PREMADOMA:

An Operational Solution for DNS Registries to Prevent Malicious Domain Registrations

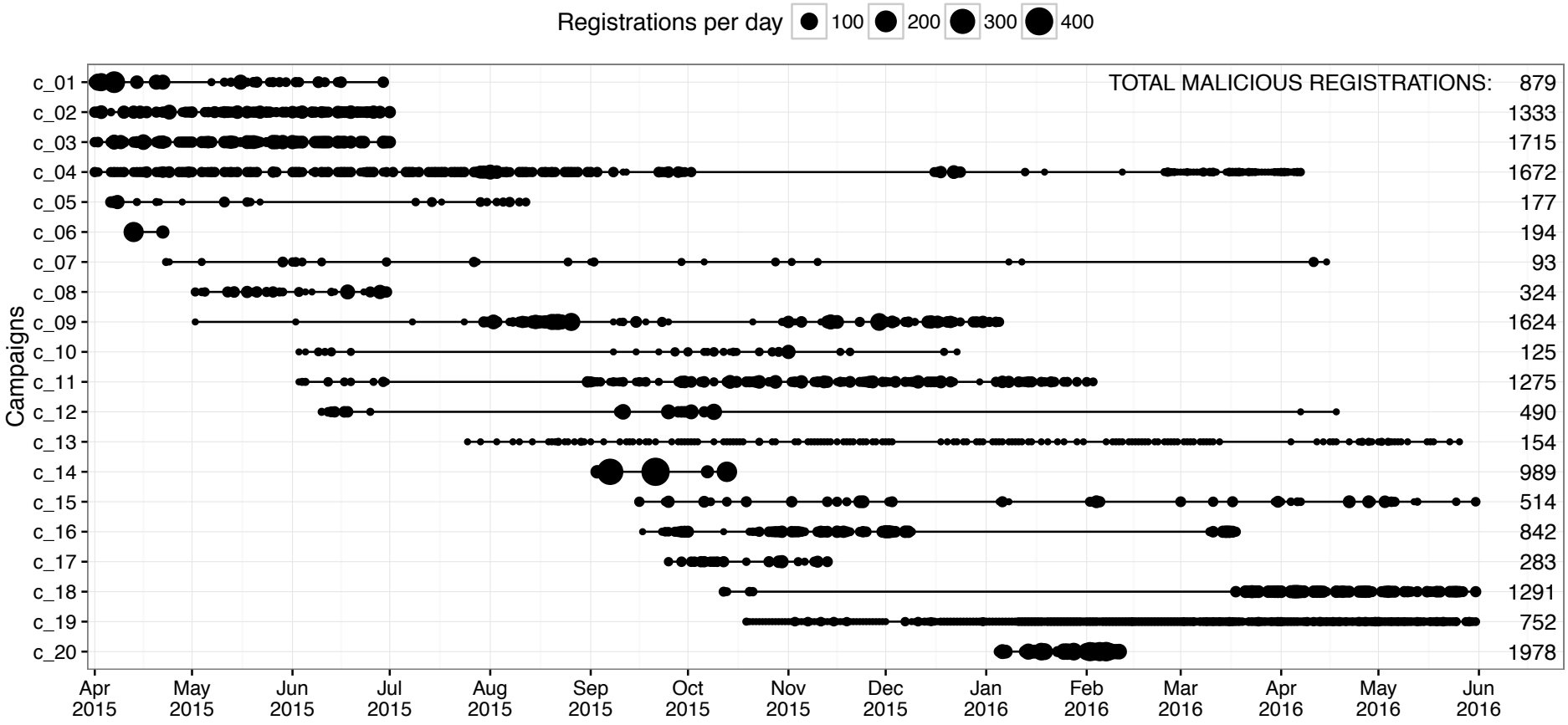
Jan Spooren, Thomas Vissers, Peter Janssen, Wouter Joosen, *Lieven Desmet*

University of Leuven – EURid (registry of .eu)

Malicious use of domain names

- Domain names are often abused by cyber criminals
  - Spam, botnet C&C infrastructure, phishing, malware, …
- To counter blacklisting, malicious actors often deploy a hit-and-run strategy
  - 60% are only active for 1 day after registration [Hao et al, 2013]

