Detecting and preventing DNS abuse in .eu

Lieven Desmet, KU Leuven – lieven.desmet@cs.kuleuven.be

Malicious use of domain names

- Domain names are often abused by cyber criminals
  - Spam, botnet C&C infrastructure, phishing, malware, …
- To avoid blacklisting, malicious actors often deploy a hit-and-run strategy
  - 60% are only active for 1 day after registration [Hao et al]

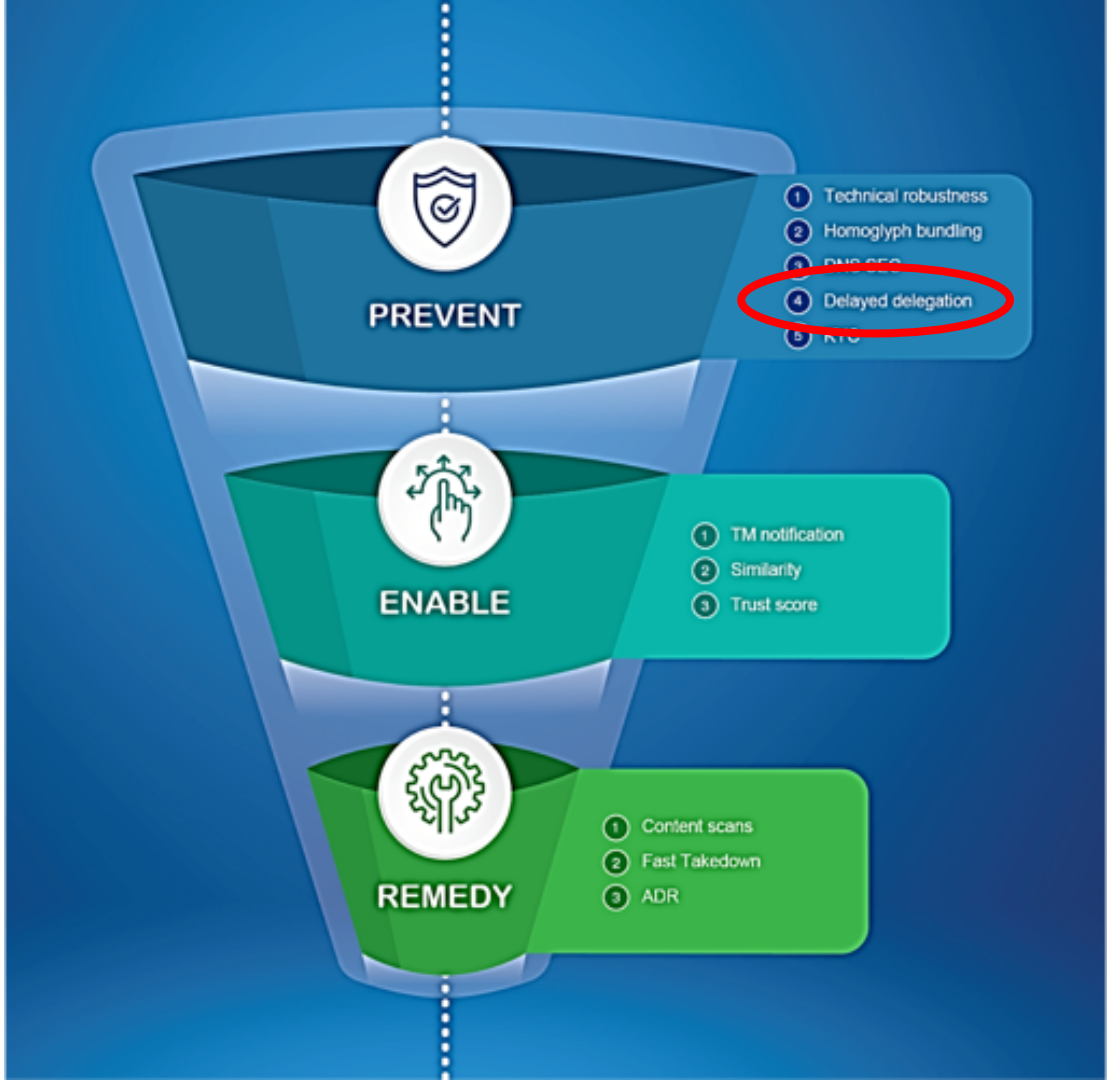
[Hao et al] “Understanding the Domain Registration Behavior of Spammers” IMC 2013

Research hypothesis:

“Malicious actors register domains in bulk, and do so for longer periods of time.”

The .eu trust strategy

- Delayed delegation
  - Predict at time of registration whether a domain name will be used abusively



Powered by EURid

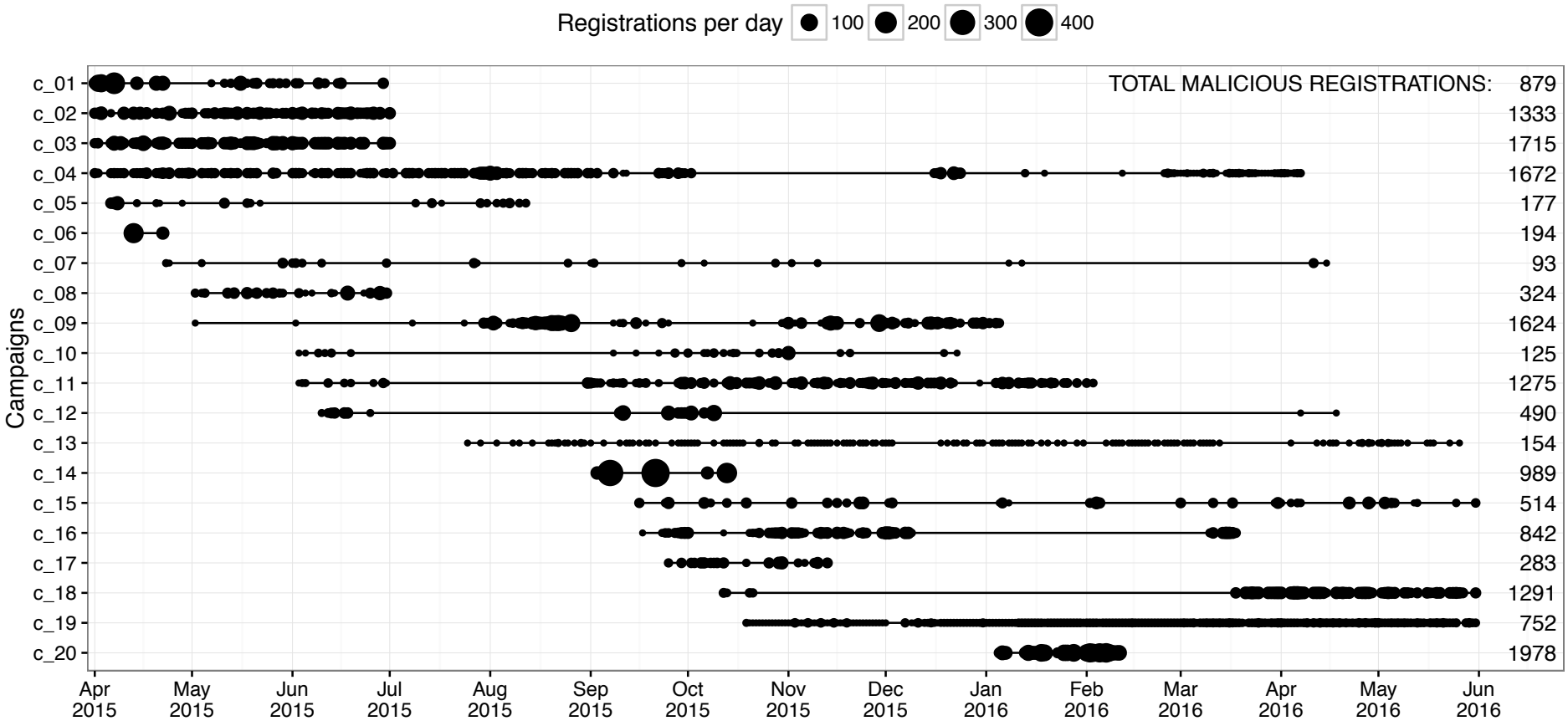
https://link.eurid.eu/prediction1


Exploring the ecosystem of malicious domain registrations in the .eu TLD

Insights in malicious domain registrations

T. Vissers et al., ***Exploring the ecosystem of malicious domain registrations in the .eu TLD****, Research in Attacks, Intrusions, and Defenses (RAID 2017), September 2017.*

Activity of identified campaigns



Insight 1: Varying campaign characteristics

- Simple campaign (c_14)
- Single (fake) registrant used throughout the campaign

- **41 days active**
- **989 blacklisted registrations**
  (= 95.37%)

