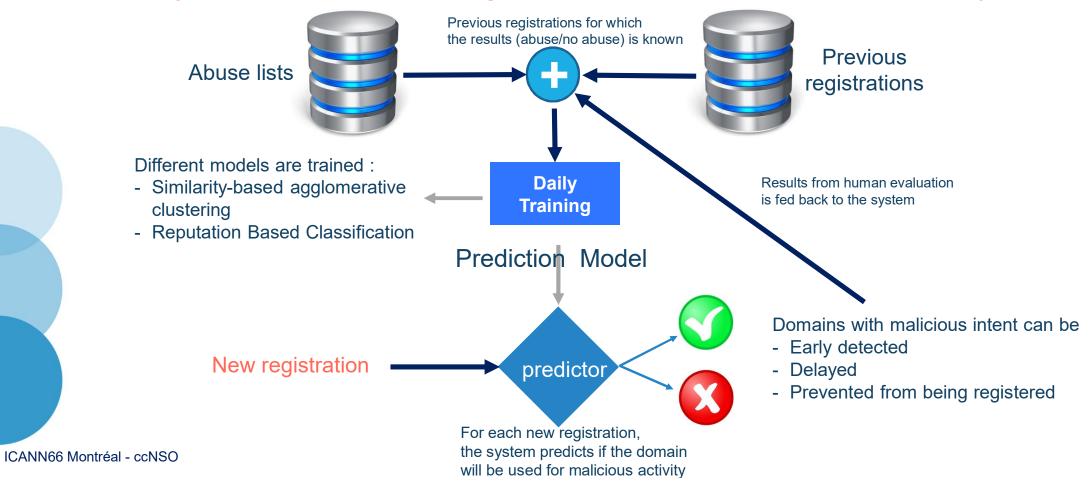


Abuse Prevention and Early Warning System (APEWS)

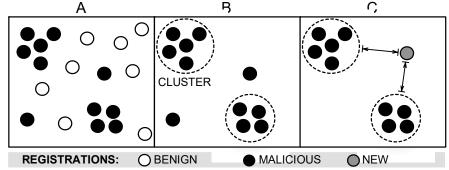
Predictive Model

Objective : Predict at time of registration whether a DN will be used abusively



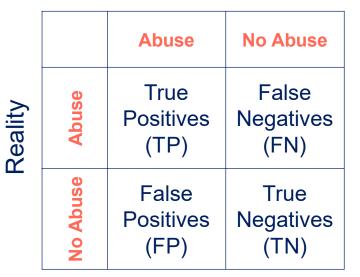
Similarity Based Clustering

- Rationale : Domains belonging to the same campaign have very similar registration data
- For all malicious registrations in the past period, the similarity with other malicious registrations is calculated and expressed as a metric
- Based on the inter-registration similarity, registrations are clustered into clusters of 'very similar' registrations, i.e. 'campaigns'
- For each new registration, the distance to the malicious clusters is calculated A B C



Results test phase

Prediction



Results

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$False \ Positive \ Rate = \frac{FP}{FP + TN}$$

How many did we find ? (of the category we were looking for)

How many were correct ? (of those we predicted as a hit)

How many were incorrectly classified as a hit ? (of those that were not abusive)

Optimization

What is most important?

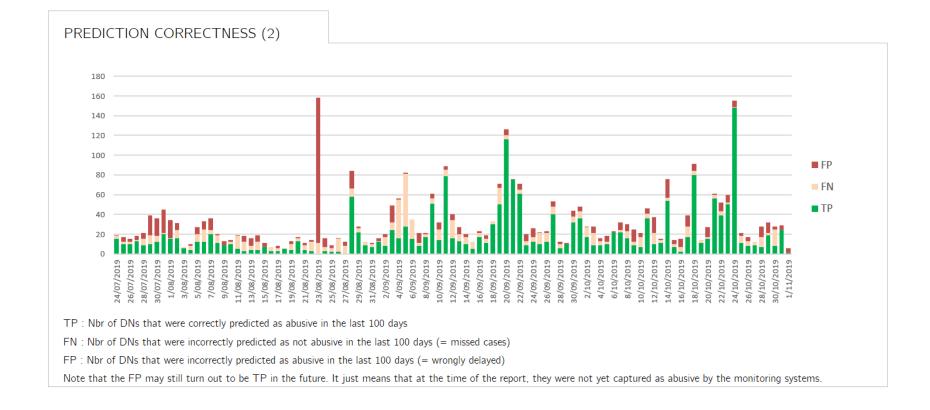
- Find all the cases (recall /) with low precision ?
- Predict correctly (precision /) and miss a lot of cases ?
- As accurate as possible ?

