

Dr. Marco Balduzzi
@embyte, madlab.it

Using Machine-Learning to Investigate Web Campaigns at Large

2nd Nov 18

JD-HITB PEK

Beijing, China

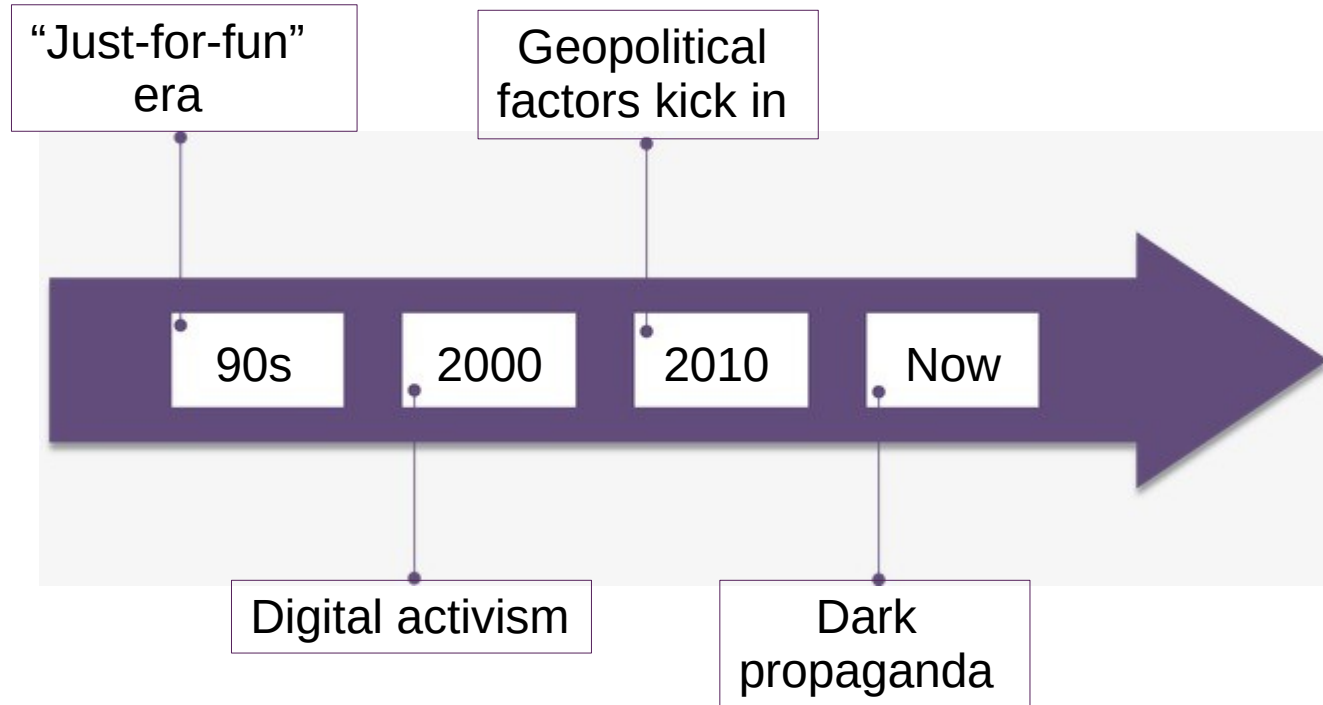
Who am I?

- HITB aficionado
- Computer security geek (since 2002)
- Love research, Ph.D. (in 2011)
- “Community centric”
- Work for Trend Micro



Web Defacement
=
Website Compromise
+
Homepage Hijacking

Evolution



“Just-for-fun” Era

```
=====
HACKED BY LIBERO
=====
Login  : Libero    Ok !
Password : ***** Ok !
=====
Conection Ok !
=====
Fuck CFK !!!
=====
WE ARE ARGENTINA HACK TEAM
=====
Twitter : @LiberoamericaMu
=====
UN SALUDO PARA EL MEJOR DE CHILE FAILROOT :)!
```

```
| owned by ssh-2 | CHILE |
```

```
uname -a;id
```

```
Linux obelix 2.6.15-1-686 #2
Mon Mar 6 15:27:08 UTC 2006
i686 GNU/Linux
```

```
uid=0(root) gid=33(www-data)
groups=33(www-data)
```

Digital Activism

We hack because we are fighting against bad information, all of us ain't terrorists, or stupid boyz who don't know what they're doing. We're just some guys, maybe we've grown fastly, but we're curious. Is it a crime to be interested in learning? Each of us has an own life, which can be good or bad. We don't like to live in misinformation or ignorance. We want a better world; a world where there ain't inequalities, differences, injustices: a new world! We don't want that the world will be guided by multinational companies, we don't want a world like this. Our world has to be of everyone who passes his/her own believings, religions, races, politic ideas, social conditions and everything which makes differences between people. We want this, and we're fighting for this. We are against each kind of war, against each kind of oppression, against each form of abuse of power: we're fighting and we'll fight in each place where the individual freedom is threatened. We will fight under each flag besides people we don't even know but whose idea of freedom is the same of our one. We haven't open access to the means of communication which are false, they say what they want and in the way they prefer, this is the only way to spread our ideas. Northern Ireland, Kosovo, Iran, Iraq, Afghanistan, Chile, Peru, Myanmar, Brazil, Philippines, Palestine, Pakistan, Kurdistan, Turkey, Algeria, South Africa, China, Taiwan, Kenya are places where freedom is a dream, open your eyes, open your mind, wake up, and all together we'll turn freedom from dream to reality. You can agree with us or not. If you agree with us let diffuse this message to make our ideas become everybody's ideas

Geopolitical Factors

msnbc.com news services
updated [12/10/2006 11:47:52 PM ET](#)

Print | Font:  A  +  -

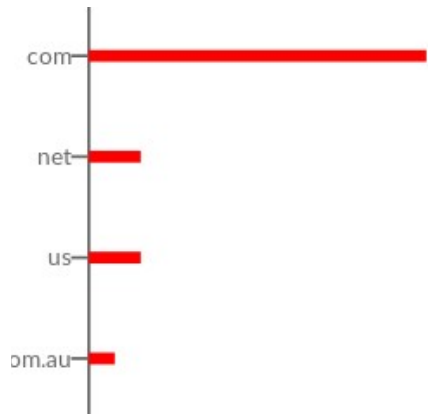
[SANTIAGO, Chile](#) — Gen. Augusto Pinochet, who terrorized his opponents for 17 years after taking power in a bloody coup, **died Sunday**, putting an end to a decade of intensifying efforts to bring him to trial for human rights abuses blamed on his regime. He was 91.

Supporters saw Pinochet as a Cold War hero for overthrowing democratically elected President Salvador Allende at a time when the U.S. was working to destabilize his Marxist government and keep Chile from exporting communism in Latin America.

Hacked by @br  ix15
R.I.P Pinochet

Coordinated campaigns: From one to many targets

- Death statement is 2nd of May 2011
- Campaign is 6th-12th of May
- Targets:





The Israeli-Palestinian conflict



HACKED BY **Anonymous Ghost Gaza**
|Free Palestine|

We are Muslims, we the people will not accept that Palestine be occupied for Israel will not eliminate Israel electronically on screens today and tomorrow with machine guns in the next Trqbona our attacks we Army Long live Palestine, Hamas Qassam
Op #Electroin
Cyber war has been declared on Zionist cyber space and you will see exactly what we are capable of.
Our Target is your Governemnt's Websites ... We Will Take You Off From The Internet !
Your Credit Cards , Your Bank's Accounts , Your Servers , Your Facebook , Twitter ... Are In Danger !
We Never Forget what You Do against the Humanity ... In GaZa Millions Of People Were Dead ... Palastinian's Muslims, Innocent Childrens, Womens Are Killed In The Force's Attacks...
Zionist Forces Destroyed Palestnian's Families , Homes ,schools, hospitals,mosks ... But No One Cares ! its the time to stop killing innocent peoples , and stop killing palastinians childs coz all da world know that zionist are the real terrorist.
Government and Peopel of the thing named "Israel" will be a goal of Palestnian resistance rockets. To save your self your Family... You MUST forced your government to release every prisoners on hunger strike... and to get out from Palestine Coz:
IN OUR MAPE IT WILL STAY PALASTINE FOR EVER , TILL WE DIE ..



HACKED BY **Anonymous Ghost Gaza**
|Free Palestine|

#Save Gaza
|Death To Zionist|

We are Muslims, we the people will not accept that Palestine be occupied for Israel will not eliminate Israel electronically on screens today and tomorrow with machine guns in the next Trqbona our attacks we Army Long live Palestine, Hamas Qassam
Op #Electroin , #OP IL , #OP FR , #OP US , #OP UK
Cyber war has been declared on Zionist cyber space and you will see exactly what we are capable of.
Our Target is your Governemnt's Websites ... We Will Take You Off From The Internet !
Your Credit Cards , Your Bank's Accounts , Your Servers , Your Facebook , Twitter ... Are In Danger !
We Never Forget what You Do against the Humanity ... In GaZa Millions Of People Were Dead ... Palastinian's Muslims, Innocent Childrens, Womens Are Killed In The Force's Attacks...
Zionist Forces Destroyed Palestnian's Families , Homes ,schools, hospitals,mosks ... But No One Cares ! its the time to stop killing innocent peoples , and stop killing palastinians childs coz all da world know that zionist are the real terrorist.
Government and Peopel of the thing named "Israel" will be a goal of Palestnian resistance rockets. To save your self your Family... You MUST forced your government to release every prisoners on hunger strike... and to get out from Palestine Coz:
IN OUR MAPE IT WILL STAY PALASTINE FOR EVER , TILL WE DIE ..