Example campaign (c_11)

- Multiple fake registrant details
  - Combinations of
    - 2 email accounts,
    - 3 phone numbers,
    - 4 street addresses

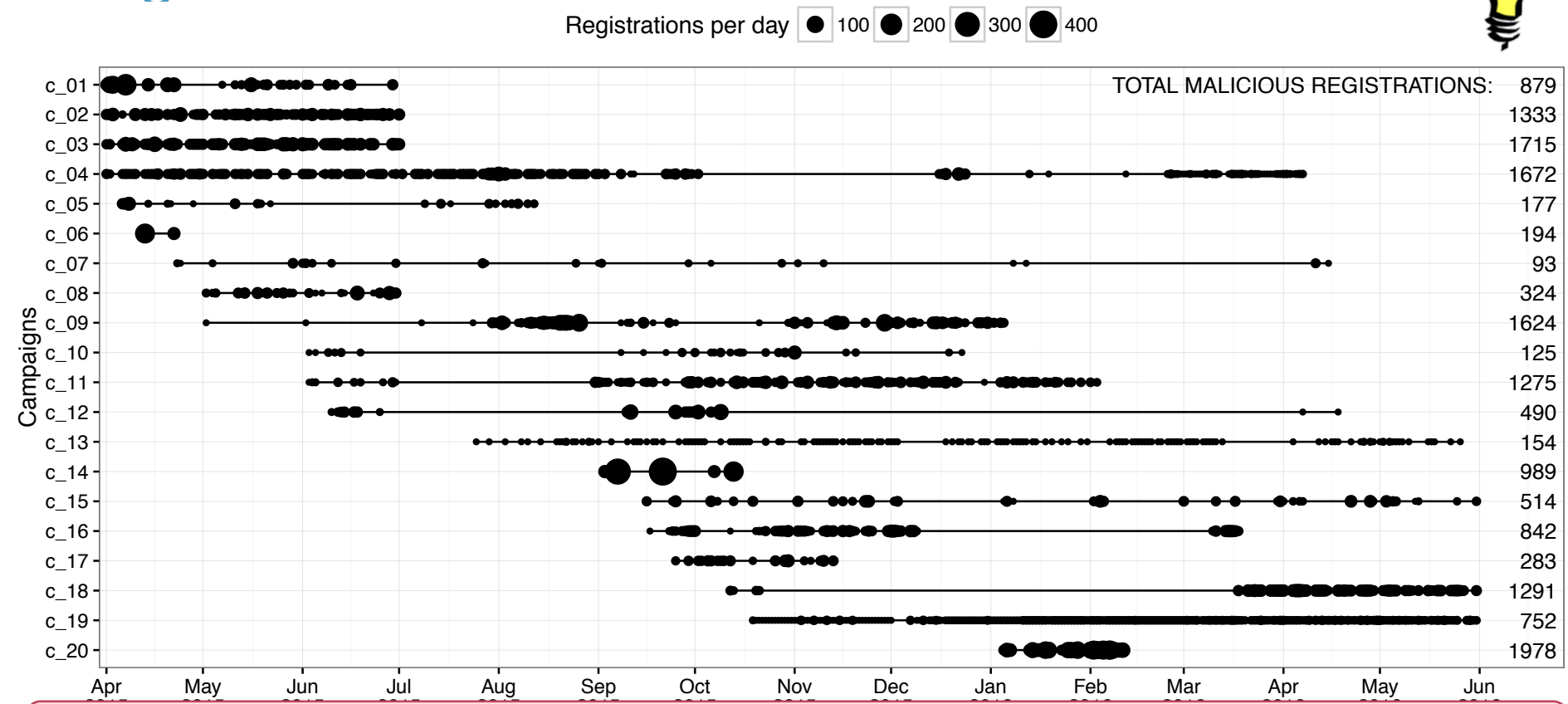
- **8 months active**
- **1,275 blacklisted registrations**
  (= 53.96%)

Example of an advanced campaign (c_15)

- Registrant details:
  - 98 fake registrants
  - Generated by Laravel Faker tool
- Domain names:
  - Consist out of 2-3 Dutch words
  - Dutch words are reused across registrants
- Batches of 8, 16, 24 or 32 registrations

- **8+ months active**
- **514 blacklisted registrations**
  (= 26.95%)