Campaigns of malicious domain name registrations

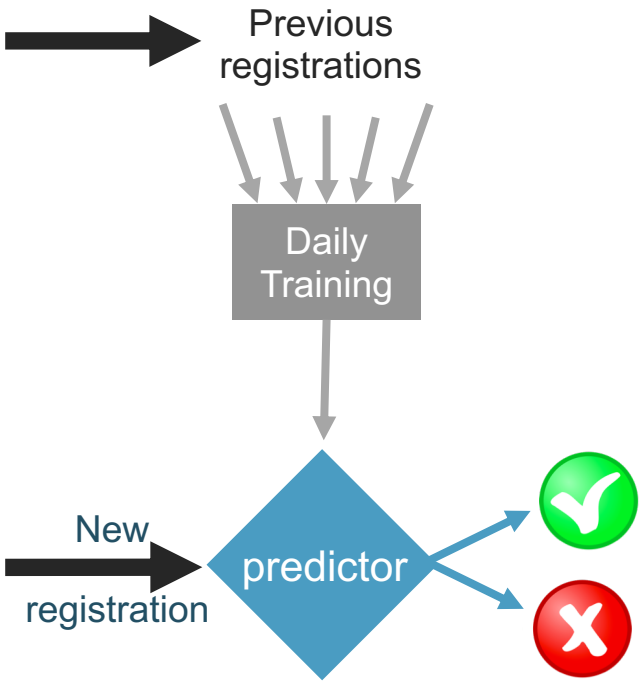


[Vissers et al, 2017] “Exploring the ecosystem of malicious domain registrations in the .eu TLD” RAID 2017

PREMADOMA: Pro-active detection and prevention

Previous registrations for which is known if they have been used maliciously

SPAMHAUS SURBL Google Safe Browsing



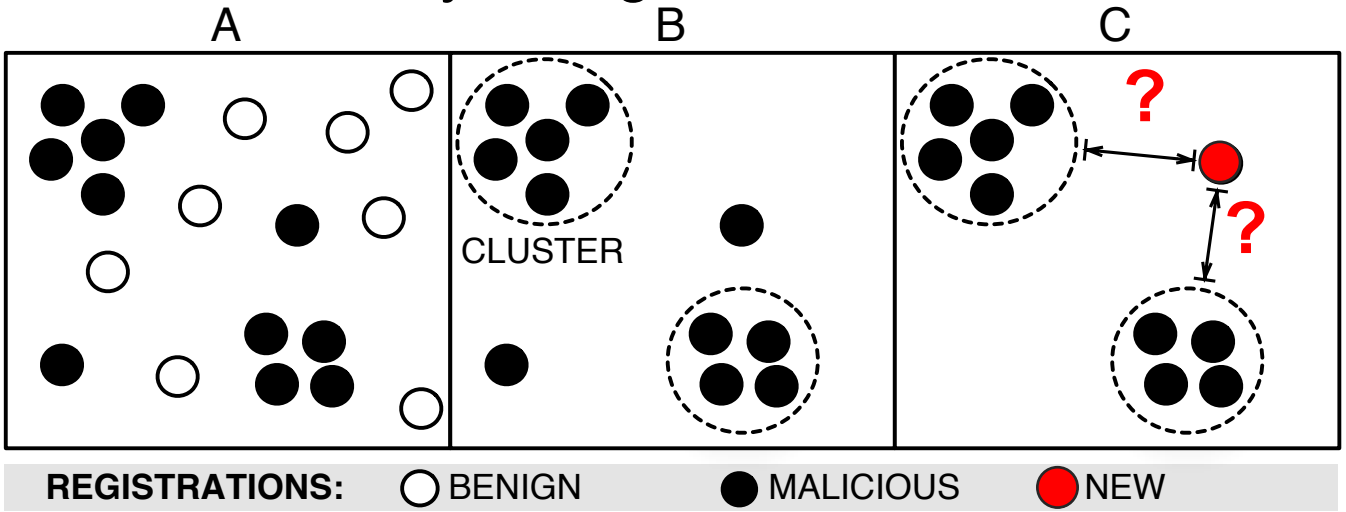
For each new registration, the system predicts if the domain will be used for malicious activity