Data Collection

Public Repositories



Search :

Notifier ▾

Search




Notifier

Mass Notifier

Top100 User

TOP30 User

- No1 : 越南邻国宰相 [34333]
- No2 : Jack Riderr [19199]
- No3 : Mr.Kro0oz.305 [10672]
- No4 : AlfabetoVirtual [8907]
- No5 : 宇少 [6559]
- No6 : 大圣 [6486]
- No7 : 域血 [5408]
- No8 : That is me [5282]
- No9 : 憔悴 [5013]
- No10 : 颓废 [4791]
- No11 : 童梦玉 [4685]

Date	Notifier	Domain	Country	PR	View
2016-02-12	To	http://bananabags.co.za/	Country	0	View
2016-02-12	To	http://bsvrenault.co.za/	Country	0	View
2016-02-12	To	http://bsgtriathlon.org/	Country	0	View
2016-02-12	To	http://broadbanddemand.co.za/	Country	0	View
2016-02-12	To	http://bettacars.co.za/	Country	0	View
2016-02-12	To	http://blog.camelotkennels.co.za/	Country	0	View
2016-02-12	Al	Virtual		0	View
2016-02-12	Al	Virtual	Country	0	View
2016-02-12	Al	Virtual		0	View
2016-02-12	ch	14		0	View

Others



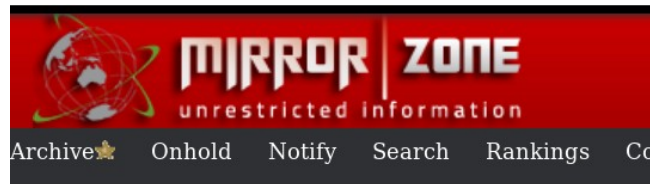
Hack Mirror
@hackmirror

The place where we make you a legend

The Universe

hackmirror.com

Iscrizione a marzo 2012



Hi There!

I am so happy to announce the new Mirror-zone.org unrestricted information site, with few nice changes.

We are now at testing this script, this means the script is not stable, and you may get some errors, if you face any issues please contact us and we will be glade to work with you to in order to resolve such issues. Also, this not mean that the Gold release will have the some futures with the bugs fixes, and we will continue to add new futures to the site.

Best Regards,
Administration

Hello

mirror-zone.org followers,
Did your defaced site mirror goes to onhold ?
Because your notify name are missing on your deface page
if you want your mirror goes to archive directly
than follow the instructions

Welcome to
MyDeface.com || Let's Show off....

Latest Verified Mirrors

Name	Domain
Defaced	http://upsanddowns-hordaland.no/web
Defaced	http://bradfordcurtainsandblinds.co.uk/w
Defaced	http://leatherheaddistrictscouts.org.uk/w
Defaced	http://leatherheaddistrictscouts.org.uk/w
Defaced	http://pickwickassociation.org.uk/web
Defaced	http://rhodapartridge.co.uk/web

THE site :)



ZONE-H In Numbers

News: **4.738**

Admins: **4**

Registered Users: **142.949**

Early Warning subscriptions: **7312**

Digital Attacks: **13.498.756**

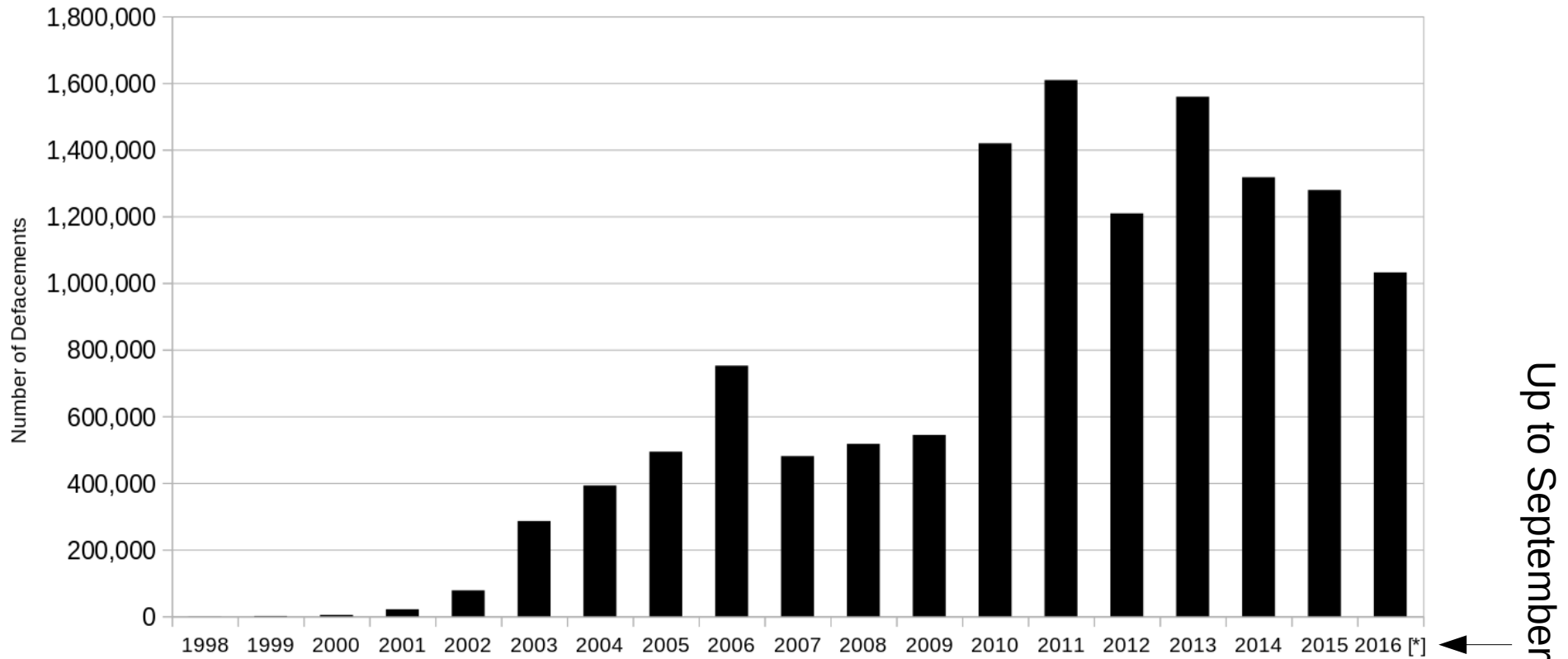
Attacks On Hold: **333.287**

Online Users: **268**

Collected Data

Source Name	Website URL	Acquired Records
Zone-H	www.zone-h.org	12,303,240
Hack-CN	www.hack-cn.com	386,705
Mirror Zone	www.mirror-zone.org	195,398
Hack Mirror	www.hack-mirror.com	68,980
MyDeface	www.mydeface.com	37,843
	TOTAL	12,992,166

Timeline Evolution



Data Format

Metadata →

```
Mirror saved on: 2018-05-24 01:28:32
Notified by: Mr [redacted]69   Domain: http://www.redraven.co.za/files.html   IP address: 75 [redacted]54 🇺🇸
System: Linux                 Web server: Apache                 Notifier stats
This is a CACHE (mirror) page of the site when it was saved by our robot on 2018-05-24 01:28:32
```

Raw content →

```
[ Xai Syndicate - ( )

~~ Mr.DreamX196 - D4RKNE55 - 0xd3vs - ML7C - ./51N1CH1 - ./R015 - Mr.AchanX48 -
Mr.BucketHead - Angel Dot Id - Ups1337 ~~
[< Littlebear69 - I3mot_n3t - Laser69 - Gend3ruw0 - SPEDY-03 - magelang9etar -
civiliant - KATENBAD - ./RosesDie - indonesia6etar! - Mr.Lucifers - Con7ext - LCR999X
- Mr.7z>]

root@x69 $ Hacked by Mr [redacted] hun with style 69 <3 --[ ]~
```

Data Trustworthiness

Type	Attribute	Example	Trustworthiness
Metadata	URL	<code>http://target.gov</code>	High
	Timestamp	<code>2010-01-02 15:00</code>	Medium
	Nickname	Neo Hacker	Medium-Low
	Webserver; Reason; Hack Mode	Nginx; Political; SQLi	Low
Raw content	Main page	HTML or TXT file	High
	Embedded resources	Various format	High
	External resources	Various format	Medium-High

General Trends

Topics Over The Years

Security Problems

Year	Most relevant topics
------	----------------------

1998	question, student, security , number, place
------	--

1999	cowboy, <i>team</i> , security , think
------	---

2000	baby, tabloid, people, provided
------	---------------------------------

2001	lord, prime, provided, saved, better
------	--------------------------------------