Insight 2: Small set of malicious actors

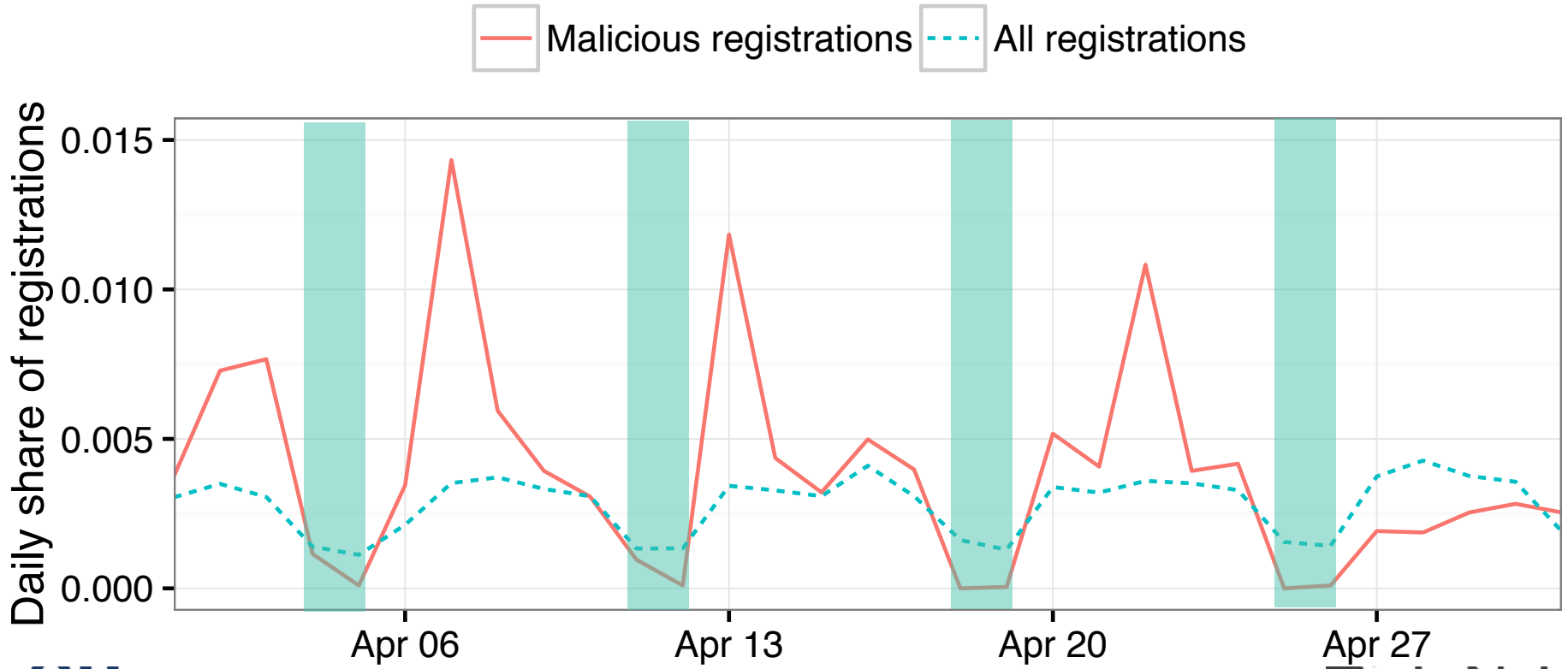


At most 20 actors represent 80% of malicious registrations

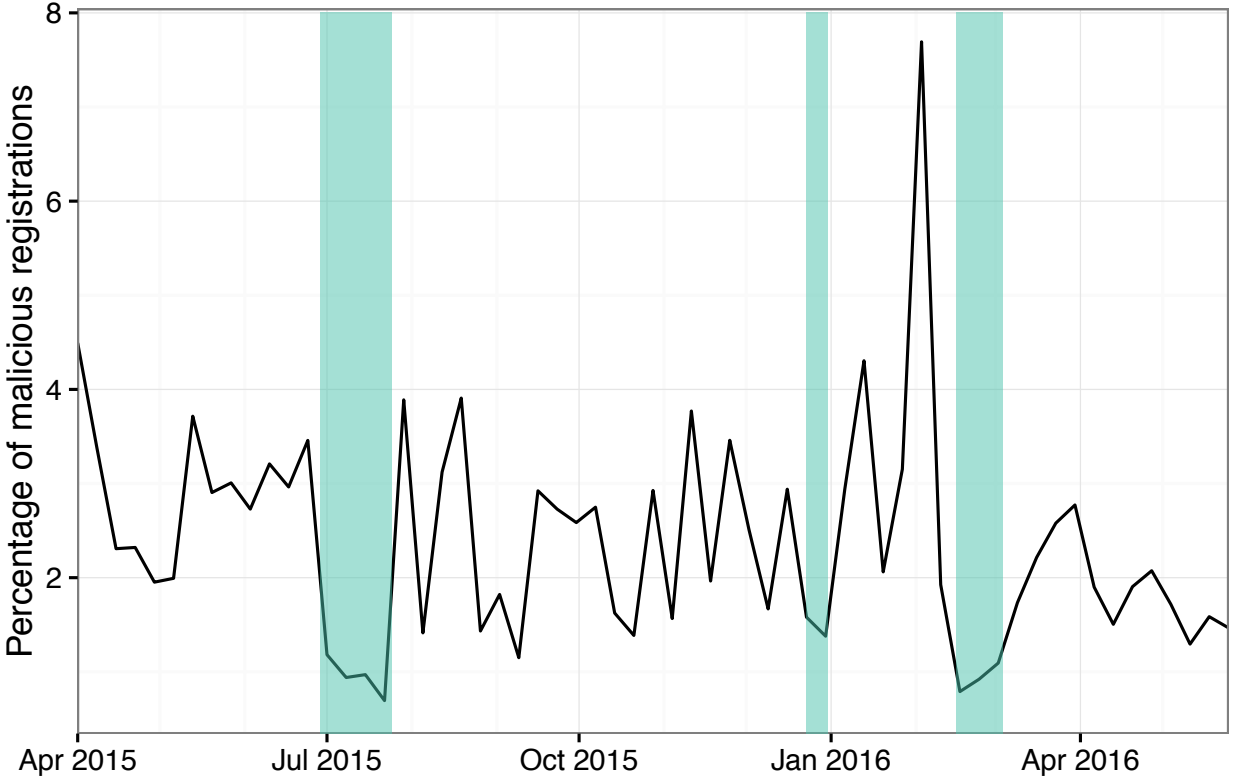
Insight 3: Top facilitators for malicious registrations

	Nb of malicious	Contribution Malicious	Benign	Toxicity
1. registrar_5	10,353	49.61%	2.27%	36.25%
2. registrar_3	3,004	14.39%	2.64%	12.41%
3. registrar_7	2,327	11.15%	0.46%	38.67%
1. gmail.com	4,221	20.23%	24.79%	2.08%
2. yahoo.com	3,348	16.04%	1.49%	21.85%
3. aol.com	2,134	10.23%	0.31%	46.28%

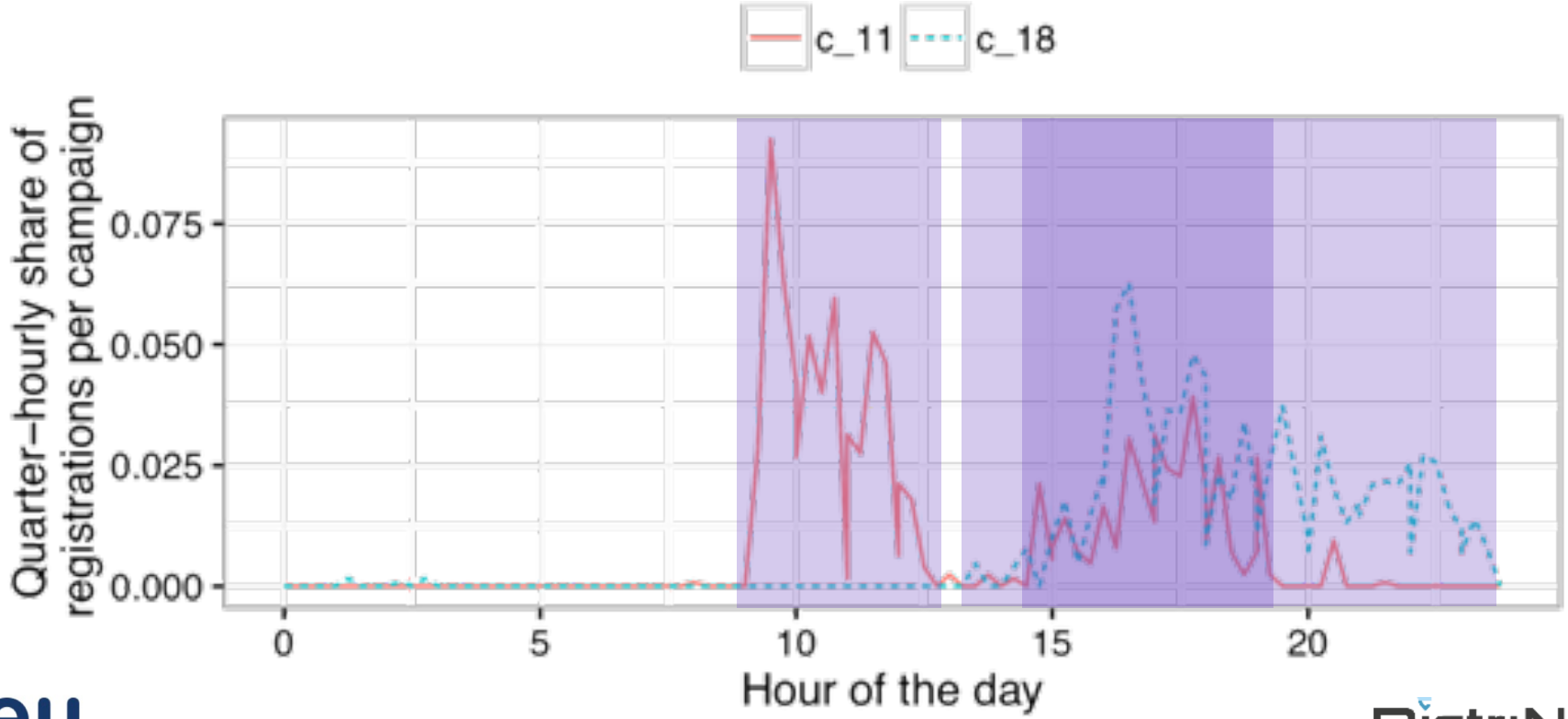
Insight 4: Some campaigns align with regular business activity patterns (1)



Insight 4: Some campaigns align with regular business activity patterns (2)



Insight 4: Some campaigns align with regular business activity patterns (3)



https://link.eurid.eu/prediction4


PREMADOMA: An Operational Solution for DNS Registries to Prevent Malicious Domain Registrations

Abstract

The Domain Name System is one of the most essential components of the Internet, mapping domain names to the IP addresses behind almost every service on the Internet. Domain names are therefore also a fundamental tool for attackers to quickly locate and relocate their malicious activities on the Internet. In this paper, we design and evaluate PREMADOMA, a fully-operational machine learning system which enables a DNS registry to predict malicious intent well before a domain name becomes operational. In contrast to blacklists, which only offer protection after some harm has already been done, this system can prevent domain names from being used before they can pose any threats. We advance the state of the art by leveraging recent insights into the ecosystem of malicious domain registrations, focusing explicitly on bulk registration behavior and similarity patterns in registrant information. We succesfully deploy PREMADOMA in the production environment of a top ccTLD registry and contribute to the take down of 74,036 registrations in 2018.

1 Introduction

Domain names remain a major facilitator of cyberattacks. Malicious actors continuously deploy domains in their cybercriminal operations, such as spam, phishing, malware distribution and botnet C&C. Due to this crucial role in cybercriminal operations, stopping malicious domain names has become a highly important security objective.

The most well-known countermeasure for malicious domains is a *blacklist*. So-called "reputation providers" curate lists of domain names that are associated with internet-based attacks. Typically, they use honeypot tactics, such as spam traps, to detect new malicious domains. Various software and services consult these blacklists and block incoming or outgoing communication with listed domains accordingly. Blacklists have become more agile and at this time domain names are blocked quickly after exhibiting attacking behavior.

In response, miscreants have adopted hit-and-run strategies. Specifically, they anticipate their malicious registrations to have a short lifespan and counter this by using a series of disposable "*burner domains*" to sustain their malicious operations. This results in large-scale *campaigns*, i.e. malicious actors that register thousands of domains [3]. Therefore, post-factum detections, such as blacklists, are becoming limited in their effects [14].

This situation expresses the need to block malicious domain registrations before they are able to execute any attack behavior. Hence, more recent security research aims to shift to earlier detection of malicious domain names. In particular, research by Hao et al. [10] proposed to determine the maliciousness of domain names *at the time of registration*. To be practically implemented, such a strategy requires cooperation of a party involved in the registration procedure, i.e. DNS registries or registrars.

