Domain names abuse and TLDs: from monetization towards mitigation

Giovane C. M. Moura, Moritz Müller, Marco Davids,
Maarten Wullink, and Cristian Hesselman
SIDN Labs
{first.lastname}@sidn.nl

DNS OARC Meeting
May 15th, 2017
Madrid, Spain

Introduction

- DNS provides a simple label for hosts, services, applications on the Internet
- Often, it is misused in malicious activities such as:
  - phishing campaigns
  - malware
  - spam
- Underlying each type of abuse, a different business model
  - provides the incentives for the crooks to keep on

Introduction

- Plenty of research work in curbing DNS-related abuse [1, 2, 3, 4, 5, 6]
  - With a clear contribution
- But, they suffer from similar issues:
  1. Bound by dataset type/duration
  2. Cover specific attacks; missing broader view on all abuses
- **This paper:**
  1. Cover first issue with longitudinal measurements and registration (`.nl`)
  2. Present a survey on domain abuses from the point of view of a TLD operator (centralized view)

SIDN LABS

Motivation: why doing this?

Came from a situation we faced :

- **There's no one size fits all**
- we have all this data
- how to better use it?
- where to begin with?
  - e.g.: malicious registered phishing or compromised phishing?
  - or other sort of abuse?
  - how to prioritize it?
  - **Which datasets too look first?**
- Other TLD operators may be facing the same problem

Understanding business models

- Helps you to understand how money is made
- And how it impact your datasets
- It's been done many times in Internet abuse. E.g.: PharmaLeaks[7].
- Business model $\rightarrow$ abuse $\rightarrow$ money

TLD Operations and Datasets

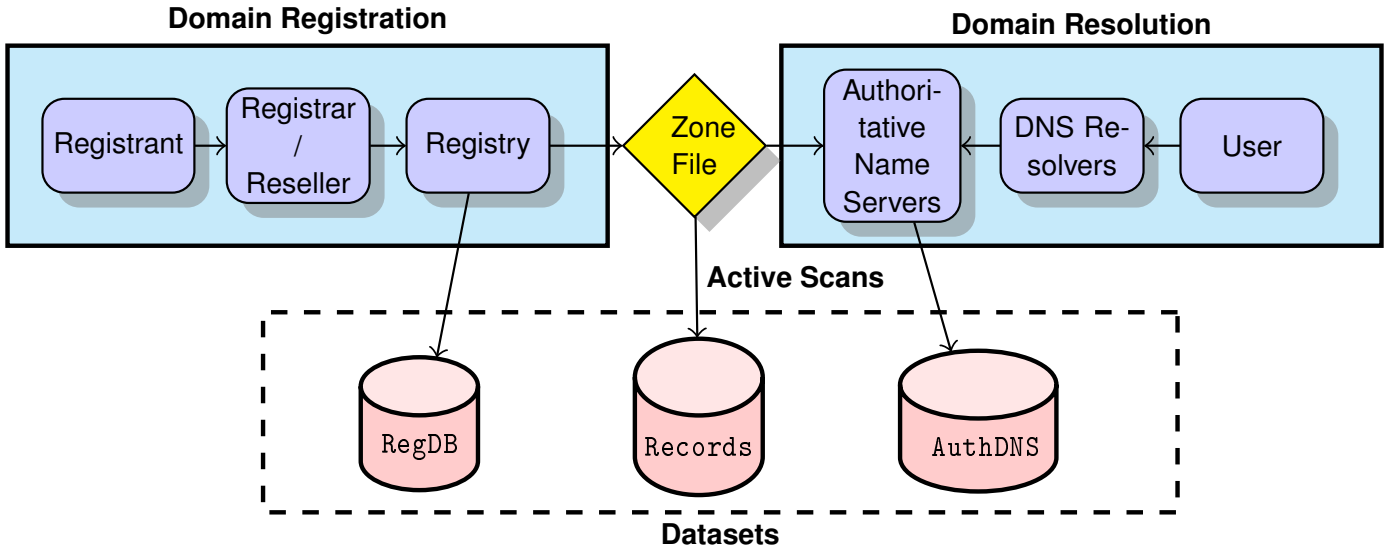

Figure: TLD Operations: registration (left), domain name resolution (right), and derived datasets.