2002	worry, sind, lame , care, encryption
------	--

2003	backup , gift, <i>team</i> , came, take
------	--

2004	best, <i>group</i> , micro, look, total
------	---

2005	normal, pope , time, familia, contact
------	--

2006	terror , saved, intruder, energy, user
------	---

2007	badger, since, high, turk , turkey
------	--

2008	<i>crew</i> , speech, warning, saved, <i>team</i>
------	---

2009	knowledge, acker, <i>team</i> , album, country
------	---

2010	posted, member, protocol, kernel, security
------	--

2011	contact, security, village, holding, highlander
------	---

2012	saved, contact, <i>team</i> , underground
------	---

2013	<i>team</i> , forgive, security
------	---------------------------------

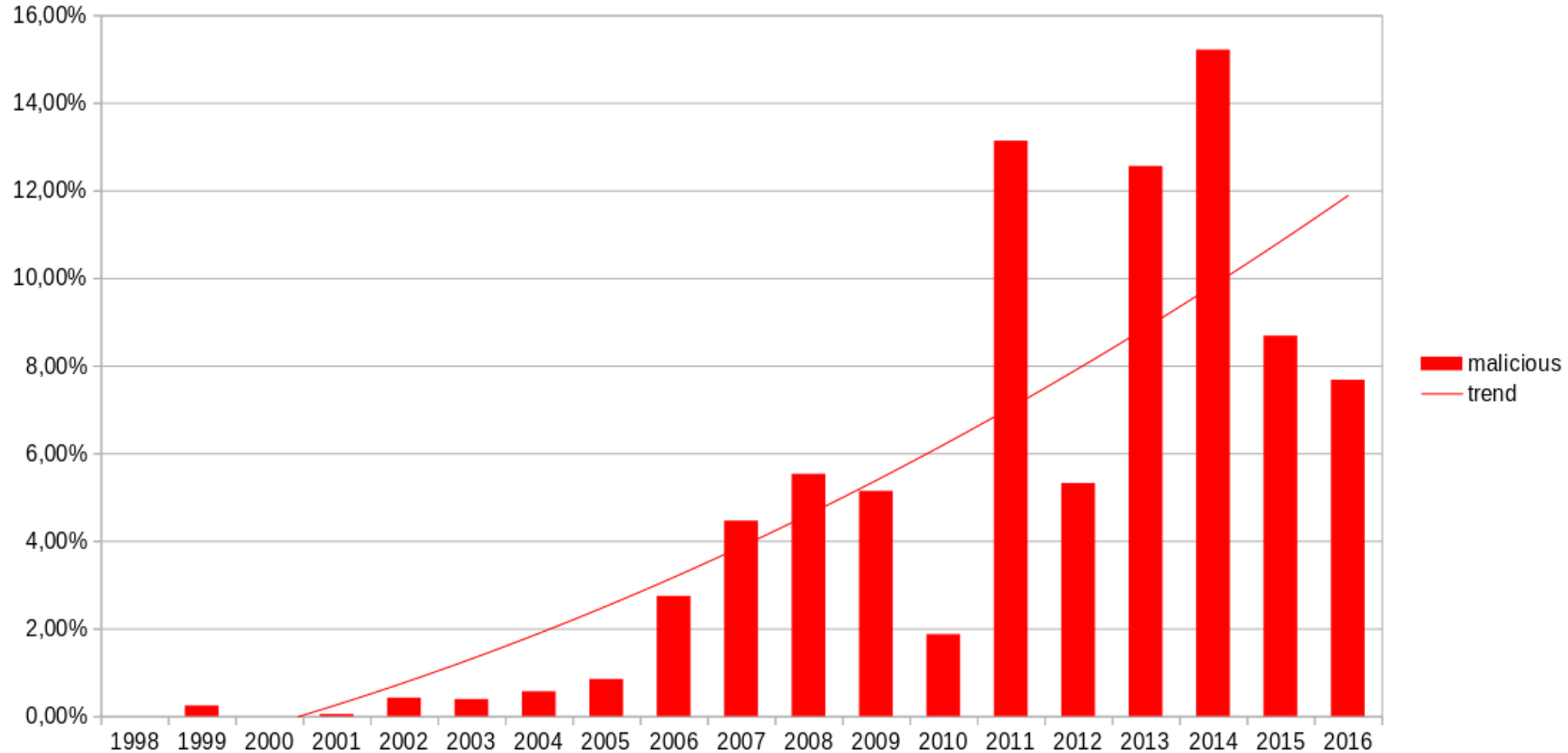
2014	eagle, crew , electronic
------	---------------------------------

2015	clash, king, terrorism , visit, alligator
------	--

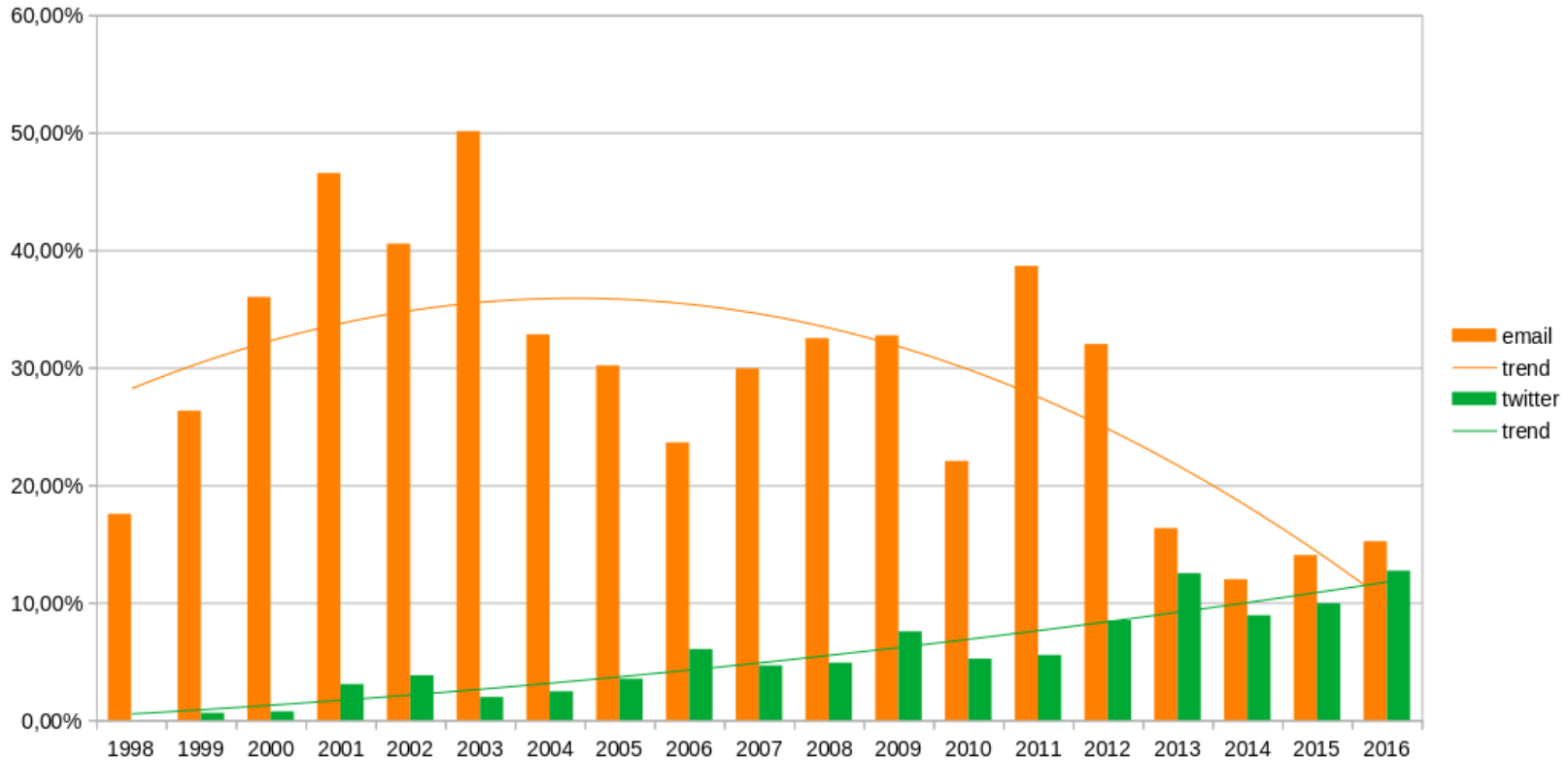
2016	marocain , turk , steel, anonymous, <i>team</i>
------	---

Real World Events

Adoption of Malicious Content



Adoption of email & Twitter handlers



Detection Engineering

Key Observations

#OpFrance
Owned !



Your System Owned By

United Islamic Cyber Force
U Remember Us ? We Are Back Again !
#We Are Stand With Muslims in Europe#
I Challenge Your Governement To Stop Us !