In this paper, we focus on the *real-world operational aspects* of designing and implementing a DNS registry's security system that is able to detect malicious domains at registration time. We take into account the operational and quality-related aspects of deploying such a system in the context of critical internet infrastructure environment at a top ccTLD registry.

1.1 PREMADOMA prediction strategy

The main goal of the PREMADOMA system is to reduce the amount of cybercriminal operations by detecting and preventing malicious registrations at registration time. Based on insights of the malicious domain registration ecosystem, we aim to design PREMADOMA such that it accurately predicts whether or not a domain registration has malicious intent. By applying an automated and adaptive mitigation strategy, PREMADOMA aims to substantially increases the cost for attackers in order to disincentivize malicious actors to launch campaigns.

Ecosystem insights Malicious online activities do typically not occur in an isolated or dispersed fashion [6, 11]. Instead, cybercriminals involve multiple, tightly related abusive strategies, techniques and targets.

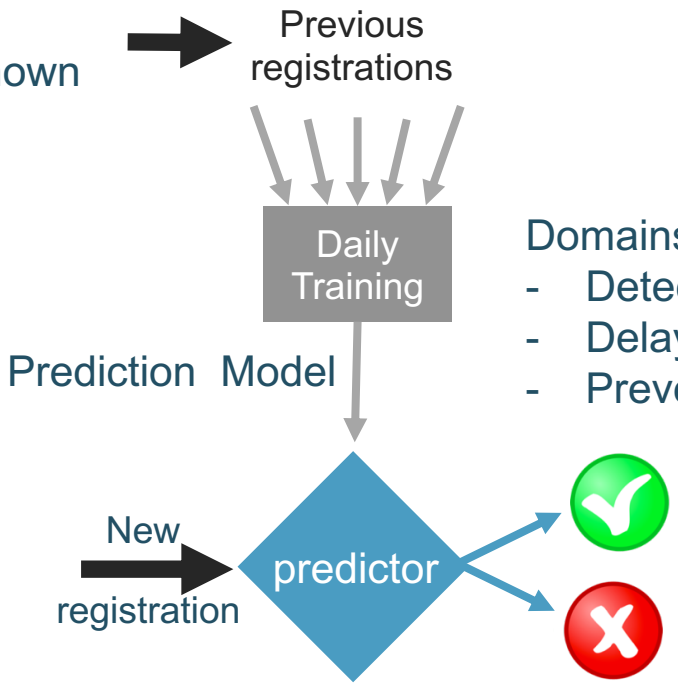
1

Registration-time prediction of malicious intent

J. Spooren et al., ***PREMADOMA: An Operational Solution for DNS Registries to Prevent Malicious Domain Registrations***, Annual Computer Security Applications Conference (ACSAC 2019), *December 2019.*

Pro-active detection and prevention

Previous registrations for which the results (abuse/no abuse) is known



Domains with malicious intent can be
- Detected early
- Delayed
- Prevented from being registered

For each new registration, the system predicts if the domain will be used for malicious activity

Underlying assumptions/rationales for our predictors

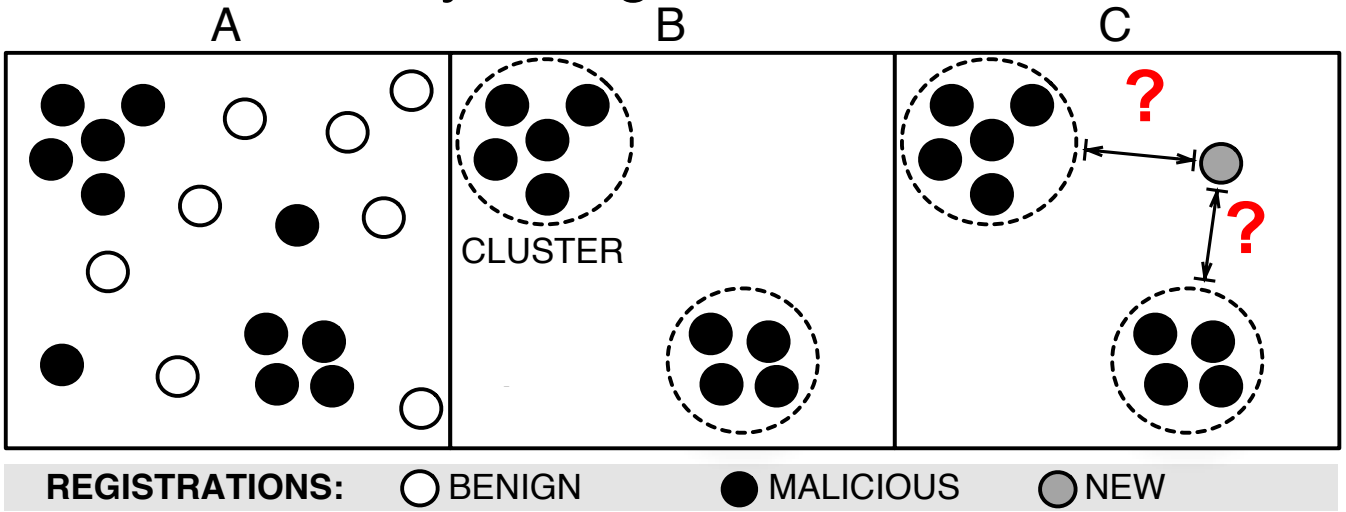
- Similarity-based agglomerative clustering
  - Domains belonging to the same campaign have very similar registration details
- Reputation-based classification
  - Domains using registration facilitators with a bad reputation (e.g. email providers or registrars), are likely to be malicious as well

Predictor 1:
Reputation-based classification

- Reputation features of “facilitators”
- Facilitators:
  - Technical facilitators: registrar, name servers
  - Communication means: email provider and phone number
- Reputation score:
  - Represent contribution and toxicity of facilitator to malicious registrations

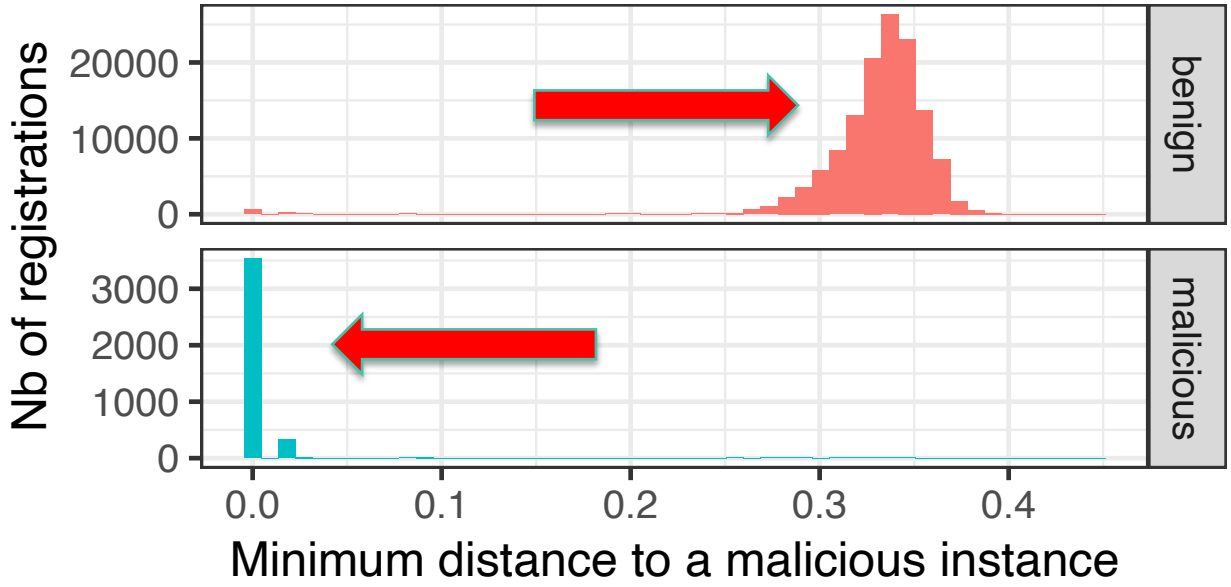
Predictor 2:
Similarity-based clustering

- Agglomerative clustering of malicious samples
- Based on the similarity of registration data



Can we differentiate between benign and malicious samples?