- `RegDB`: your registration DB
- `Zone File Scans`: in our case, the `OpenIntel.nl` project
- `AuthDNS`: data from auth servers, we use ENTRADA [8]

Business Models Survey and Lit Review

Business	**Spam**	`RegDB`	`AuthDNS`	`Records`	**Lit**
Phishing(0-day)	Yes	Weak	Strong	Weak	[3, 6]
Phishing(comp.)	Yes	None	Strong	Weak	[9]
Parking (Ads)	No	Strong	Weak	Strong	[10, 11]
Parking (Mal)	No	Strong	Weak	Strong	[10, 11]
Fake Goods	Yes	Weak	Weak	Medium	[6, 7]
Drop-Catch	No	Medium	Medium	Weak	[12]
Botnet C&C	No	Medium	Strong	?	[13]
Blackhat SEO	No	Medium	Medium	Strong	[14, 15]

Table: Business Models and Datasets/signal “strength”, and research works that cover those.

SIDN LABS

Phising (0-day)

- Two types of phishing: compromised and 0-day (newly registered)
- 0-day phishing business model:
  1. Registered domain(s)
  2. Large spam campaign at the same time
  3. ID theft (ID, credit card, etc).
  4. Money: selling the data, using it themselves

Phising (0-day)

- Datasets:
  1. `Records`: harder to detect, IP/registrar reputation
  2. `RegDB`: hard but possible to detect (it's been done for spamming domains [16])
  3. `AuthDNS`: strongest signal, but after attack has started [3, 6]

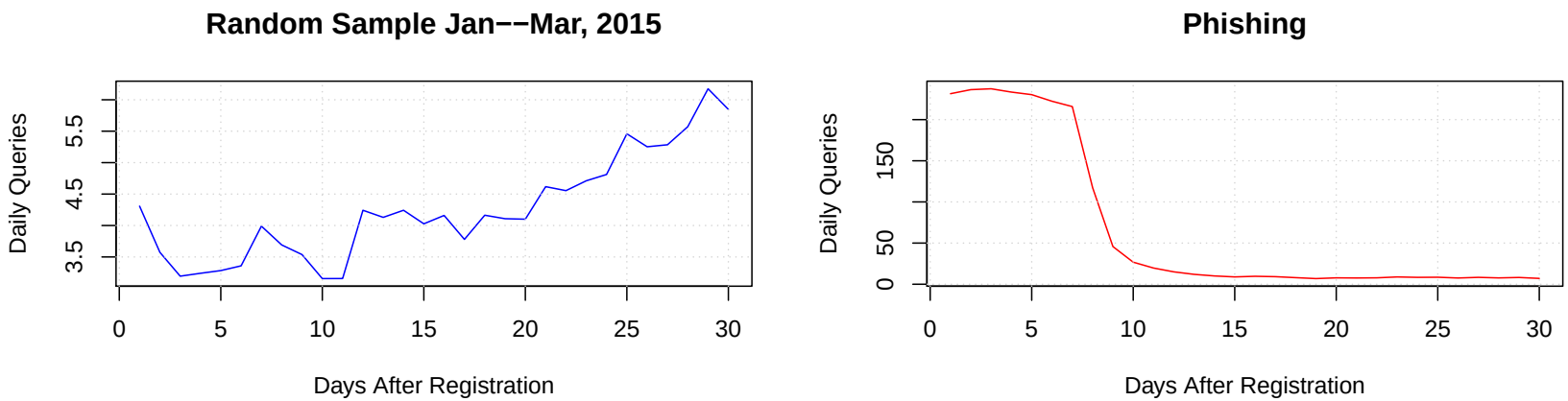

Figure: .nl Random vs Phishing new domains average daily queries [6]

Phising (compromised)