We are : # Dr. [redacted] erHack AI

#OpFrance
France Coupes Website Hacked



Your System Owned By

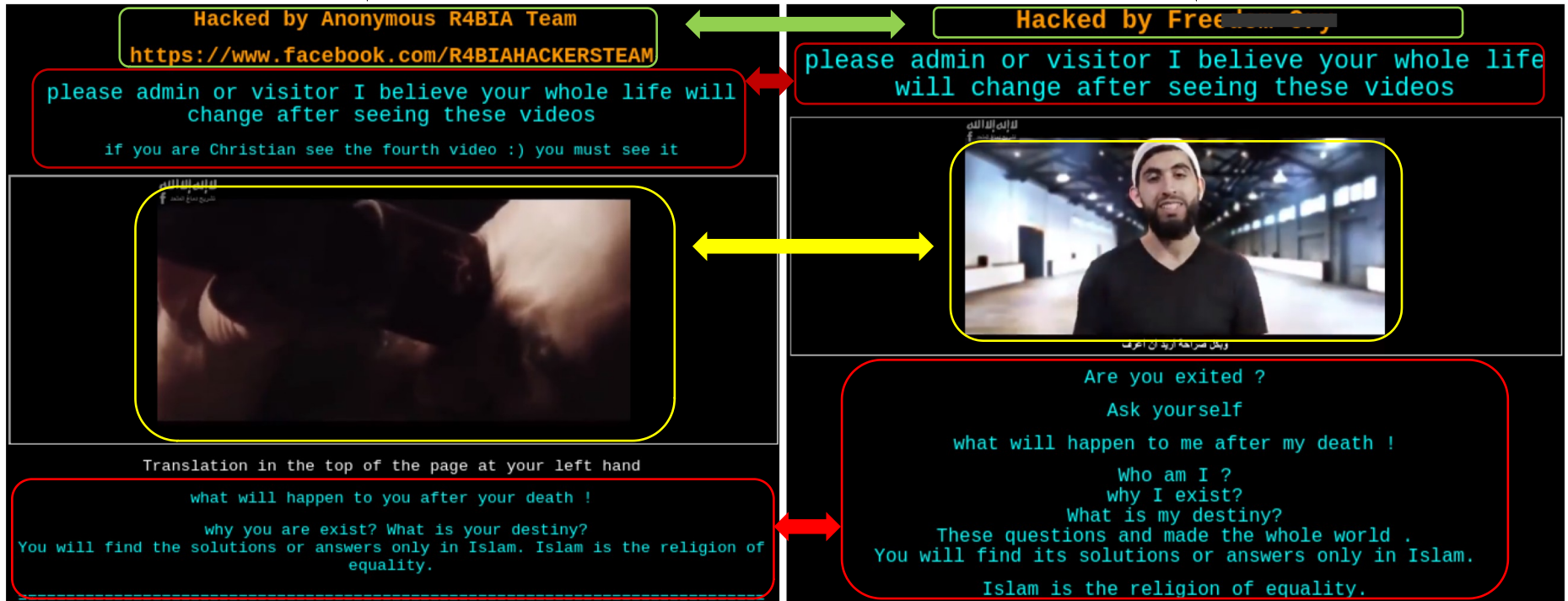
United Islamic Cyber Force
#We Are Stand With Muslims in Europe#
We Are Not Terrorist , And Stop Calling Muslim Terrorist
Respect Exisitance ,Or Expect Resistance
We Are Muslims, We Love Peace
I Challenge Your Governement To Stop Us !

We are : Dr. [redacted]

Key Observations

Template

Customization



Key Observations

1. Actors **cooperate** in teams

Especially if driven by strong ideologies

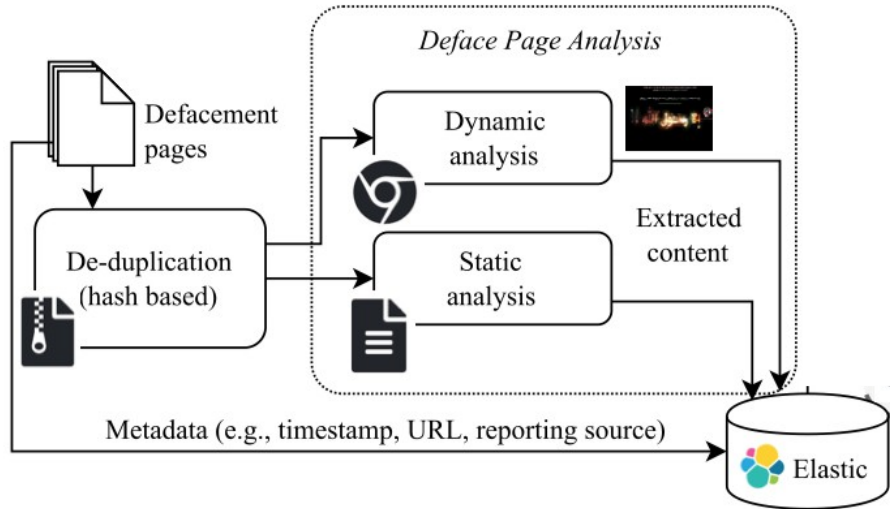
2. Defacements are organized around **campaigns**

3. When a team prepares and runs a campaign, it tends to re-use a common **template** that each member can customize

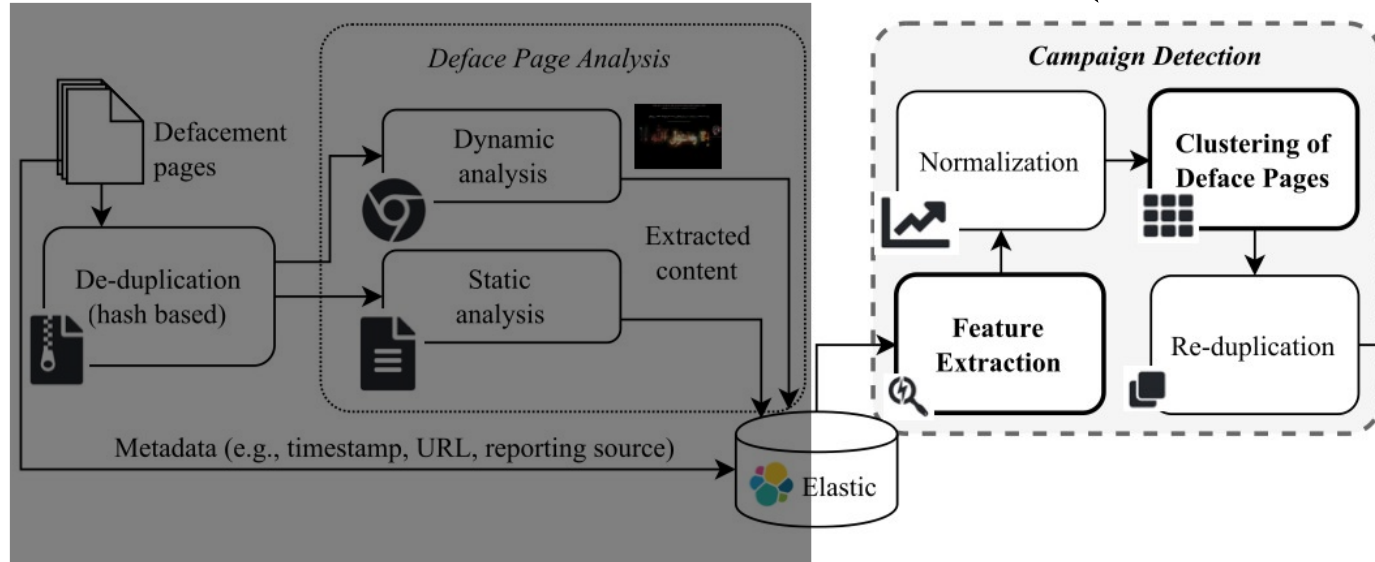
Next Generation* Defacement Explorer (DefPloreX-NG)

(*) 1st generation presented at BH Arsenal

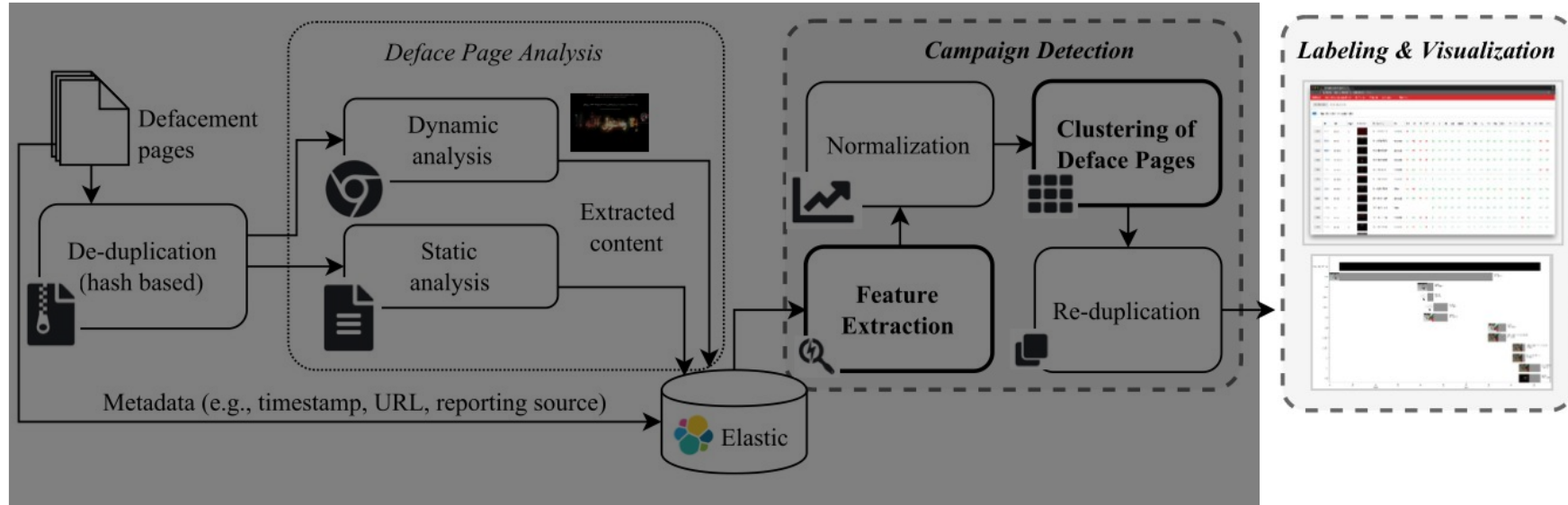
Deface Page Analysis



Campaign Detection



Labeling & Visualization



3,730 records "Free Kashmir" AND tld_name:"in"



