PHOENIX & CERBERUS

We haz botnets!

#Honeynet2014

Stefano Schiavoni, Edoardo Colombo

Federico Maggi

Lorenzo Cavallaro

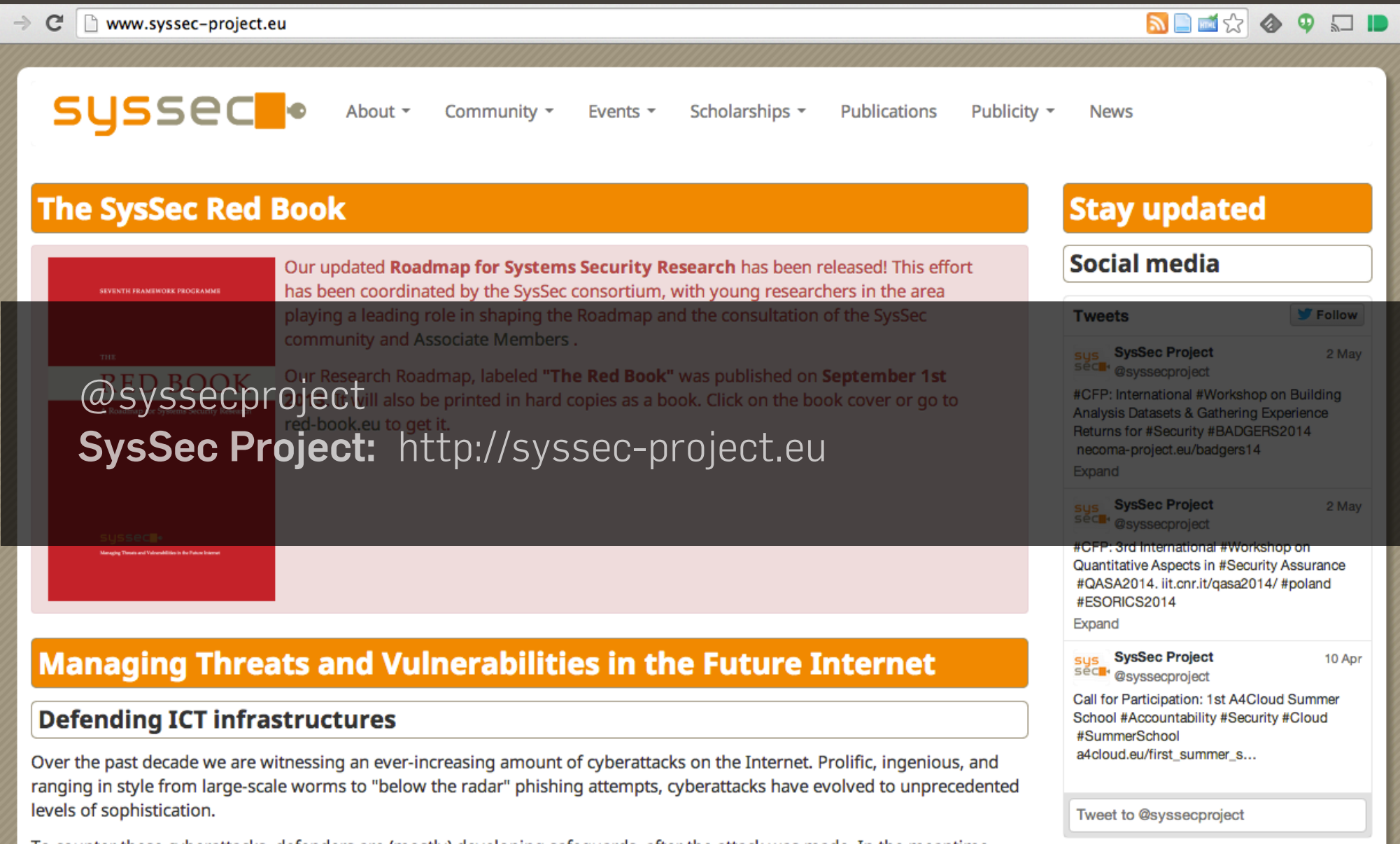
Stefano Zanero

Politecnico Di Milano & Royal Hollway, University of London

POLITECNICO
DI MILANO

BOTNETS

FUNDING ETC.


www.syssec-project.eu
syssec
About
Community
Events
Scholarships
Publications
Publicity
News
The SysSec Red Book
Our updated Roadmap for Systems Security Research has been released! This effort has been coordinated by the SysSec consortium, with young researchers in the area playing a leading role in shaping the Roadmap and the consultation of the SysSec community and Associate Members .
Our Research Roadmap, labeled "The Red Book" was published on September 1st
red-book.eu to get it.
@syssecproject
SysSec Project: http://syssec-project.eu
Managing Threats and Vulnerabilities in the Future Internet
Defending ICT infrastructures
Over the past decade we are witnessing an ever-increasing amount of cyberattacks on the Internet. Prolific, ingenious, and ranging in style from large-scale worms to "below the radar" phishing attempts, cyberattacks have evolved to unprecedented levels of sophistication.
Stay updated
Social media
Tweets
Follow
SysSec Project @syssecproject 2 May
#CFP: International #Workshop on Building Analysis Datasets & Gathering Experience Returns for #Security #BADGERS2014 necoma-project.eu/badgers14
Expand
SysSec Project @syssecproject 2 May
#CFP: 3rd International #Workshop on Quantitative Aspects in #Security Assurance #QASA2014. iit.cnr.it/qasa2014/ #poland #ESORICS2014
Expand
SysSec Project @syssecproject 10 Apr
Call for Participation: 1st A4Cloud Summer School #Accountability #Security #Cloud #SummerSchool a4cloud.eu/first_summer_s...
Tweet to @syssecproject

The bot **encrypts** files on the victim's computer and asks for a **ransom** to recover them.