Domains with malicious intent are rendered harmless

Insights into the predictors

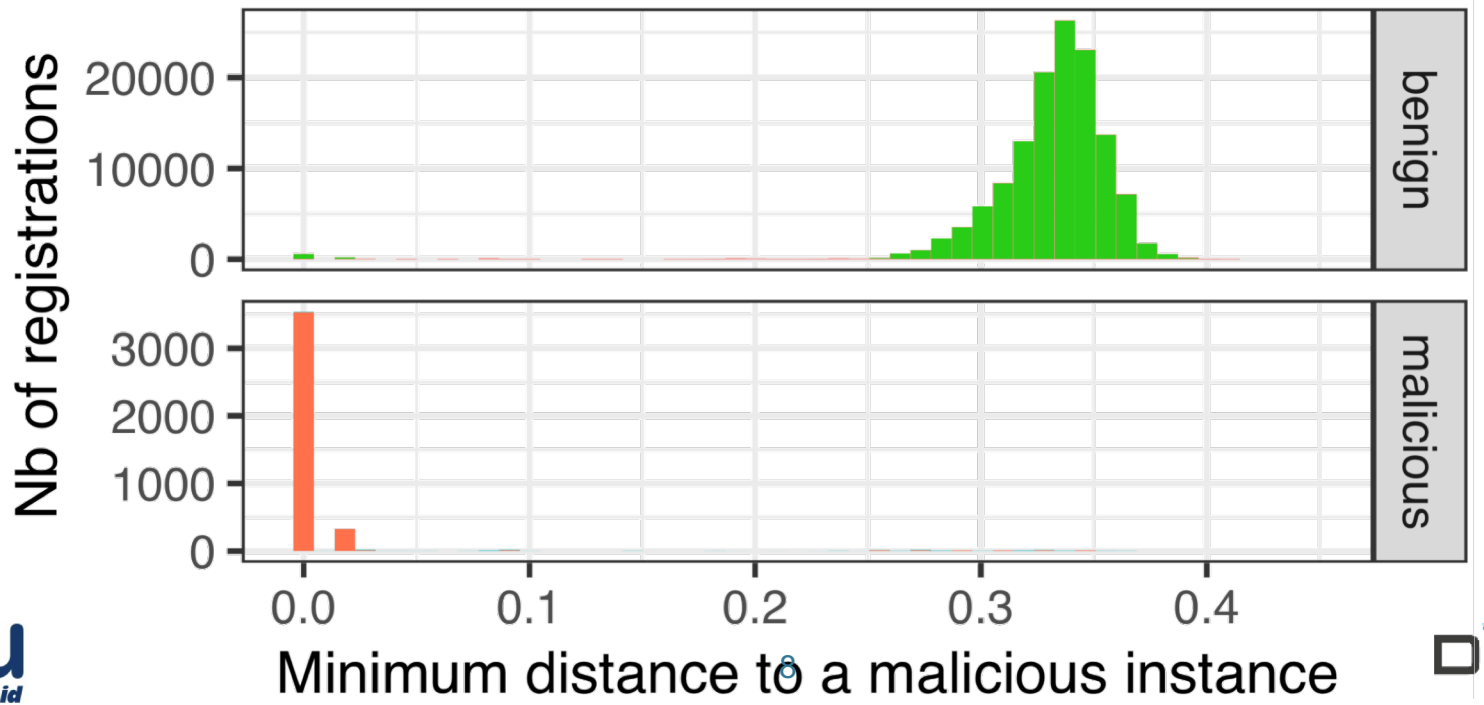
Predictor 1:
Similarity-based clustering

- 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



Predictor 2:
Reputation-based classification

- Reputation features of “facilitators”

- Technical facilitators:
  - registrar
  - name servers
- Communication facilitators:
  - email provider
  - phone number

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%

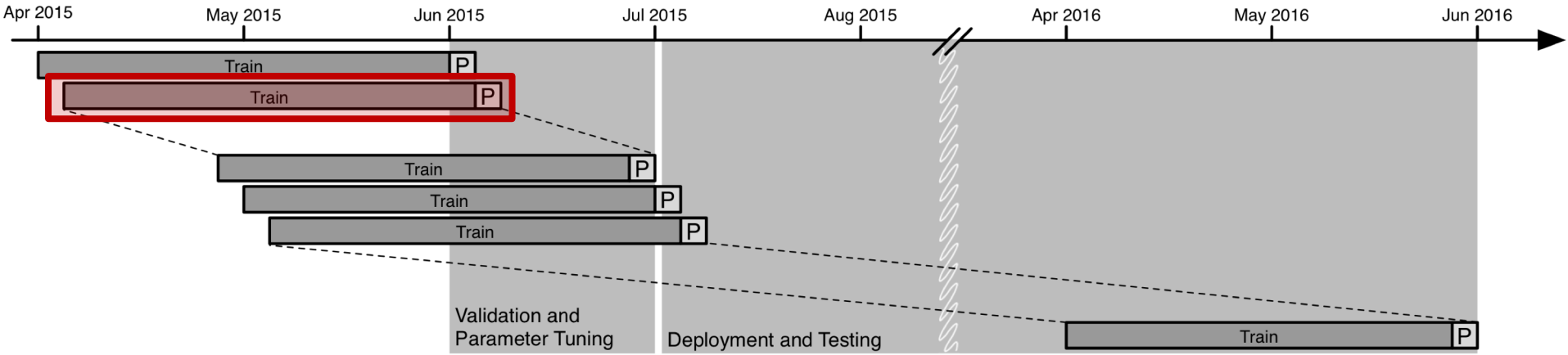
[Vissers et al, 2017] “Exploring the ecosystem of malicious domain registrations in the .eu TLD” RAID 2017