- Most common sort of phishing
- Typically on hacked CMSes, instead of newly registered
- Business model:
  1. Hack a website
  2. Sam campaign at the same time
  3. ID theft (ID, credit card, etc).
  4. Money: selling the data, using it themselves

Phising (compromised)

- Datasets:
  1. `Records`: harder to detect, typically no changes
  2. `RegDB`: also, usually no changes in here
  3. `AuthDNS`: possible to detect, very hard to tell false positives source

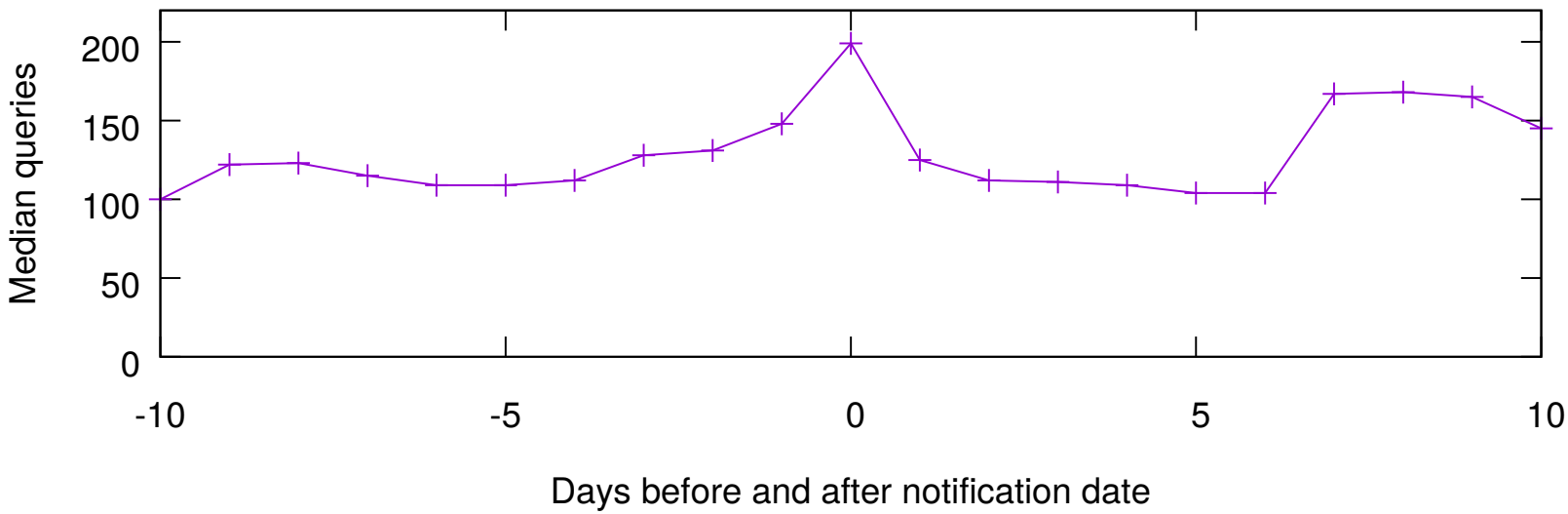

Figure: Median daily queries for 1,374 compromised phishing sites on `.nl`, before and after Netcraft's notification

Parking

- Parking is a big industry
- Business model:
  1. Register many domains (bulk)
  2. Wait for traffic to come in
  3. Redirect to ad networks
  4. Money:
     - Legal: ad networks
     - Illegal: malicious, ID-theft

Parking

- Datasets:
  1. `Records`: can be done, same ASes, IPs, etc
  2. `RegDB`: Yes, bulk registrations, same registrar, etc.
  3. `AuthDNS`: usually not the case