› Closest distance of a registration to malicious domain

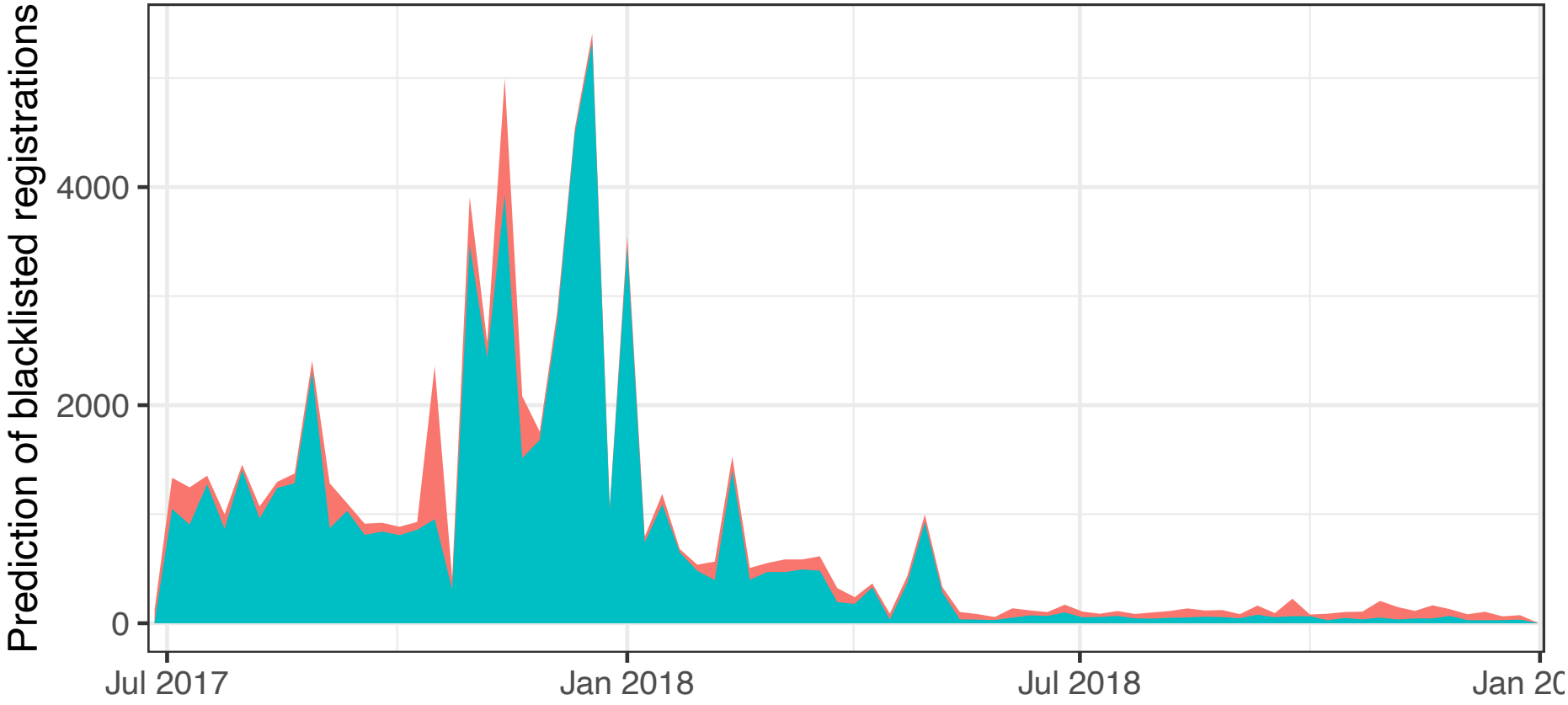


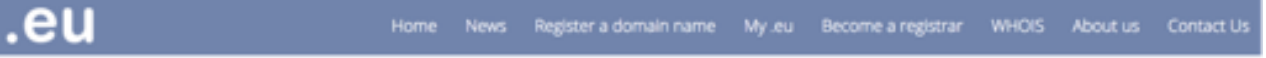
Evaluation on historical data

- Ground truth-based evaluation
  - Recall: 66.23%
  - Precision: 84.57
  - False positive rate: 0.30%

- Campaign-based evaluation
  - 17 out of the 20 campaigns are well predicted

Detecting and preventing abuse in .eu: “1 picture …”





Over 25 000 domain names suspended with ties to identity fraud

Tweet

« Back to the news page

On 29 January 2018, EURid susp

With actions as such, our focus i
enforcement, both on a nationa
towards building the most trust
illegal activity online. "With our t
names for potential abuse, leadi
EURid Legal Manager.

In 2017, we suspended 20 126 c
enforcement.

Over 11 000 abusive domain names suspended

Tweet

« Back to the news page

On 21 June 2018, EURid suspended 11 760 domain names that were registered with non-eligible registration data, of which some have been reported for abuse.

With actions as such, our focus is on the safety of online consumers. Via close collaborative efforts with law
l, as well as with our registrar channel, we
in name space, taking a stand against

onitor our domain names for potential
pared to 2017, where we suspended 20 126
abusive domain names, we're up to 36 336 abusive domain name suspensions thus far in 2018." said Geo Van Langenhove, EURid Legal Manager.

Learn more about the ways we're building a trustworthy .eu and .ею domain name space at trust.eurid.eu.

Predictive Algorithms

Through the use of historical data and self-learning algorithms, we are working to predict at the time of registration whether or not a domain name might be used in an abusive way in an effort to prevent such malicious domain names from becoming active in the first place.

As part of the EURid's Trust & Security program, 58,966 domains were suspended in 2018.

Powered by EURid

Operational results

- Period: July 2017 – December 2018 (18 months)
  - Recall: 85.51%
  - Precision: 72.04%
  - False positive rate: 2.86%

- Very big campaigns (October 2017 - March 2018)
- Incomplete ground truth

https://link.eurid.eu/prediction3


Assessing the Effectiveness of Domain Blacklisting Against Malicious DNS Registrations

Thomas Vissers*, Peter Janssen†, Wouter Joosen*, Lieven Desmet*
*imec-DistriNet, KU Leuven
†EURid VZW

Ground truth analysis

T. Vissers et al., ***Assessing the Effectiveness of Domain Blacklisting Against Malicious DNS Registrations****, IEEE Workshop on Traffic Measurements for Cybersecurity (WTMC 2019), May 2019.*

Sources of ground truth

Google Safe Browsing

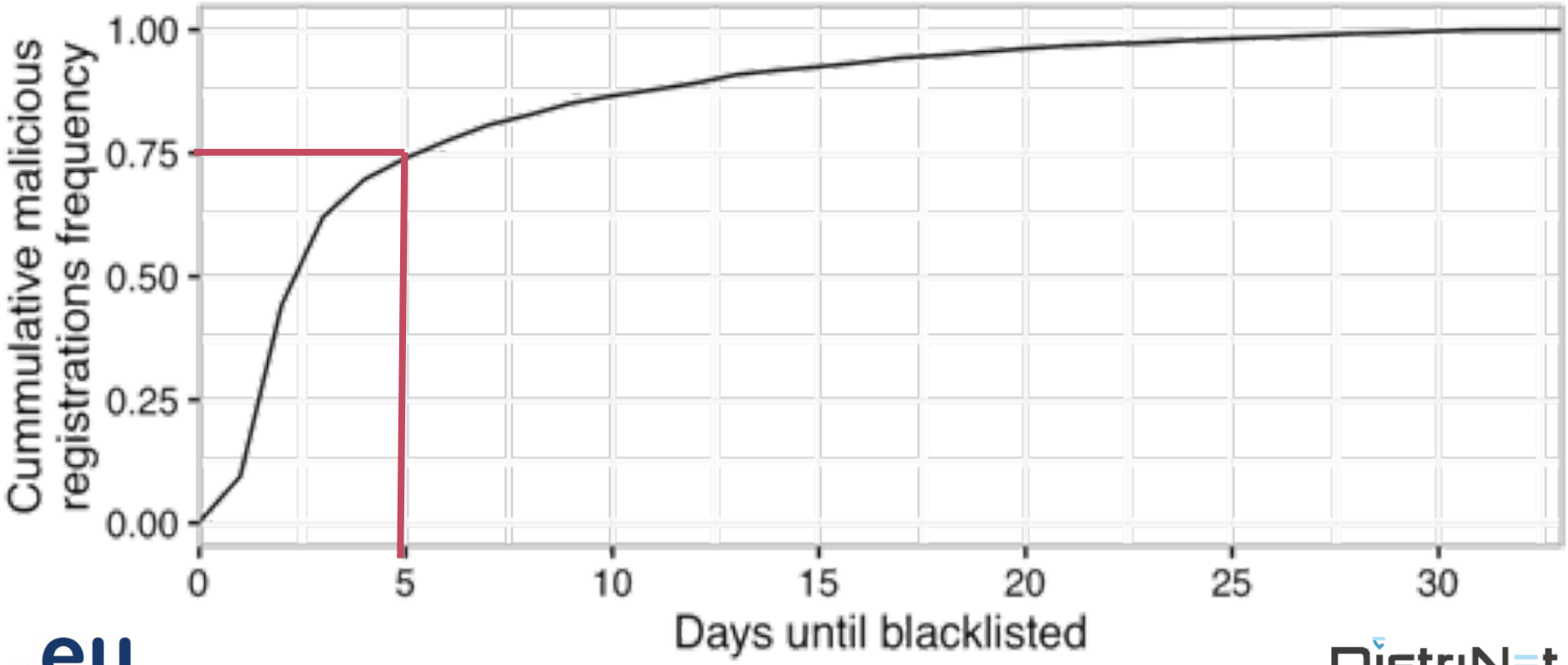
- Around 60K domains to check per day
- Simplified view: once on a abuse list, always considered malicious