Pre-Filtering Search

The Indo-Pakistani conflict





503 clusters params.cluster_size > 1



Campaigns Search

Choose clustering features

	Size	Key	Start	End	N. Attackers	N. Text Hashes	N. TLDs	N. Domains	Timeline
Details	183	European	2013-11-01 12:00	2015-04-01 12:00	4	14	1	183	
Details	153	European	2011-09-01 12:00	2016-04-01 12:00	4	7	1	135	
Details	137	Other	2012-03-01 12:00	2015-11-01 12:00	3	3	1	120	
Details	132	European	2011-06-01 12:00	2016-03-01 12:00	18	39	1	131	
Details	122	European	2015-03-01 12:00	2016-05-01 12:00	3	4	1	122	

JSON	17264154	■ ■ ■ ■ ■		2012-03-18 01:10
JSON	17264170	■ ■ ■ ■ ■		2012-03-18 01:10
JSON	17264172	■ ■ ■ ■ ■		2012-03-18 01:10
JSON	17264197	■ ■ ■ ■ ■		2012-03-18 01:12



Muslim Liberation Army
Security Compromised by ~~XTRMIST~~

Free Kashmir .. Freedom is our goal..// End the Occupation

"Indian Panel Code (Act No. 45 of 1860) CHAPTER -II 18: India . India means the territory of India excluding the State of Jammu and Kashmir ."

This institutionalized impunity with which the killings of civilians by military and police forces in Jammu and Kashmir continues should be a source of shame for India which propagates to be a democracy!

Implementation Details

Features Engineering

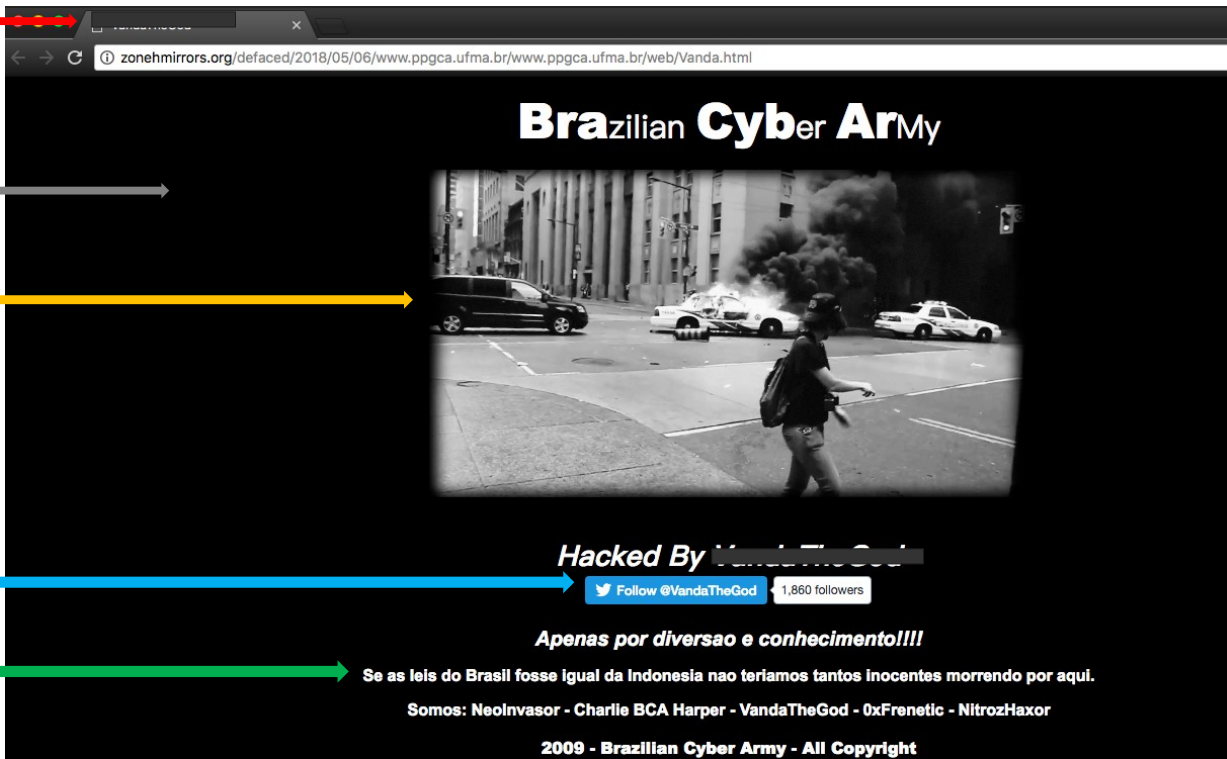
Format of Title →

Visual Features (colors) →

Visual Features (images) →

Social Features →

Structural Features →



The image shows a screenshot of a defaced website with several annotations. The browser address bar shows the URL: `zonehmirrors.org/defaced/2018/05/06/www.ppgca.ufma.br/www.ppgca.ufma.br/web/Vanda.html`. The main content of the page is a black background with the following elements:

- Title:** "Brazilian Cyber Army" in white, bold, sans-serif font.
- Image:** A black and white photograph of a street scene with a person in the foreground, a car, and a building with smoke in the background.
- Social Media:** A Twitter follow button for "@VandaTheGod" with "1,860 followers".
- Text:** "Hacked By VandaTheGod" in white, italicized font.
- Text:** "Apenas por diversao e conhecimento!!!!" in white, italicized font.
- Text:** "Se as leis do Brasil fosse igual da Indonesia nao teriamos tantos inocentes morrendo por aqui." in white font.
- Text:** "Somos: Neolvasor - Charlie BCA Harper - VandaTheGod - 0xFrenetic - NitrozHaxor" in white font.
- Text:** "2009 - Brazilian Cyber Army - All Copyright" in white font.

Annotations on the left side of the image point to these features:

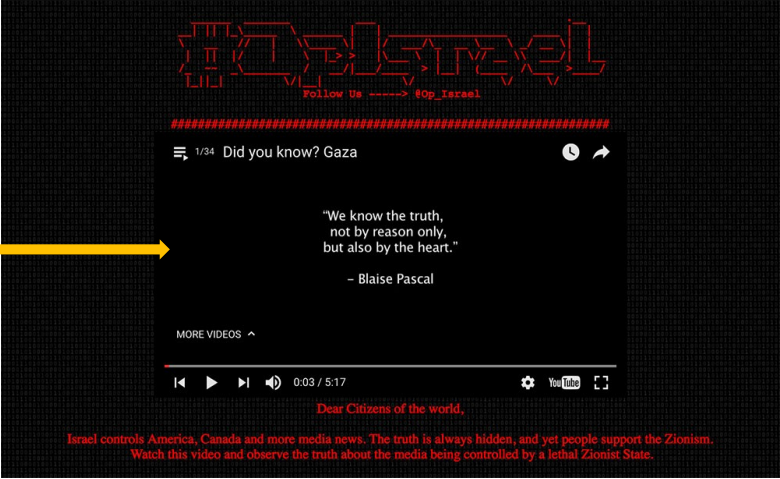
- Format of Title:** A red arrow points to the "Brazilian Cyber Army" title.
- Visual Features (colors):** A grey arrow points to the black background.
- Visual Features (images):** A yellow arrow points to the street scene image.
- Social Features:** A blue arrow points to the Twitter follow button.
- Structural Features:** A green arrow points to the text "Se as leis do Brasil fosse igual da Indonesia..."