Results test phase

	ТР	FP->TP	FP	TN	FN	Recall	Prec.	FPR	$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$	-How accurate is our prediction ?
10/01/2019 - 02/01/2019	64	254	248	28045	60	84.13%	56.18%	0.88%		
02/06/2018 - 10/01/2019	1575	3919	1311	334821	1759	75.75%	80.73%	0.39%	$Recall = \frac{TP}{TP + FN}$	How many did we find ? (of the category we were looking for)
02/04/2018 - 20/06/2018	1996	1301	488	93023	378	89.71%	87.11%	0.52%	ТР	
28/03/2018 - 24/04/2018	643	1085	222	37504	140	92.51%	88.62%	0.59%	$Precision = \frac{TP}{TP + FP}$	How many were correct ? (of those we predicted as a hit)
10/01/2018 - 28/03/2018	4055	24	1089	80551	867	82.47%	78.93%	1.33%	FP	How many were wrong ?
			Average						$False \ Positive \ Rate = \frac{FP}{FP + TN}$	(on total benign)
		TPR: 82.32% (pct reported abuses found)								
	Precision: 81.62% (pct correct on predicted abuses)									
ICANN66 Montréal - ccNSO)			FPR (abuses			58% al benign))		

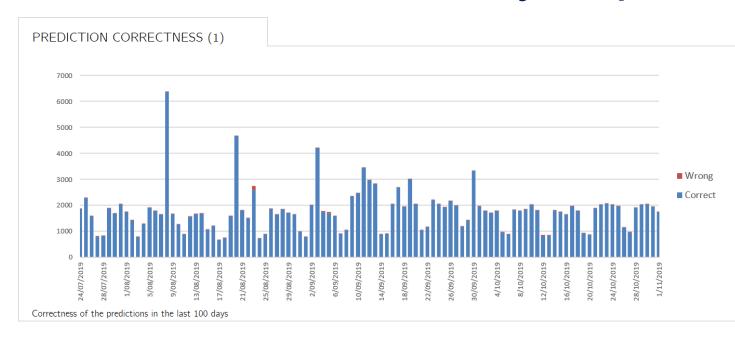
Production phase (no delay)





ICANN66 Montréal - ccNSO

The Accuracy trap



Pct of the prediction that was correct : 99.33%

But ... if we would always predict *no abuse*, accuracy would be 98.53% ! Typical for unbalanced data.



Effectiveness

1

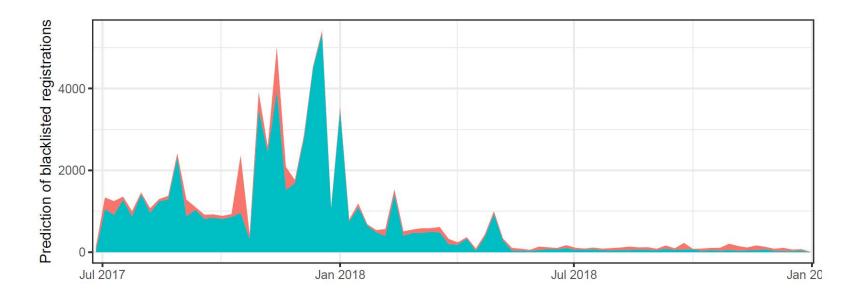


Figure 8: The weekly prediction of blacklisted registrations for the selected ensemble predictor during operations. The red area plots the total number of blacklisted registrations on that week, whereas the green area represents the predictions.

