,269	21.72	China	24,098,150	8.96
Gibraltar	3,028,035	10.47	Germany	18,044,735	6.71
Cayman Is.	2,069,919	7.16	Canada	16,704,693	6.21
Singapore	1,870,886	6.47	India	11,135,408	4.14
Japan	1,741,228	6.02	Japan	7,935,585	2.95
India	1,323,117	4.57	Australia	6,425,896	2.39
Germany	1,105,708	3.82	France	4,988,581	1.86
Hong Kong	836,069	2.89	UK	4,511,714	1.68
France	450,371	1.56	Turkey	2,418,232	0.9

Geographical Location

SURBL Distribution

New gTLD Country	#Incidents	Percentage	Rate
Gibraltar	751,748	49.44	2482.63
Japan	295,647	19.44	976.37
China	214,332	14.1	707.83
United States	109,989	7.23	363.24
India	54,782	3.6	180.92
United Kingdom	24,955	1.64	82.41
France	20,121	1.32	66.45
United Arab Emirates	11,793	0.78	38.95
Cayman Islands	8,912	0.59	29.43
Canada	6,494	0.43	21.45

Legacy gTLD Country	#Incidents	Percentage	Rate
United States	1,985,574	47.06	130.18
Japan	1,190,409	28.21	78.05
China	319,546	7.57	20.95
India	268,812	6.37	17.62
Germany	73,185	1.73	4.8
Ireland	58,292	1.38	3.82
Canada	40,355	0.96	2.65
Australia	33,080	0.78	2.17
Turkey	32,266	0.76	2.12
Bahamas	28,918	0.69	1.9

Registrar Reputation

- Method
 - Filter out registrars designed for sinkholing domains.
 - Count number of incidents per registrar.
 - Calculate percentage of total abuse linked to registrar.