Features Engineering



Email Addresses

Multimedia URLs

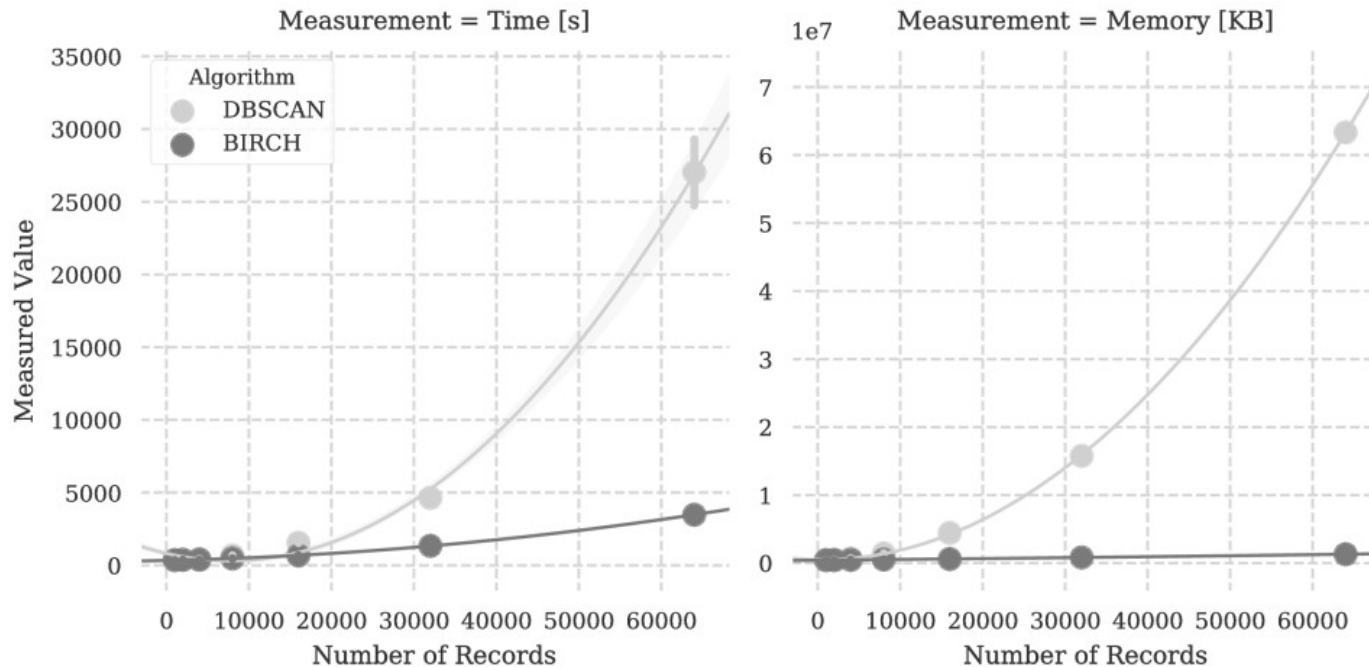


Clustering

- BIRCH
 - Balanced Iterative Reducing and Clustering Hierarchies
 - Do **not** materialize the entire distance matrix
 - Statistical values are efficient to compute
 - Quickly find the closest cluster for each new data points
-

Clustering

- Scalability of BIRCH vs. DBSCAN (10 runs)

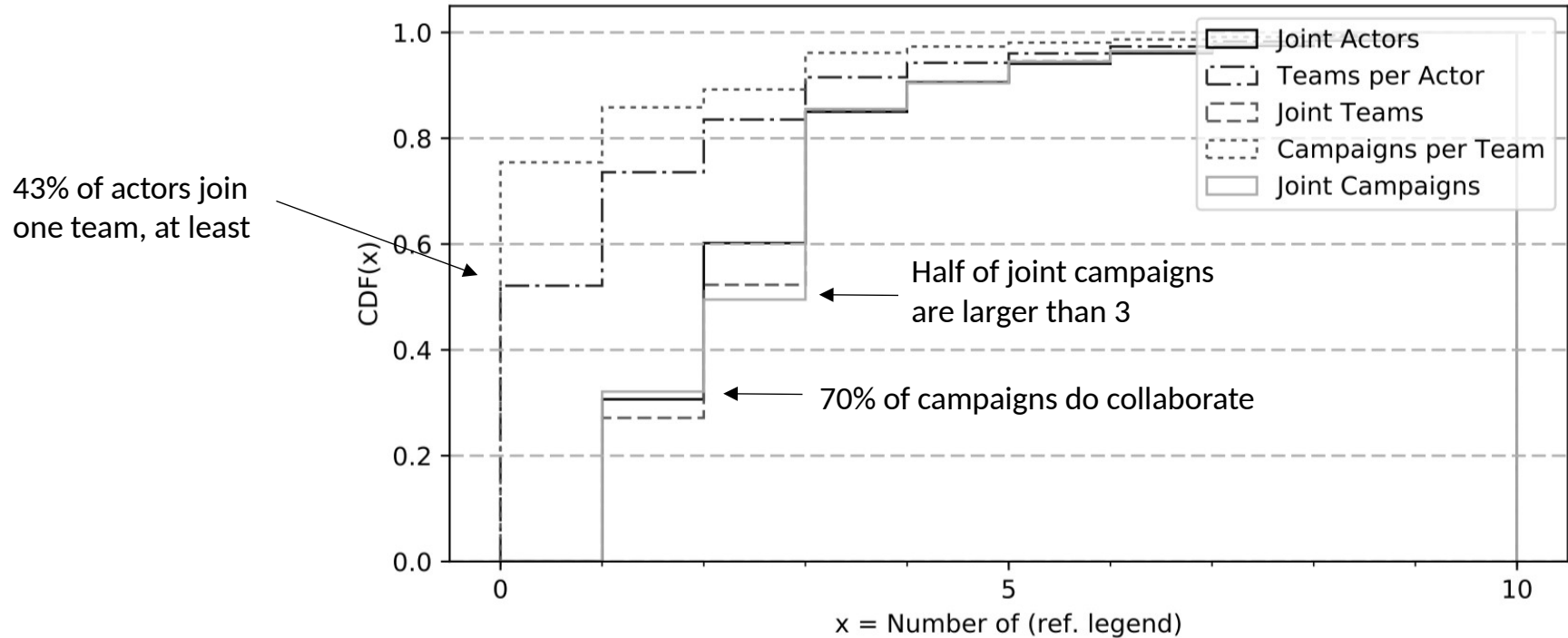


Labeling

- Each cluster is represented by a succinct report
 - Time span
 - Screenshot thumbnails (by perceptual hash)
 - Name of actors and teams
 - Keywords used in campaigns (e.g. #opfrance)
 - Category of targets (e.g, news, governmental sites)
-