ICANN66 Montréal - ccNSO

Delayed Delegation

Predict at time of registration whether a DN will be used abusively

Status :

- Running in production without delayed delegation
- Currently 80% Recall and 80% Precision

Next Steps :

- Improve algorithms (add categorisation)
- Explore to include other abuse lists
- Start delaying

More information

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

Thomas Vissers¹, Jan Spooren¹, Pieter Agten¹, Dirk Jumpertz², Peter sen², Marc Van Wesemael², Frank Piessens¹, Wouter Joosen¹, and Lieven Desmet¹

¹ imee-DistriNet, KU Leuven, Belgium {firstname.lastname}@cs.kuleuven.be, ² EURid VZW, Belgium {firstname.lastname}@eurid.eu

Abstract. This study estensively scrutinizes 14 months of registrations data to identify hipsyncoke malidious emappings present in the aver TLD. We explore the ecceptem and modus operandi of elaboratic ephercimal emitties that recurrently register large amounts of domains for cons-hot, malicious use. Although these malicious domains are short-lived, by incorporating registration information, we exhibit that at lasous 40.06% of them can be framed in to 20 larger campaigns with varying duration also intensity. We further report on insights in the operational aspects of this business and observe, amongst other findings, that their pr are only partially automated. Finally, we apply a post-factum cla process to validate the campaign identification process and to an alidate the campaign identification process and to a em analysis of malicious registrations in a TLD zone

Keywords: malicious domain names, campaigns, DNS security

1 Introduction

The Domain Name System (DNS) is one of the key technologies that has allowed The romain scatter system (rocks) is one to the large viewinnedges that in an answer the web to expand to be current dimensions. Virtually all communication on the web requires the resolution of domain names to IP addresses. Malicions activi-ties are no exception, and attackers constantly depend upon throughing domain names to execute their absiste operations. For instance, pholing attacks, dis-tributing span machine, botter command and control (C&A) and and control (C&A). Widely-used domain blockits are currented and used to to spanific on domain and the start of the start widely one domain blockits are currented and used to to start of the sta sames³ shortly after abusive activities have been observed and reported. As a consequence, attackers changed to a hit-and-run strategy, in which malicious domain names are operational for only a very small time window after the initial registration, just for a single day in 60% of the cases [11]. Once domain names ³ We use the term malicious domain name whenever we refer to a domain name that is registered to be bound to a malicious service or activity.

https://link.eurid.eu/prediction1

ABSTRACT

ABSTRACT Makews typically use Donain Generation Agenthus (DOA) at a mechanism to ionistic that Consumal and Canad arrays. In me-dation of the term of the difficulty for an adverary to circumvent these charafters when the matchine learning model behavior the term of the term of an in-thibit of the term of the difficulty for an adverary to circumvent these charafters when the matchine learning model behavior the term of the term of the lattice send 3 does provide the term of the term of the lattice send 3 does provide the term of the term of the lattice send 3 does provide a provide the term of the lattice send 3 does the term of the term in the deep beauting approved performs constantially terms of \$47 term out \$37 term of the term of the term of the term of the \$47 term of \$37 term of the term of term of the term of term of term of the term of term of term of term of the term of term of term of the term of as on the extent DAPA, with an average classification exclusion of sectors of the sector of the sector of the sectors of the sectors show that one of the dangers of manual feature engineering is that DAGs can adapt their start segt sease on knowledge of the features used to detect them. To demonstrate this, we use the knowledge of the used feature set to design a new DAG which makes the random forces classifier powerless with a classification accuracy of 59.9%.

Jan Spooren

imec - Distrinet - KULeuven

Heverlee, Belgium jan.spooren@cs.kuleuven.be

Peter Janssen

EURid VZW. Belgium

Diegem, Belgium Peter, Janssen@eurid.eu

The deep learning classifier is also (albeit less) affected, reducing its accuracy to 85.5%. CCS CONCEPTS

- Security and privacy -> Malware and its mitigation; - Com-puting methodologies -> Neural networks, Classification and regression trees;

KEYWORDS Malware Detection, D ing

ion to make digital or hard copier of all or part of this wor

Figure 1: Bot using DGA to connect to a C&C Serve

https://link.eurid.eu/prediction2

Assessing the Effectiveness of Domain Blacklisting Against Malicious DNS Registrations

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

Absence—Biackins are widely-used in security research. Here, rever, fairs in limit, imply into how they operate, what the combine DN traffic measurements with deman regarization and biacking that. This allows to seems and support to what the data security of the securit

operations. This enables us to observe campaign specific attack patterns. Following these insights, we can further assess the effectiveness of domain blacklisting of campaign registrations.