Types of abuse recorded

- Majority of abuses are related to spam (93.68%)

- Different coverage statistics per abuse list for .eu:
  - Spamhaus DBL: 81.07%
  - SURBL multi list: 50.04%
  - Google Safe Browsing: 1.81%

Delay of the ground truth



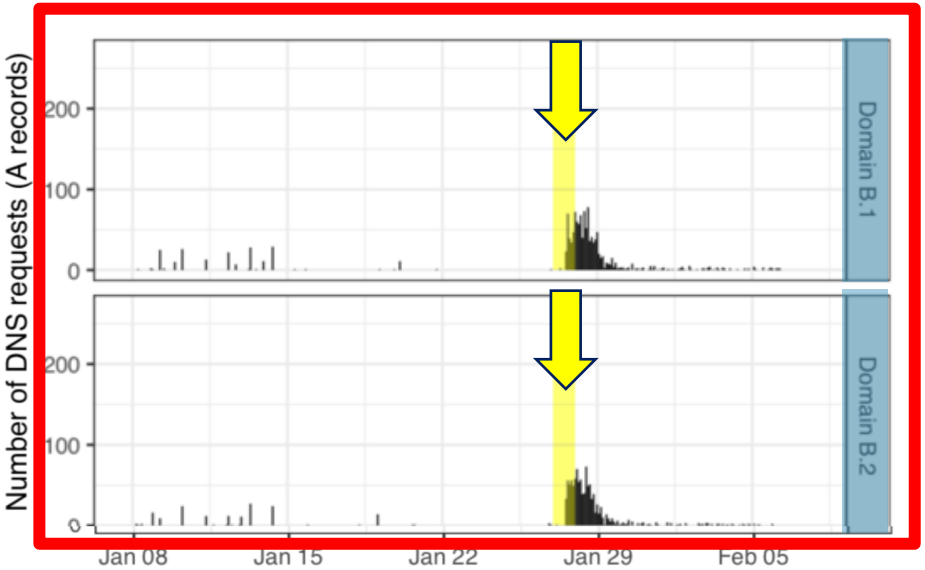
Incompleteness of the blacklists

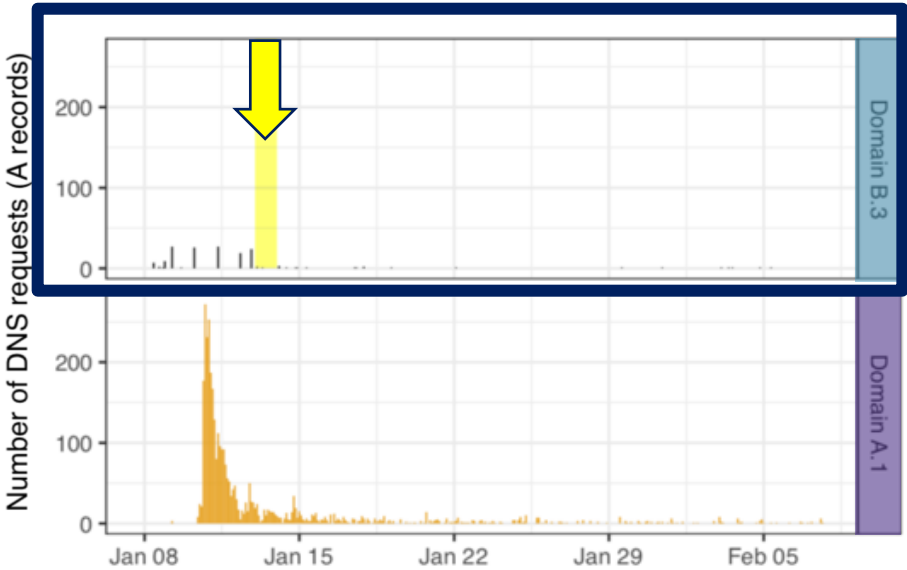
- Failed to detect?
- Never active/malicious?

	Active	Dormant
Blacklisted	Blocked	Pro-actively blocked
Non-blacklisted	**Missed**	**Unused**