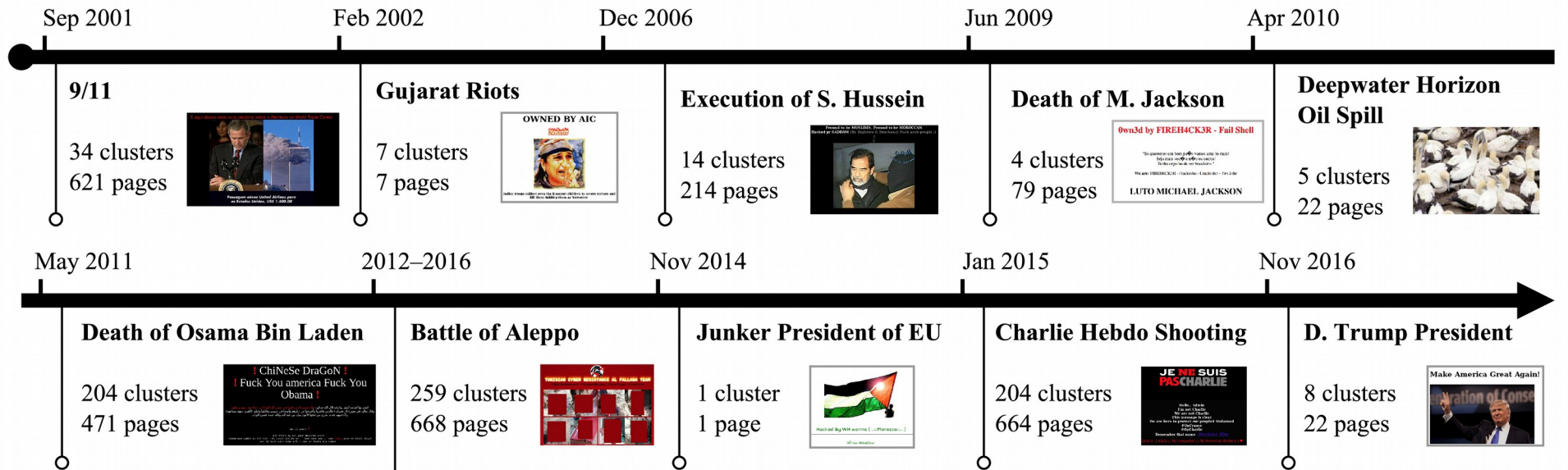
Findings

Organization of Actors



Geopolitical Real-World Events

- Successfully detected

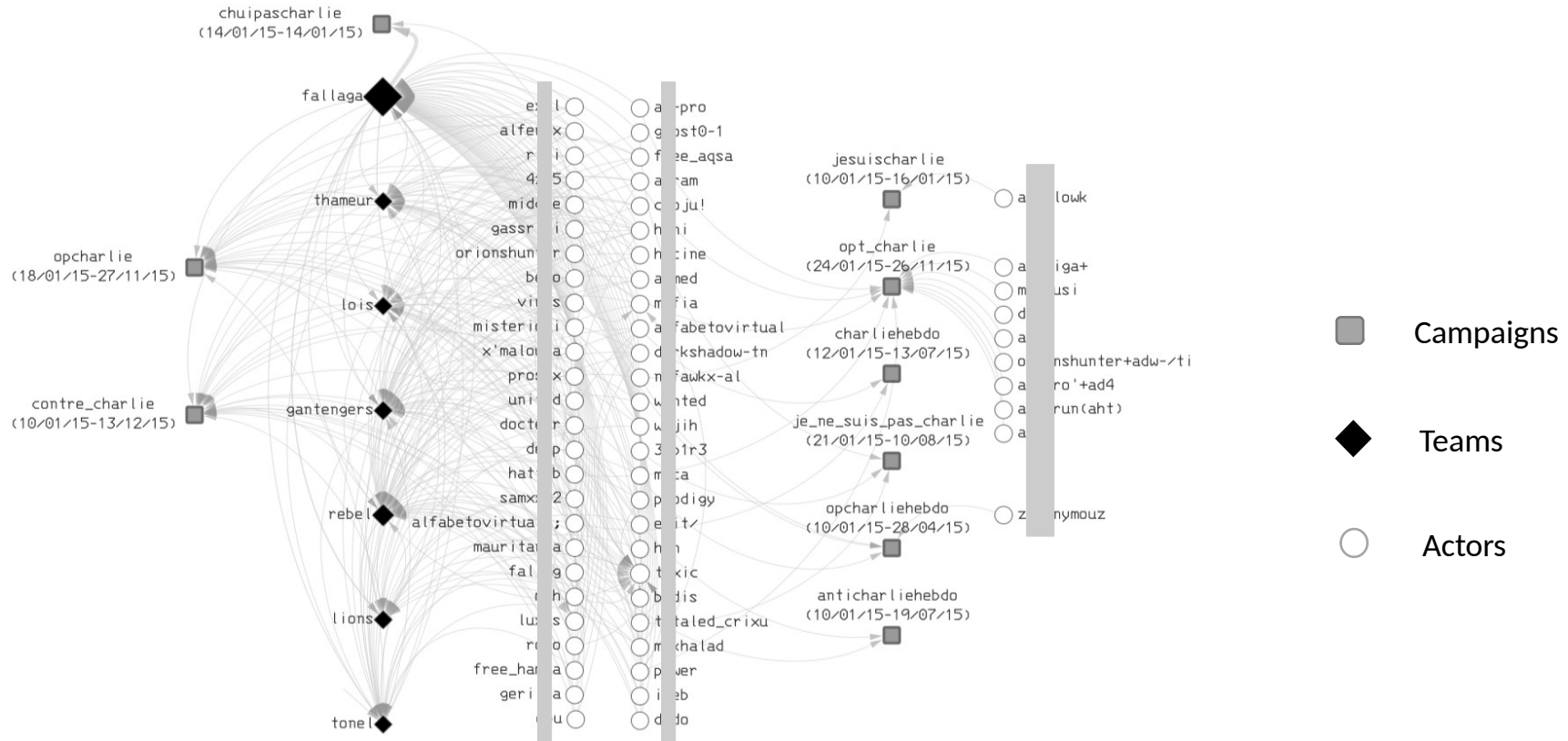


The “Charlie Hebdo” case

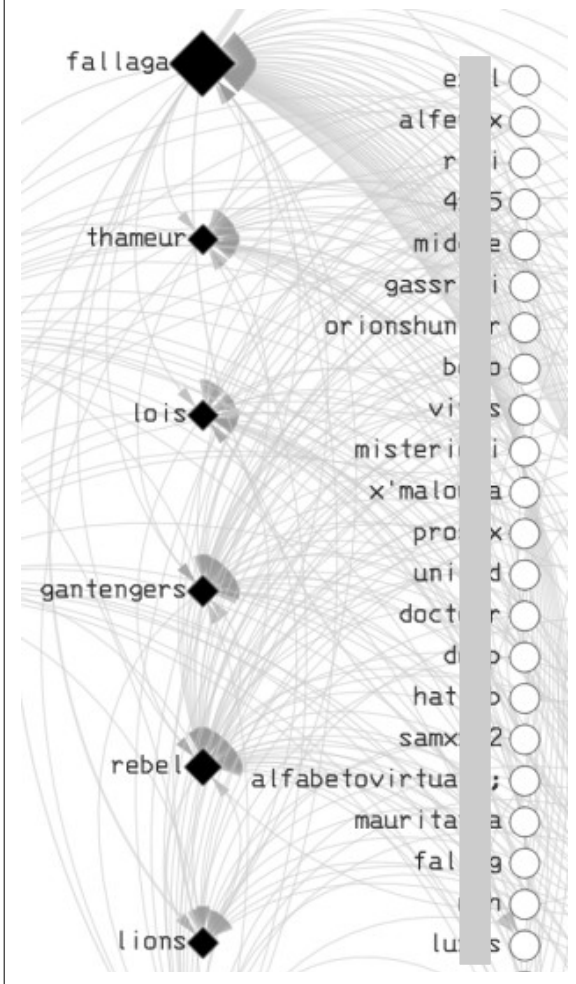
**JE NE SUIS
PAS CHARLIE**

Hello.. Admin
I'm not Charlie
We are not Charlie
This message is clear
We are here to protect our prophet Mohamed
#OpFrance
#OpCharlie
Remember that name [Docteur Hm](#)