Features used for classification

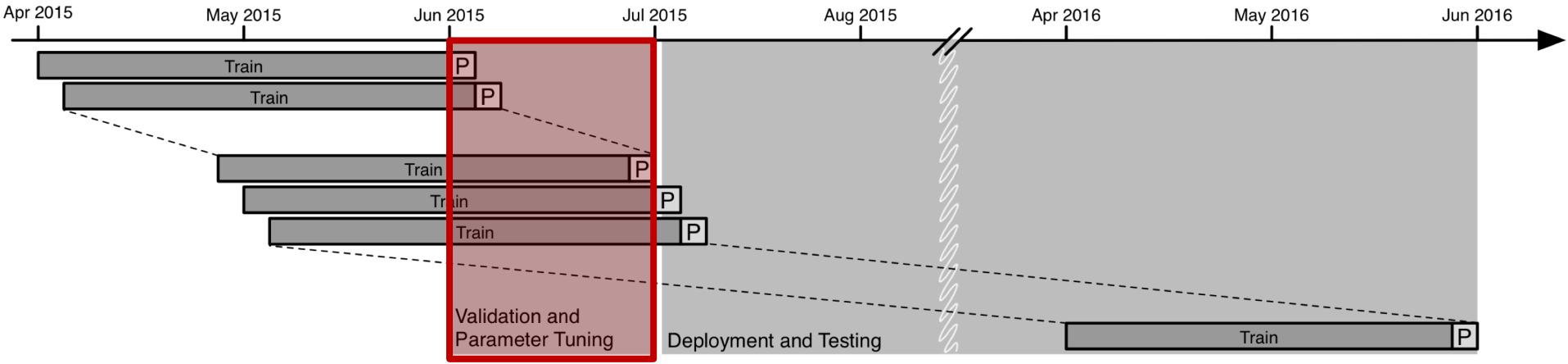
Feature	New?	Feature	New?
domain_length	[9]	domain_digits	[3]
domain_max_digit_len	✓	domain_max_digit_offset	✓
domain_max_hex_len	✓	domain_max_hex_offset	✓
email_provider	✓	hour_of_registration	[9]
registrant_country_code	✓	registrant_address_score	✓
registrar	[1,5]		
registrar_reputation_pct	✓	registrar_reputation_pct_14d	✓
registrar_reputation_pct_30d	✓	registrar_reputation_pct_60d	✓
nameservers_reputation_pct	✓	nameservers_reputation_pct_14d	✓
nameservers_reputation_pct_30d	✓	nameservers_reputation_pct_60d	✓
email_provider_reputation_pct	✓	email_provider_reputation_pct_14d	✓
email_provider_reputation_pct_30d	✓	email_provider_reputation_pct_60d	✓
phone_number_reputation_pct	✓	phone_number_reputation_pct_14d	✓
phone_number_reputation_pct_30d	✓	phone_number_reputation_pct_60d	✓

Training, validation and testing

1 month validation (June 2015)
11 month testing (July 2015 – May 2016)