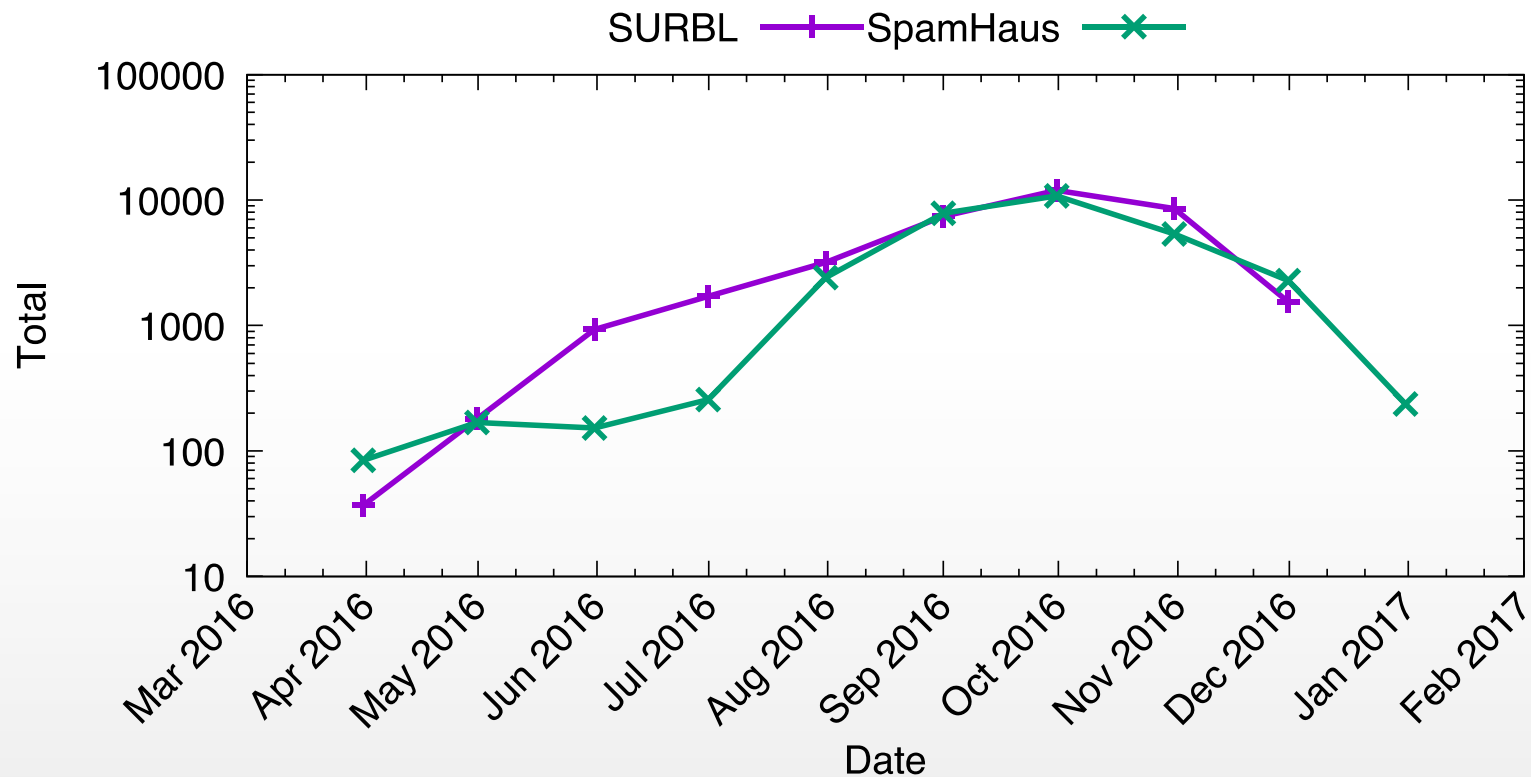
Registrar Reputation

SURBL Distribution

new gTLD registrar	#Domains	#Incidents	Percent
Nanjing Imperiosus Technology	38,025	35,502	93.36
Intracom Middle East FZE	20,640	11,255	54.53
Dot Holding Inc.	153	76	49.67
Alpnames Limited	3,028,011	751,748	24.83
Todaynic.com, Inc.	329,399	69,404	21.07
Web Werks India Pvt. Ltd	785	146	18.6
GMO Internet, Inc. d/b/a Onamae.com	1,734,775	295,641	17.04
TLD Registrar Solutions Ltd.	163,988	24,700	15.06
Xiamen Nawang Technology Co., Ltd	282,925	42,089	14.88
Insta Corporation Pty Ltd.	77,642	6,200	7.99
Legacy gTLD registrar	#Domains	#Incidents	Percent
HOAPDI INC.	141	126	89.36
asia registry r2-asia (700000)	1,379	598	43.36
Nanjing Imperiosus Technology	35,309	10,834	30.68
Paknic (Private) Limited	10,525	3,083	29.29
OwnRegistrar, Inc.	22,188	5,238	23.61
Eranet International Limited	6,109	1,339	21.92
BR domain Inc. dba namegear.co	847	158	18.65
Netlynx Inc.	17,612	3,030	17.2
AFRIREGISTER S.A.	1,551	266	17.15
GMO Internet, Inc. d/b/a Onamae.com	7,306,312	1,177,886	16.12