Campaign related activity

- E.g. spam triggers multiple DNS requests:
  - SPF, DMARC, DKIM, MX, A



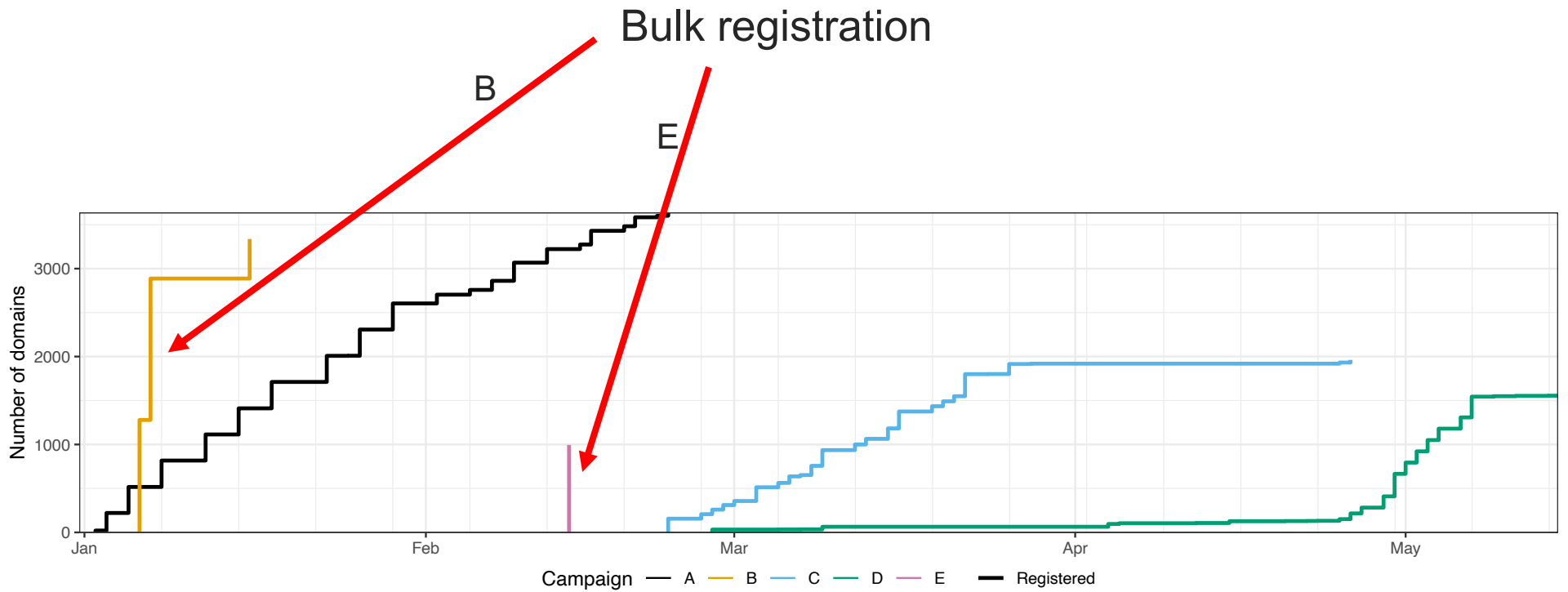


Active vs Dormant – Blacklisted vs Non-blacklisted

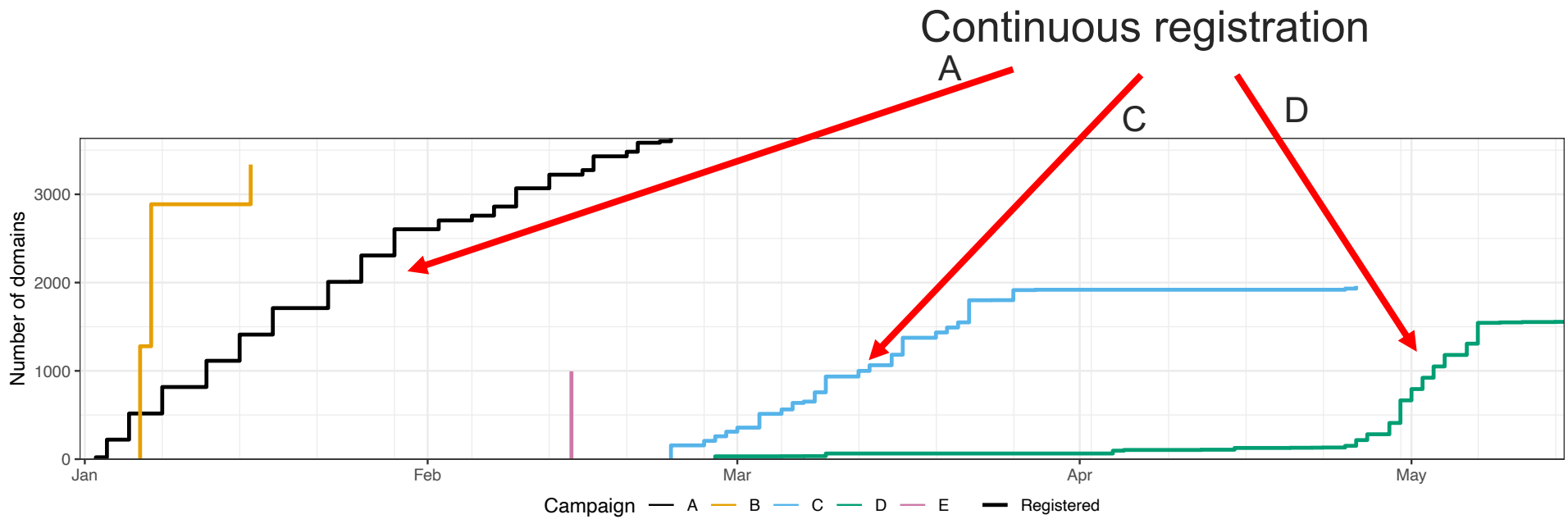
- 5 largest campaigns in .eu (Q1-Q2 2018)
- Based on passively-logged DNS requests (.eu TLD server)

	Active	Dormant
Blacklisted	**Blocked** 54.8%	**Proactive** 2.9%
Non-blacklisted	**Missed** 14.1%	**Unused** 14.0%