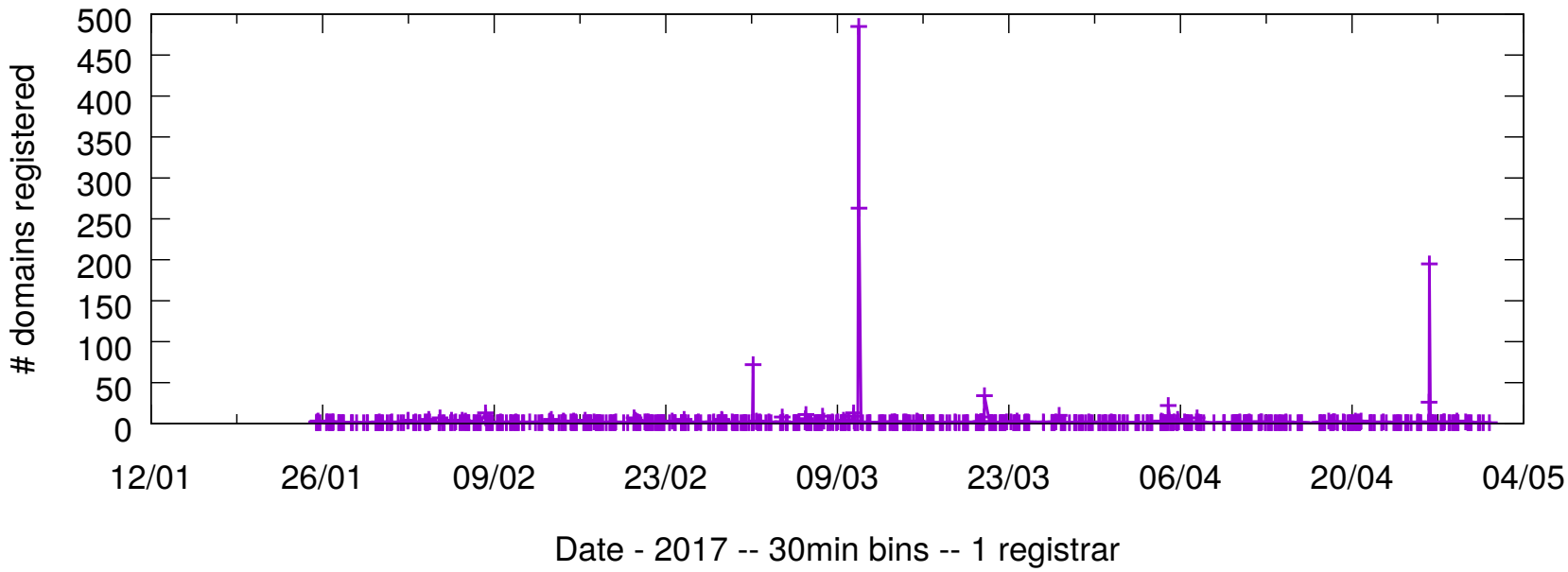

Figure: Number of domains registered for one registrar, in every 30min – spikes indicates anomalies

Parking

- Datasets:
  1. `Records`: can be done, same ASes, IPs, etc
  2. `RegDB`: Yes, bulk registrations, same registrar, etc.
  3. `AuthDNS`: usually not the case