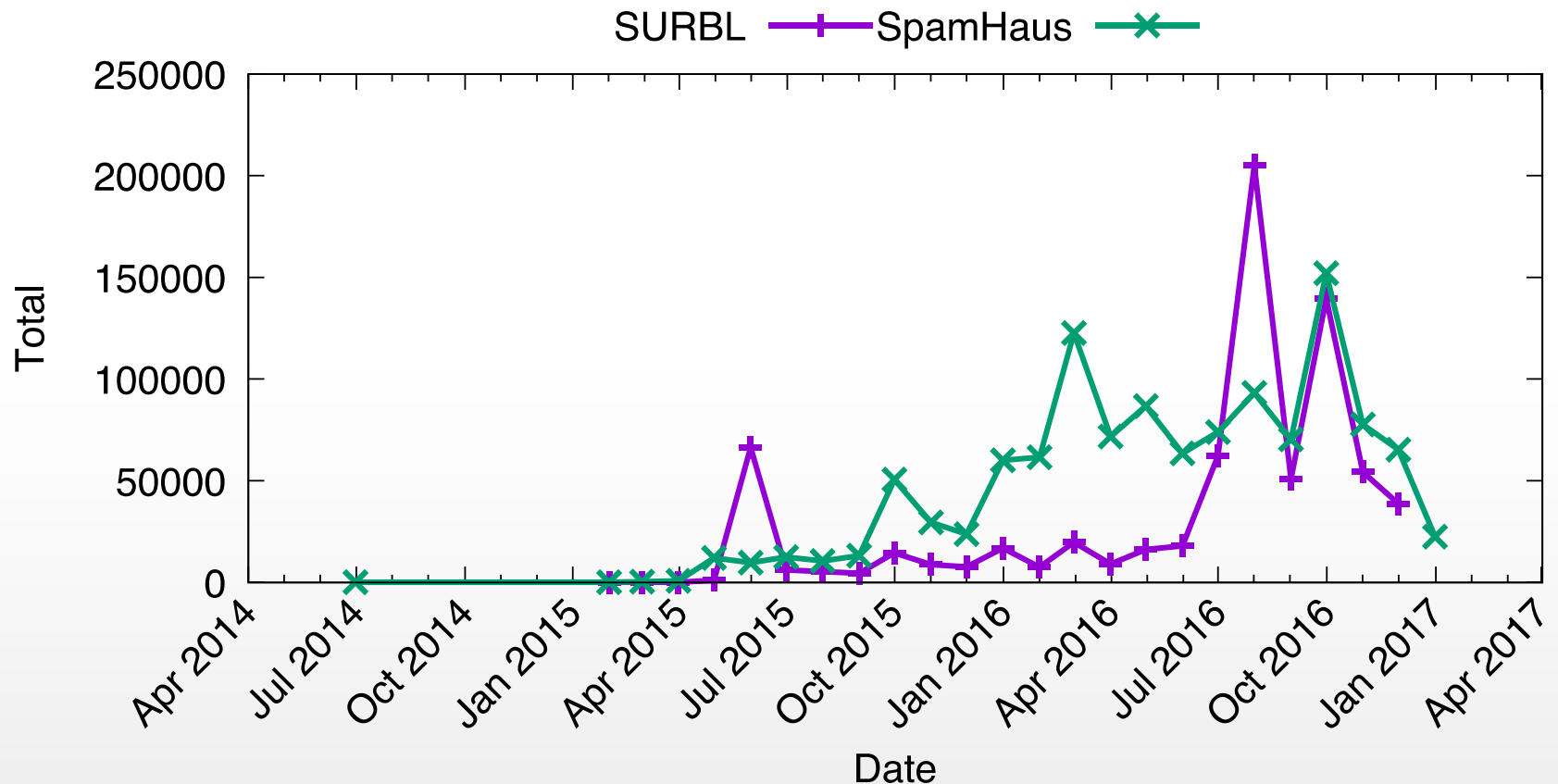
Registrar Reputation

Nanjing Imperiosus Technology Co. Ltd.



Registrar Reputation

Alpnames Ltd.



Questions?



Contact information

Maciej Korczyński

Grenoble INP - Grenoble Alps University

maciej.korczynski@univ-grenoble-alpes.fr

Maarten Wullink, SIDN Labs

maarten.wullink@sidn.nl

Acknowledgements

This study was commissioned by the Competition, Consumer Trust, and Consumer Choice Review Team with the support of ICANN.

We would like to thank ICANN, Domain-Tools, Whois XML API, Spamhaus, SURBL, StopBadware, CleanMX, Secure Domain Foundation, Anti-Phishing Working Group for providing access to their data.

Authors also thank Roland van Rijswijk for his help in obtaining additional domain data.