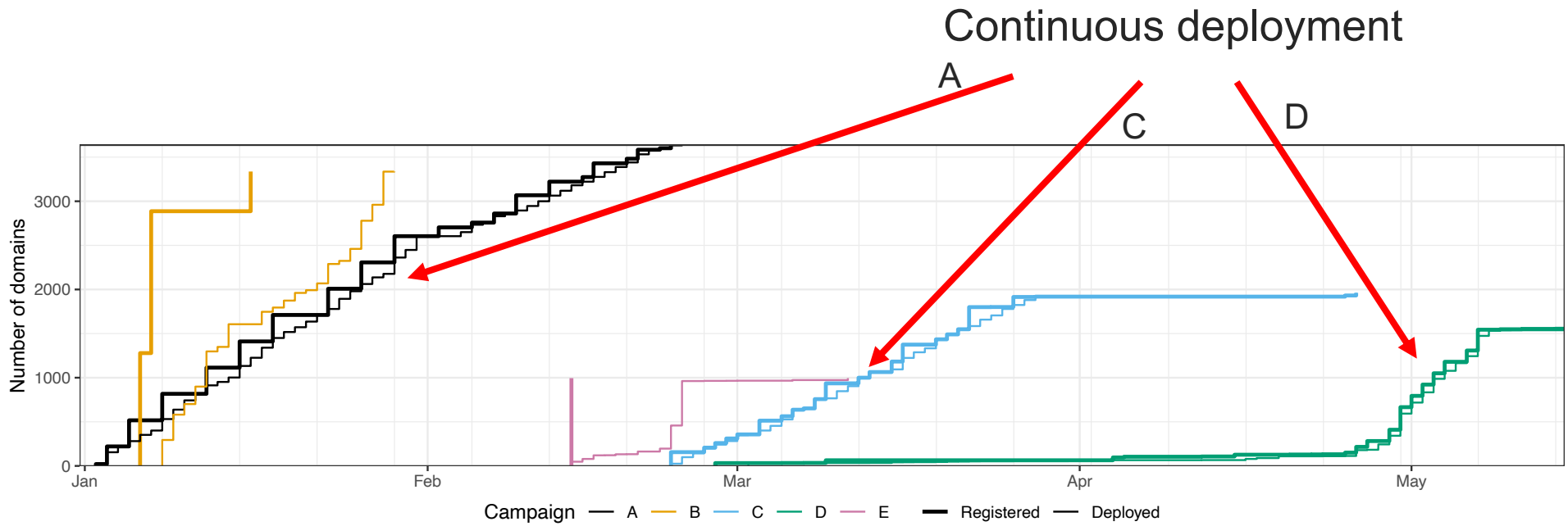
1. Registration strategy



1. Registration strategy

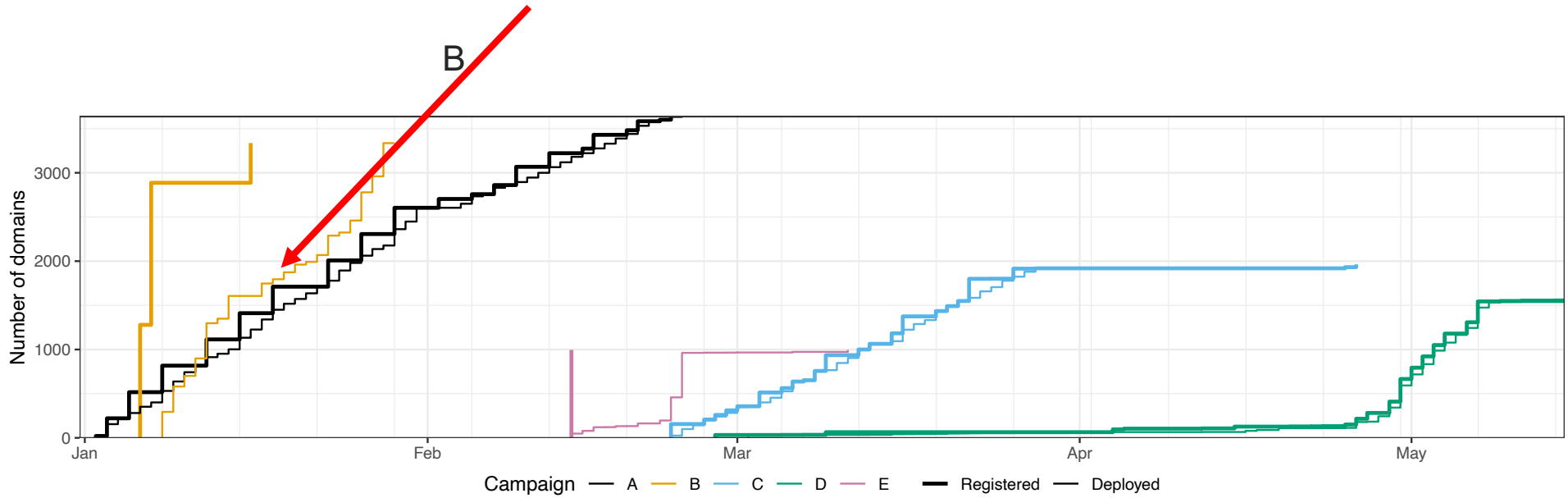


2. Deployment strategy (thin line)

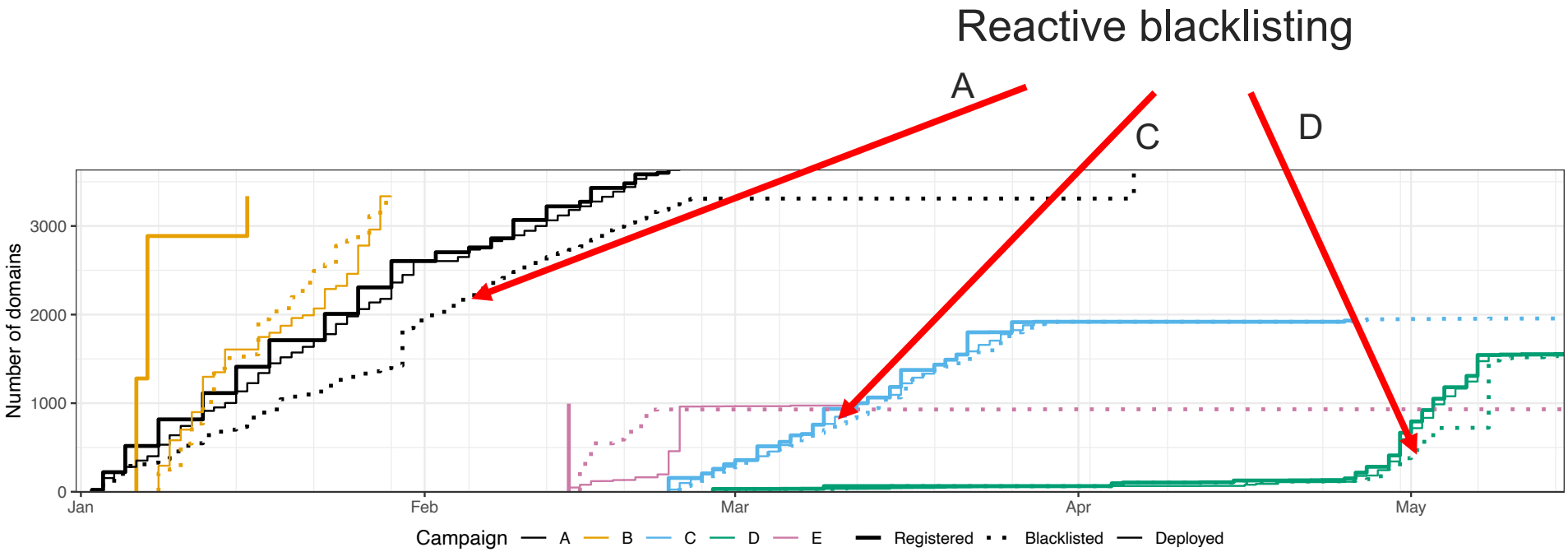


2. Deployment strategy (thin line)

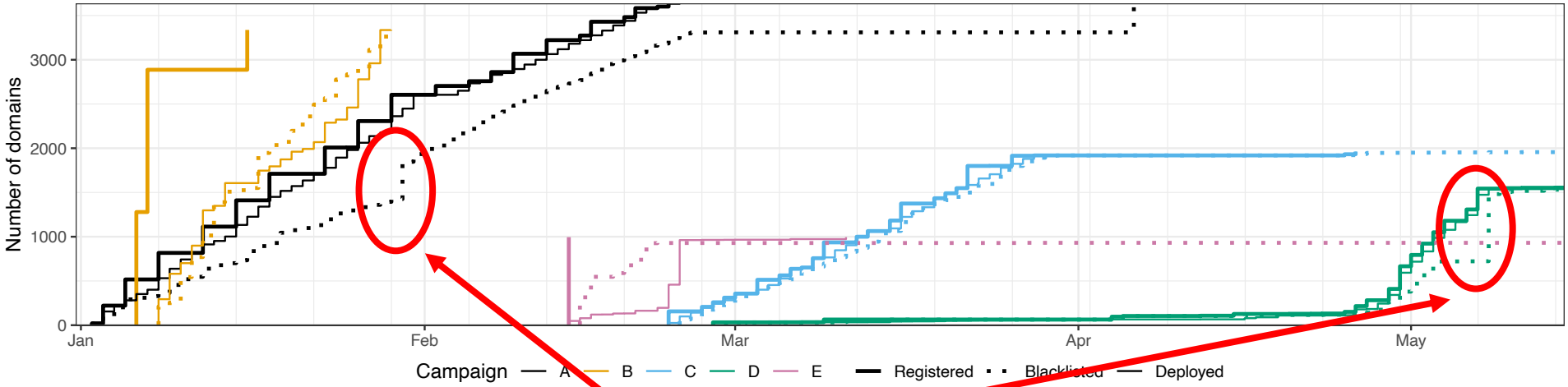
Gradual deployment,
although registered in bulk



3. Domain blacklisting (dotted line)

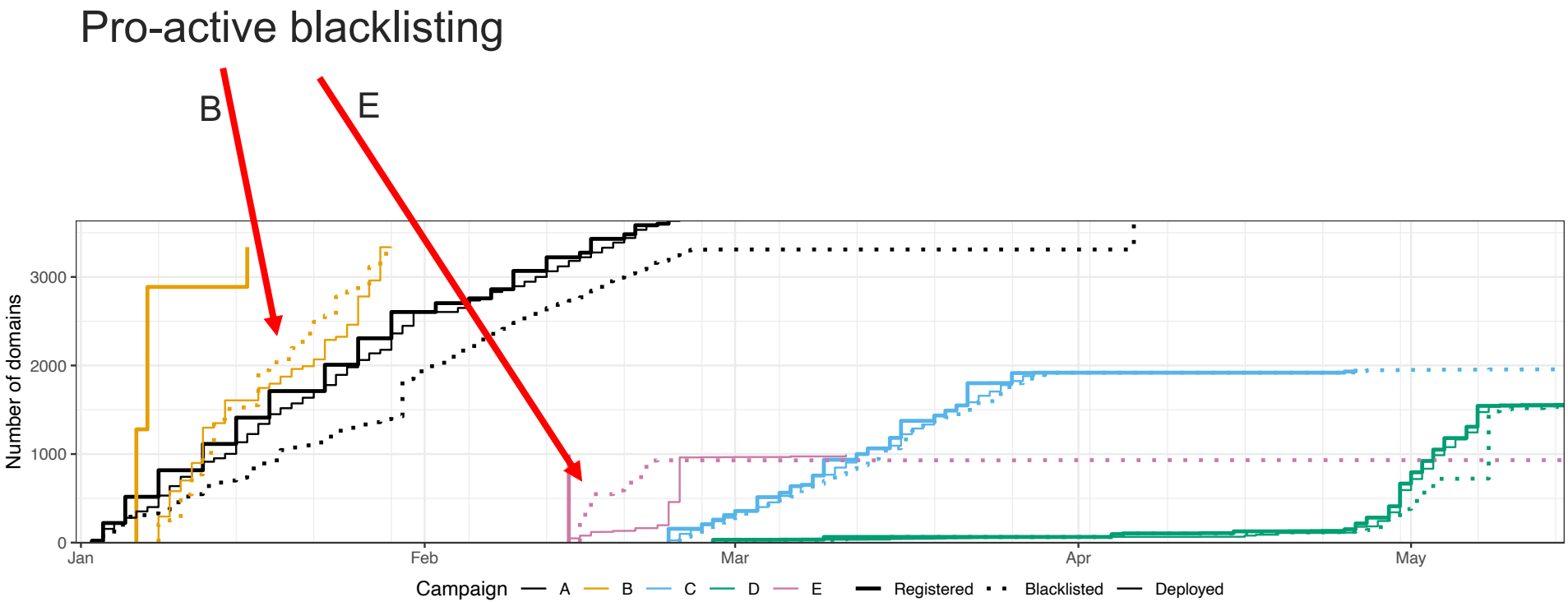


3. Domain blacklisting (dotted line)



Blacklisting in batch

3. Domain blacklisting (dotted line)



Key takeaways

Rather small set of bad actors

- Up to 20 campaigns are responsible for 80% of malicious registrations

- Top facilitators:
  - About half of the malicious registrations via 1 registrar
  - 1 public email provider are malicious with a high toxicity

Registration-time detection and prevention

- Two prediction models predict at registration-time the malicious intent

- Captures the majority of malicious domain registrations
- Incompleteness of ground truth makes analysis hard

- Interesting to see how this will further impact the security landscape

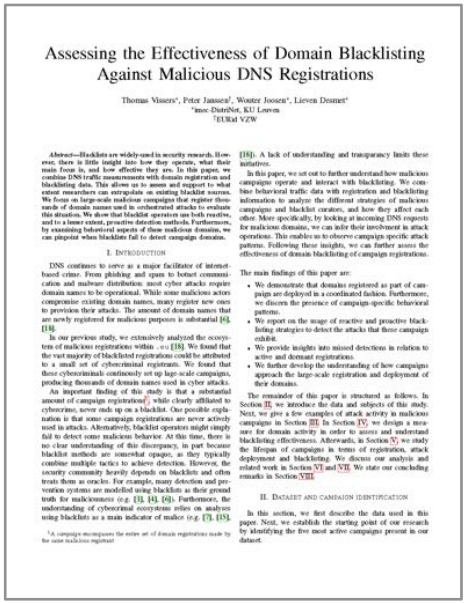
Attackers vs Defenders

- Ground truth is (somewhat) tricky
  - Bias towards spam
  - Delay in labeling
  - “Incompleteness”
- 2 different ecosystems:
  - abusive registration
  - abusive activity
- Interesting to see how it will further impact the abuse landscape

Interested in more? Some reading material…

Exploring the ecosystem of malicious domain registrations in the .eu TLD

Detection of Algorithmically Generated Domain Names used by Botnets: A Dual Arms Race.



Assessing the Effectiveness of Domain Blacklisting Against Malicious DNS Registrations

PREMADOMA: An Operational Solution for DNS Registries to Prevent Malicious Domain Registrations

https://link.eurid.eu/prediction1

https://link.eurid.eu/prediction2

https://link.eurid.eu/prediction3

https://link.eurid.eu/prediction4

https://link.eurid.eu/prediction[1-4]

Powered by EURid

DistriNet

Detecting and preventing DNS abuse in .eu

Lieven Desmet, KU Leuven – lieven.desmet@cs.kuleuven.be