DNS continues to serve as a major facilitator of internet-based crime. From phishing and spam to botnet communi. The main findings of this paper are: We demonstrate that domains registered as part of cam-paign are deployed in a coordinated fashion. Furthermore, we discern the presence of campaign-specific behavioral we discert the presence of campaign-specific behavioral patterns. • We report on the usage of reactive and proactive black-listing strategies to dotect the attacks that these campaign exhibit. • We provide insights into missed detections in relation to

active and dormant registrations. • We further develop the understanding of how campaigns

approach the large-scale registration and deployment of their domains.

An important finding of this study is that a substantial mount of compare predistrating¹ while learly affiliated in comparison of this paper is interned as ignore. In success of the study is the study of the study of the study of the study of the comparison is section of the study of the study of the study of the study is the study of the study of the study of the study and in study. A study of the stud fail to detect some makicene behavior. At this time, there is no clear understanding of this discrepancy, in part because blacklit methods are somewhat opsages, as they typically combine multiple tacks to a knive detection. However, the security community heavily depends on blacklitus and other setts them as oracles. For example, many detection and pro-vention systems (e.g. [1], [4], [6]. Furthermore, et understanding of cybercrimal consystem relies on analysis with glacklitus as main inductor or function. (e.g. [7], [5].

using blacktists as a main moncator of mainee (e.g. [/], [15], ¹A compary mempiases the entire set of domain registrations made by the same mainious registrant

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

Thomas Vissers Jan Spooren thomas.vissers@cs.kuleuven.be imec - DistriNet, KU Leuven jan.spooren@cs.kuleuven.be imec - DistriNet, KU Leuven Wouter Joosen

wouter.joosen@cs.kuleuven.be imec - DistriNet, KU Leuven

Belgium

eter.janssen@eurid.eu EURid VZW Belgium Belgium Lieven Desmet lieven.desmet@cs.kuleuven.be imec - DistriNet, KU Leuven Belgium

1 INTRODUCTION

Peter Janssen

9

ABSTRACT

Prevent Malicious Domain Registrations. In 2019 Annual Computer Security Applications Conference (ACSAC '19), December 9–13, 2019, San Juan, PR, USA, ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3359789.3359836 ADS 11KAC1 DNS is one of the most essential components of the Internet, map-ping domain names to the IP addresses behind almost every online service. 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 Parsanosa, a solution for DNS registries to predict malicious intert well be-Domain name remain a major facilitator of cyberattacks, Malicious actors continuously deploy domains in their cybercriminal oper-ations, 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 whether fore a domain name becomes operational. In contrast to blacklists which only offer protection after some harm has already been which only offer protection after some harm has already been door, this system can prevent domain mane from being used be-fore they can pose any threats. We alwance the state of the art by levergains prevent inhights in the ecosystem of malicious domain registrations, focusing explicitly on facilitators employed for bulk registrations on dimatrixip patterns in registration and damainty pattern is registration and damainty pattern is registration and damainty pattern is registrations on dimatosis and the strategistration and manifestration evaluation endines of the strategistrate and advance with the registrate endines. The strategistrate endines are strategistrate the strategistrate endines are strategistrate endines are strategistrate. The most well-known countermeasure for malicious domains is a blacklist. So-called *reputation providers* curate lists of domain manes that are associated with internet-based attacks. Various soft-ware and services consult these blacklists and block incoming or the second services consult these blacklists and block incoming or cu ccTLD registry to detect and prevent malicious registrations, and have contributed to the take down of 58,966 registrations in 2018.