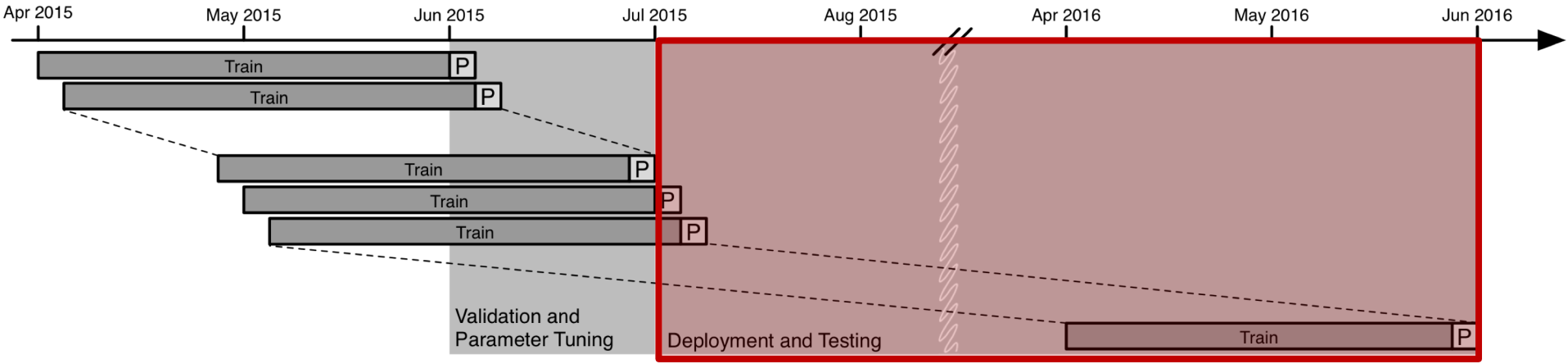


1 month validation (June 2015)
11 month testing (July 2015 – May 2016)



- Validation phase:
  - Parameter tuning
  - Ensemble model selection

1 month validation (June 2015)
11 month testing (July 2015 – May 2016)



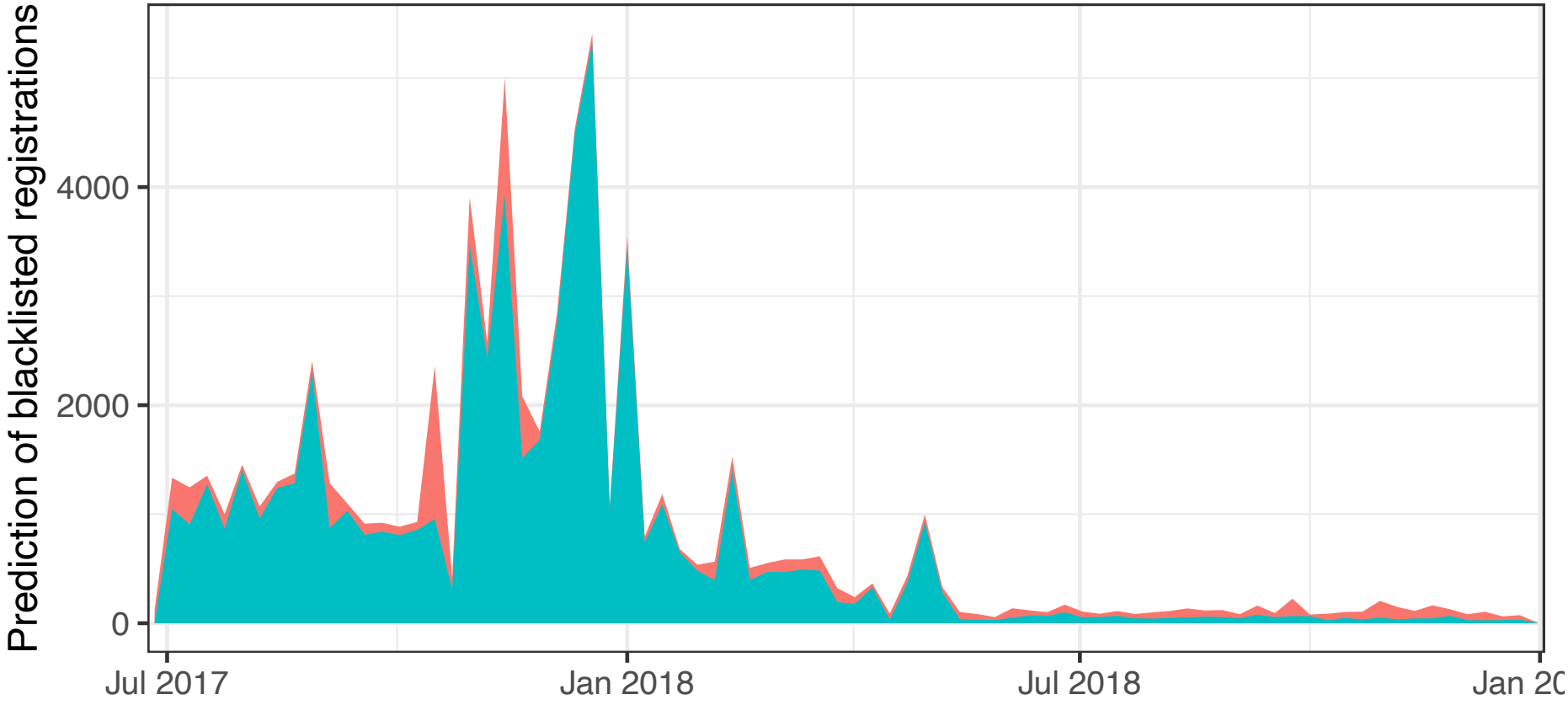
- Testing phase:
  - Evaluate ensemble model from validation phase