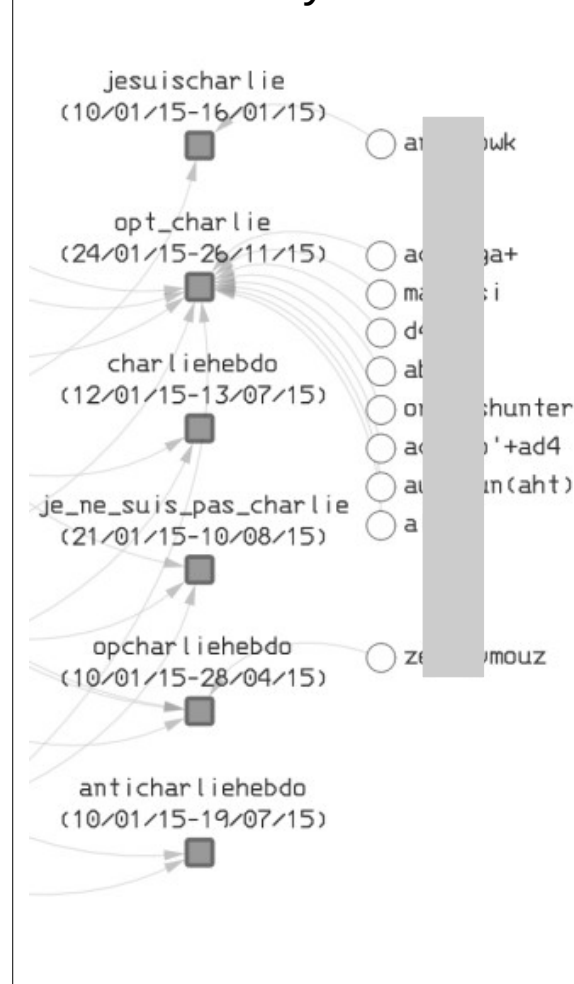
The “Charlie Hebdo” case



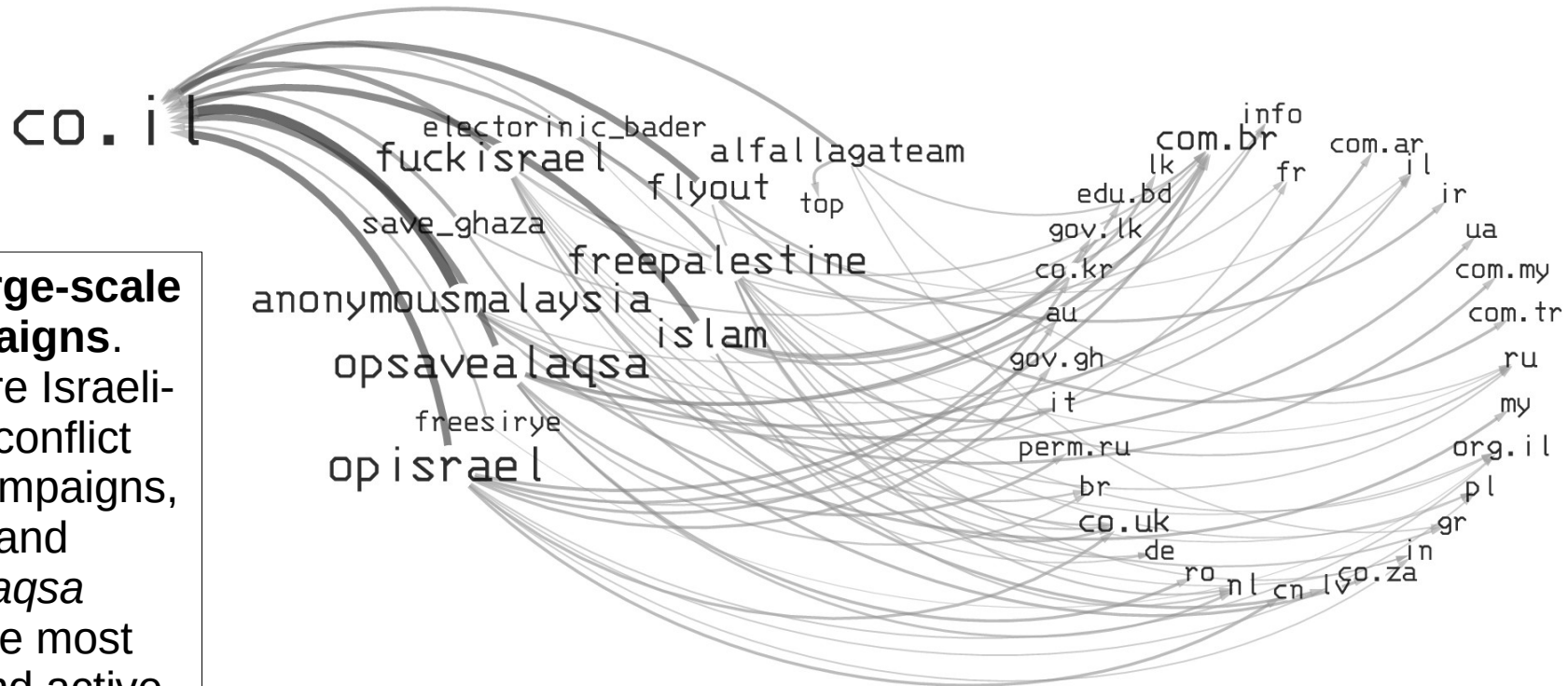
Affiliated actors



Lonely wolfs

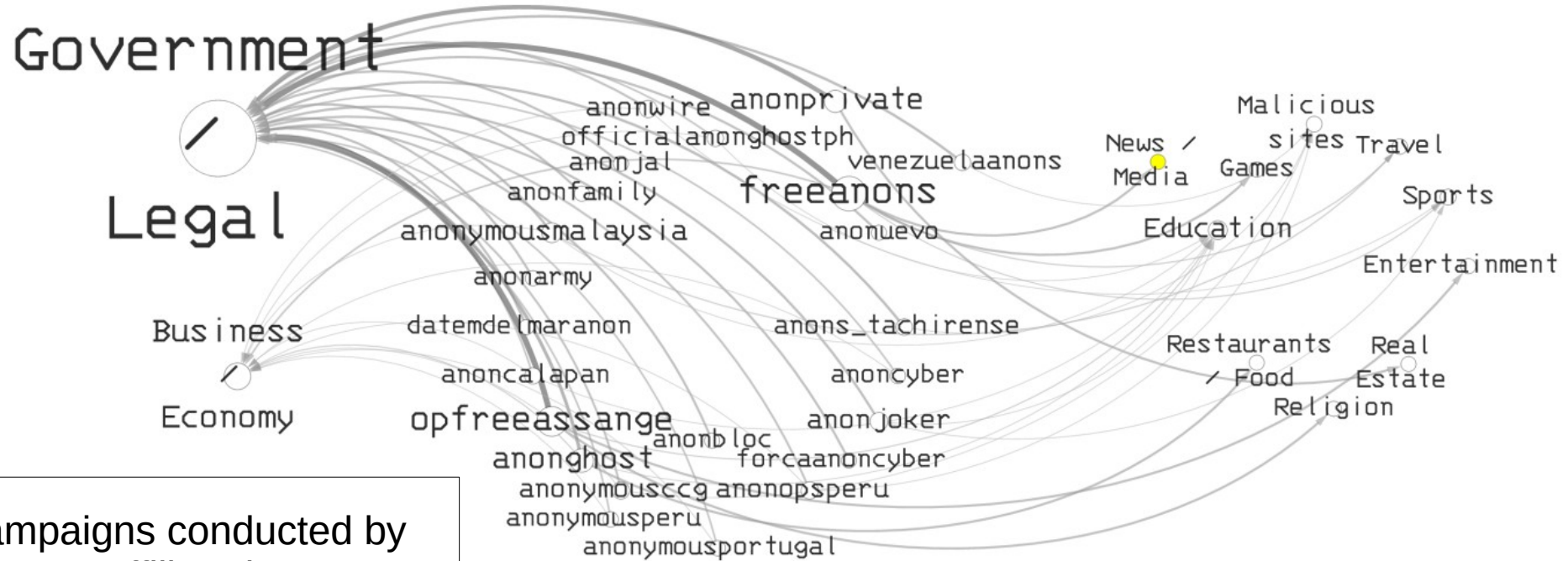


The Israeli-Palestinian conflict



Example of **large-scale joint campaigns**. While the entire Israeli-Palestinian conflict involves 12 campaigns, *opisrael* and *opsavealaqsa* represent the most aggressive and active ones.

Anonymous Operations



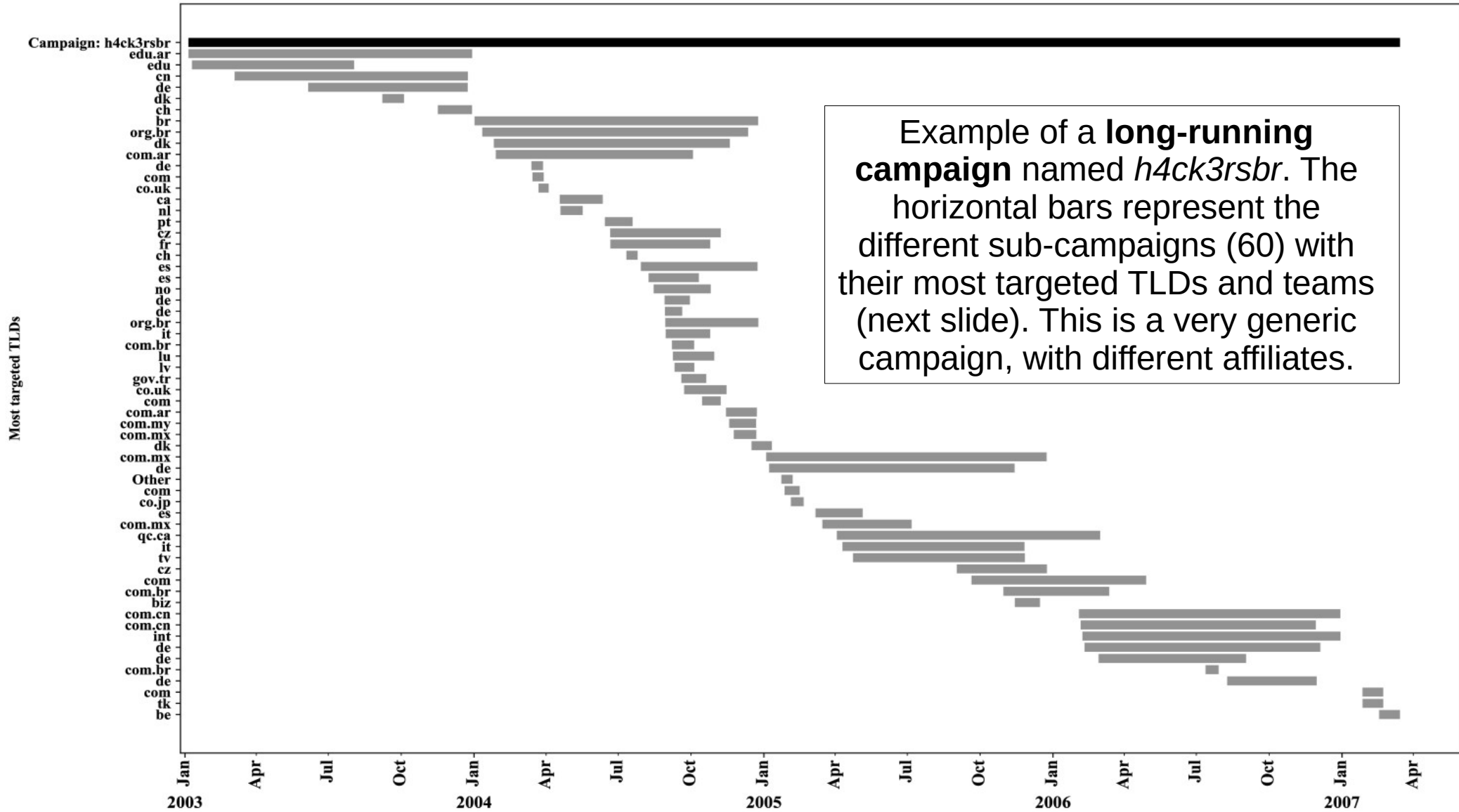
Joint campaigns conducted by *anonymous*-affiliated groups against governmental sites.

Long-Term vs. Aggressive Campaigns

palestine					1					1		29	2	98	902	1629	1124		
powhack					238	557	133	27		28				1	1				
r00t	1	3	18	4	18				5		138	117	10	217	189	375	321		
redhack							69	1973	538	3731	279	22	636	17	22	1	49	6	
rodape										1	1	14	13	11	24	69	29		
samarindahack					8	96	49	40	104	1	9	2			498	4			
syshack					2	2	2	42	2	11	28				67		1	1	
turkish							2		6		101	7	6	120	5	27	2	31	16
venezuela			1				1	18	65	62					2		54	1	

Campaign *savegaza* reacted to war events in the Gaza Strip

savegaza													3	20	12328	8436	257		
spykids					211	554	2350	9685	118		2								
team_cc															7312	1693	197		
whackerz						603	2339	3814	909	2118	1561	283							
	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016



Most targeted TLDs

com.mx

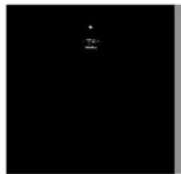


qc.ca



c0d3rz (24 def.)

it



Trustix (38 def.)

tv



Fatal Error (6 def.)



Infektion Group (17 def.)

Oct

Conclusions

- Dark propaganda
 - Prevailing phenomenon
 - Driven by geopolitical motivations
 - Key targets, influencing sites

 - Contribute to make the Internet a better world!
-

Conclusions

- DefPloreX-NG
- GitHub (old code, new ask):
<https://github.com/trendmicro/defplorex>
- Paper*:
<https://documents.trendmicro.com/assets/wp/wp-web-defacement-campaigns-uncovered-gaining-insights-from-deface-pages-using-defplorex-ng.pdf>

(*) Joint work with Federico Maggi, Ryan Flores, Lion Gu and Vincenzo Ciancaglini

Thanks! Questions?

Dr. Marco Balduzzi
@embyte, madlab.it