CCS CONCEPTS

ware and services cound these blackins and block incoming or organizing communications with lated domina secondingly. Black-last have become more agile and novadays domin names are blacked quickly there chalking attacking blacking. They counter the short sackle lifespace of their domina names by registering blackers of disposale.²⁴ *Prove domasile* '0 statistican their mildious operations, resulting in large-scale registration cou-pleagings. [10, 22] *Provements* and the statistican their mildious operations, resoluting in large-scale registration cou-pleagings. [10, 22] *Provements* and the statistican black based on the statistican black of the statistican black based on the statistican black of the statistican black based on the statistican black of the statistican black theore, more recent security research aims to shall be statistican black of malaxies. The statistican black of the statistican black theorem is a statistican black of the statistican black of the statistican black of the statistican black of the statistican black theorem is a statistican black of the statistican black of the statistican black theorem is a statistican black of the statistican black of th Security and privacy → Intrusion/anomaly detection malware mitigation: • Networks → Naming and addre Information systems → World Wide Web. KEYWORDS

Domain Name Registration, Early Detection, Malicious Domains

ACM Reference Format: Jan Spooren, Thomas Vissers, Peter Janssen, Wouter Joosen, and Lieven Desmet. 2019. PREMADOMA: An Operational Solution for DNS Registries to

Trence, more recent security resented and to share to construct acce-tion of malicious domain names. In a ground-breaking paper, Hao et al. [9] proposed to predict the maliciousness of domain names at the time of registration, using a set of 22 manually crafted fea-tures derived from data available at registration time and a Cowwe Polytope Machine (CPM) classifier. Protytope Machane (CPM) classifier. Subsequent work by Vissers et al. [22], showed that in the .eu top level domain (TLD), approximately 80% of malicious domain registration compaigns are registered by maximum 20 actors, indi-vidually using very different modi operandi. This could also explain Permission to make digital or hard copies of all or part of this work for p classroom use is granted without fee movided that coving the second secon classroom use is granted without fee provided that cogs for profit or commercial abstrategies and that cogies bear if on the first page. Copyrights for compensations of this was undrecity must be horsened. Abstratesting with coefficient and/or a fee. Request permissions from permissions/plus (ACMC '15, December 9 - 17, 2007, San Jawa, PA, USA vidually using very different modi operandi. This could also explain why the detection accuracy reported by Hao et al. [0] for the .net TLD (613) differs significantly from the .com TLD (703) at the same PRE: Different TLD will likely have different sets of malicious ac-tors, with different operational characteristics, yielding different detection results. Moreover, the actual operational deployment of such a detection system in a real and lice environment detastically

I. INTRODUCTION

cation and malware distribution; most cyber attacks require

https://link.eurid.eu/prediction3

https://link.eurid.eu/prediction4

ICANN66 Montréal - ccNSO

cation is available at Springer via http://dx.doi.org/10.1007/978-3-319-66332-6.21

 aphuTreating com
 id7ahrsids.com
 ost846is4t.com
 sd16e9w13s.com ration Algorithms, Machine Learn

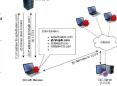
Detection of Algorithmically Generated Domain Names used by

Botnets: A Dual Arms Race.

Davy Preuveneers

imec - Distrinet - KULeuven

Heverlee, Belgium davy.preuveneers@cs.kuleuven.be



Lieven Desmet

imec - Distrinet - KU Leuven Heverlee, Belgium lieven.desmet@cs.kuleuven.be

ACM Meterence Format: Jan Spooren, Davy Preuveneers, Lieven Desmet, Peter Jansson, and Woute Joseen. 2019. Detection of Algorithmically Generated Domain Names use by Botnets: A Dual Arms Bace. In The 34th ACM/SIGAPP Symposium or

Applied Computing (SAC '19), April 8–12, 2019, Linamol, Cyprus. ACM, New Tork, NT, USA, Article 4, 8 pages https://doi.org/10.1145/3297280.5297467

1 INTRODUCTION The Internet non-nets billions of devices, ranging from servers and personal computers to tablets, mobile phones, household appli-ances, and many more. Malicious actors are constantly scanning the internet for vulnerable devices which could be compromised, or are

Wouter Joosen

imec - Distrinet - KU Leuven

Heverlee, Belgium wouter.joosen@cs.kuleuven.be

ACM Reference Format:

1 INTRODUCTION





Thanks

marc.vanwesemael@eurid.eu