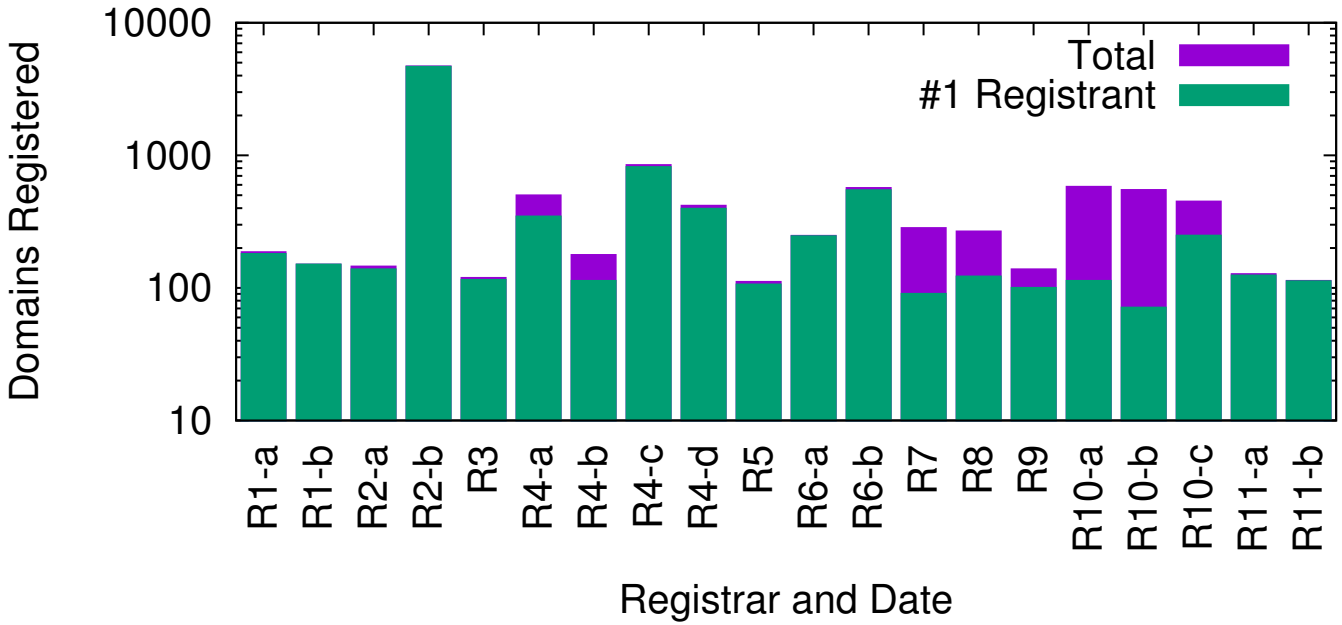

Figure: Anomalous registrations for Registrars and Top 1 registrant – most of registrations are done in bulk by 1 registrant.

Parking

- Key aspect: tell ad networks from malicious (e.g.: malicious redirection)
- Malicious redirection type has more incentives to use a new e-mail addresses during registrations (and no reuse)
- We've seen that for ad nets
- Need to develop a solution that address this (open)

Fake Goods

- When we develop nDEWS [8] to detect 0-day phishing, we notice a lot of domains were neither phishing neither false positives
- Their frequency and continuity suggested a profitable business model
- Just like phishing (0-day) business model, and detection too
- This sort of abuse falls into a “gray area”:
  - not as bad as phishing
  - still bad because of ID theft
  - hard to tell if it’s fake or not
- Detection: similar to 0-day phishing

SIDN LABS

Botnet C&C

- Domains can be used also for botnet command-and-control channels
- Domain generation algorithms (DGA) typically used for that
- Bots are supposed to contact a new domain every *x* time
- DGAs generated many, but only a few are registered , to avoid detection

Botnet C&C

- Business model: registration
- Datasets:
  1. `RegDB`: registration of “weird” looking names
  2. `Records`: in combination with the previous
  3. `AuthDNS`: NXdomain queries for non registered DGAs

Summary

- DNS abuse has been active for many years
- There are many types, which different business models
- Business models of DNS abuse impact datasets differently
- TLDs ops should develop applications according to business models
  - no one-size fits all
- Which one first?
  - that depends on the frequency of the abuse on their zone
- This paper presents a **survey** and a **discussion** on which datasets can be used
  - And some of our experience with these abuses on `.nl`

Questions?

- Contact:
  - http://sidnlabs.nl
  - giovane.moura@sidn.nl
  - Twitter: @giomourasec
- Thank you for your attention

Download our paper at:

https://www.sidnlabs.nl/downloads/publications/dissect2017.pdf

Bibliography I

[1] S. Hao, A. Kantchelian, B. Miller, V. Paxson, and N. Feamster, " PREDATOR: Proactive Recognition and Elimination of Domain Abuse at Time-Of-Registration," in *Proceedings of the 2016 ACM CCS*, October 2016.

[2] S. Hao, M. Thomas, V. Paxson, N. Feamster, C. Kreibich, C. Grier, and S. Hollenbeck, "Understanding the Domain Registration Behavior of Spammers," in *Proceedings of the 2013 Conference on Internet Measurement Conference*, ser. IMC '13. New York, NY, USA: ACM, 2013, pp. 63–76.