Testing phase: Evaluation on historical data

- Ground truth-based evaluation (11 months)
  - Recall: 66.23%
  - Precision: 84.57
  - False positive rate: 0.30%

Deployment in an operation context

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



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 [Vissers et al, 2019]

[Vissers et al, 2019] “Assessing the Effectiveness of Domain Blacklisting Against Malicious DNS Registrations” WTMC 2019

APEWS

The Abuse Prediction and Early Warning System (APEWS) is an innovative and award-winning methodology based on evaluating patterns of domain name registrations. It predicts whether a domain name may potentially be used in an abusive manner.

If the system identifies a registered domain name as potentially linked to abuse, its delegation in the .eu zone file is delayed and its status in the web-based WHOIS shows “Server Hold”.

The domain name is registered. However, any service linked to it (such as a website, email or any other service) will not function until our verification procedure is completed.

EURid manually reviews all domain names whose delegation is delayed as a result of the APEWS system. We request the domain holder to confirm his or her registration data and to submit evidence of his or her identity. The review process may lead to the delegation of the domain name in the .eu zone file or to its suspension. Should the domain name be suspended and subsequently withdrawn, it will be made available for new registration in a timely manner.

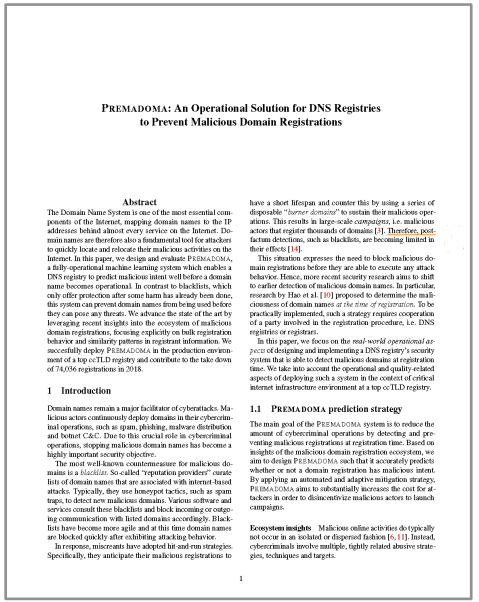
To find out more about APEWS, please c

- Detection of Algorithmically Generat
- Exploring the ecosystem of malicious domain registrations in the .eu TLD
- Assessing the Effectiveness of Domain Blacklisting Against Malicious DNS Registrations.
- An Operational Solution for DNS Registries to Prevent Malicious Domain Registrations

PREMADOMA is now fully operational
for all newly registered domain names in .eu

Challenges to go from idea to 24/7 operational system

- Inherent data set challenges
  - Strong imbalance of benign/malicious classes
  - Delays in the ground truth labelling
  - Incompleteness of the ground truth labelling
- Operational challenges
  - Trade-offs between security and performance
  - Need for predictor insights drives choice of ML
  - Strong focus on very low FPR
  - PREMADOMA itself impacts future ground truth



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 successfully 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

Key takeaways

Registration-time detection and prevention

- Two models predict at registration-time the malicious intent
- Successfully deployed at part of EURid's registrations process
- Interesting to see how this will further impact the security landscape

PREMADOMA:

An Operational Solution for DNS Registries to Prevent Malicious Domain Registrations

Jan Spooren, Thomas Vissers, Peter Janssen, Wouter Joosen, *Lieven Desmet*

University of Leuven – EURid (registry of .eu)