[3] Hao, Shuang and Feamster, Nick and Pandrangi, Ramakant, "Monitoring the Initial DNS Behavior of Malicious Domains," in *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, ser. IMC '11. New York, NY, USA: ACM, 2011.

SIDN LABS

Bibliography II

[4] M. Antonakakis, R. Perdisci, W. Lee, N. Vasiloglou II, and D. Dagon, “Detecting Malware Domains at the Upper DNS Hierarchy.” in *USENIX Security Symposium*, 2011, pp. 16–32.

[5] M. Antonakakis, R. Perdisci, D. Dagon, W. Lee, and N. Feamster, “Building a Dynamic Reputation System for DNS.” in *USENIX security symposium*, 2010, pp. 273–290.

[6] Giovane C. M. Moura, Moritz Muller, Maarten Wullink, and Cristian Hesselman, “nDEWS: a New Domains Early Warning System for TLDs,” in *IEEE/IFIP International Workshop on Analytics for Network and Service Management (AnNet 2016), co-located with IEEE/IFIP Network Operations and Management Symposium (NOMS 2016)*, April 2016.

SIDN LABS

Bibliography III

[7] D. McCoy, A. Pitsillidis, G. Jordan, N. Weaver, C. Kreibich, B. Krebs, G. M. Voelker, S. Savage, and K. Levchenko, "PharmaLeaks: Understanding the Business of Online Pharmaceutical Affiliate Programs," in *Proceedings of the 21st USENIX Security Symposium*. Bellevue, Washington, USA: USENIX Association, August 2012.

[8] Maarten Wullink, Giovane C. M. Moura, Müller, M, and Cristian Hesselman, "ENTRADA: a High Performance Network Traffic Data Streaming Warehouse," in *Network Operations and Management Symposium (NOMS), 2016 IEEE* , April 2016.

[9] A. Noroozian, M. Korczynski, S. Tajalizadehkhoob, and M. van Eeten, "Developing security reputation metrics for hosting providers," in *8th Workshop on Cyber Security Experimentation and Test (CSET 15)*, 2015.

SIDN LABS

Bibliography IV

[10] S. Alrwais, K. Yuan, E. Alowaisheq, Z. Li, and X. Wang, "Understanding the Dark Side of Domain Parking," in *23rd USENIX Security Symposium (USENIX Security 14)*. San Diego, CA: USENIX Association, Aug. 2014, pp. 207–222.

[11] T. Vissers, W. Joosen, and N. Nikiforakis, "Parking Sensors: Analyzing and Detecting Parked Domains." in *Proceedings of the 2015 Network and Distributed System Security Symposium (NDSS 2015), San Diego, California, USA.*, 2015.

[12] C. Lever, R. Walls, Y. Nadji, D. Dagon, P. McDaniel, and M. Antonakakis, "Domain-Z: 28 Registrations Later," In: *Proceedings of the 37th IEEE Symposium on Security and Privacy*. San Jose, California., 2016.

SIDN LABS

Bibliography V

[13] B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna, "Your botnet is my botnet: analysis of a botnet takeover," in *Proceedings of the 16th ACM conference on Computer and communications security*. ACM, 2009.

[14] niche site project, "Private Blog Networks http://nichesiteproject.com/private-blog-networks/," Sep. 2016.

[15] K. Du, H. Yang, Z. Li, H. Duan, and K. Zhang, "The Ever-Changing Labyrinth: A Large-Scale Analysis of Wildcard DNS Powered Blackhat SEO," in *25th USENIX Security Symposium (USENIX Security 16)*. USENIX Association.

SIDN LABS

Bibliography VI

[16] S. Hao, A. Kantchelian, B. Miller, N. Feamster, and V. Paxson, "PREDATOR: Proactive Recognition and Elimination of Domain Abuse at Time-Of-Registration," in *Proceedings of the SIGSAC Conference on Computer and Communications Security*, 2016.

SIDN LABS