[1] http://www.cybersec.kent.ac.uk/Survey2.pdf

[2] http://www.zdnet.com/cryptolockers-crimewave-a-trail-of-millions-in-laundered-bitcoin-7000024579/

[3] http://www.theguardian.com/technology/2013/nov/21/us-police-force-pay-bitcoin-ransom-in-cryptolocker-malware-scam

The bot **encrypts** files on the victim's computer and asks for a **ransom** to recover them.

- First appeared in early September 2013

---

[1] http://www.cybersec.kent.ac.uk/Survey2.pdf

[2] http://www.zdnet.com/cryptolockers-crimewave-a-trail-of-millions-in-laundered-bitcoin-7000024579/

[3] http://www.theguardian.com/technology/2013/nov/21/us-police-force-pay-bitcoin-ransom-in-cryptolocker-malware-scam

The bot **encrypts** files on the victim's computer and asks for a **ransom** to recover them.

- First appeared in early September 2013
- In the UK 41% victims paid the ransom[1]

[1] http://www.cybersec.kent.ac.uk/Survey2.pdf

[2] http://www.zdnet.com/cryptolockers-crimewave-a-trail-of-millions-in-laundered-bitcoin-7000024579/

[3] http://www.theguardian.com/technology/2013/nov/21/us-police-force-pay-bitcoin-ransom-in-cryptolocker-malware-scam

The bot **encrypts** files on the victim's computer and asks for a **ransom** to recover them.

- First appeared in early September 2013
- In the UK 41% victims paid the ransom[1]
- Earnings estimated at 27 million USD on Dec 18 2013[2]

---

[1] http://www.cybersec.kent.ac.uk/Survey2.pdf

[2] http://www.zdnet.com/cryptolockers-crimewave-a-trail-of-millions-in-laundered-bitcoin-7000024579/

[3] http://www.theguardian.com/technology/2013/nov/21/us-police-force-pay-bitcoin-ransom-in-cryptolocker-malware-scam

The bot **encrypts** files on the victim's computer and asks for a **ransom** to recover them.

- First appeared in early September 2013
- In the UK 41% victims paid the ransom[1]
- Earnings estimated at 27 million USD on Dec 18 2013[2]
- Massachusetts police have admitted to paying a ransom[3]

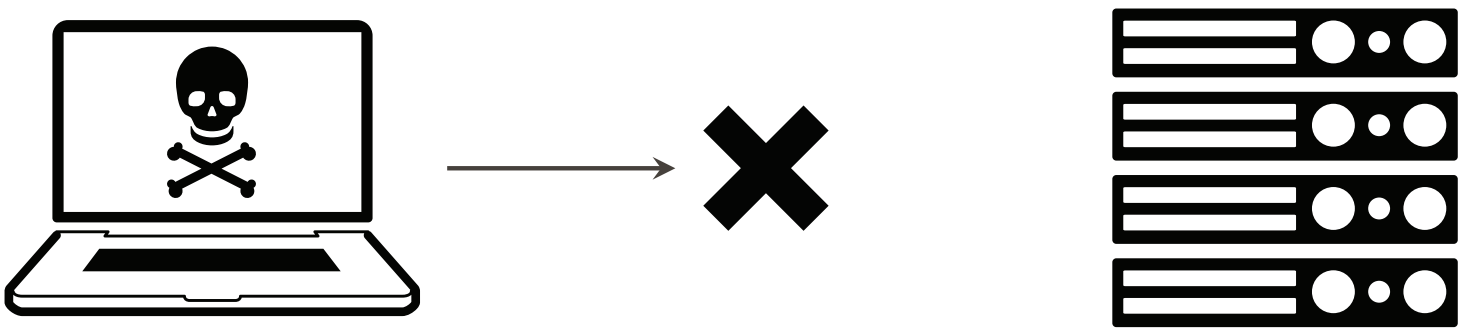
---

[1] http://www.cybersec.kent.ac.uk/Survey2.pdf

[2] http://www.zdnet.com/cryptolockers-crimewave-a-trail-of-millions-in-laundered-bitcoin-7000024579/

[3] http://www.theguardian.com/technology/2013/nov/21/us-police-force-pay-bitcoin-ransom-in-cryptolocker-malware-scam

CENTRALIZED BOTNETS > MITIGATION



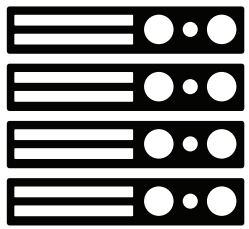
- **C&C channel:** single point of failure.
- **Rallying Mechanisms:** the countermeasure.

BOTNETS > DOMAIN GENERATION ALGORITHMS

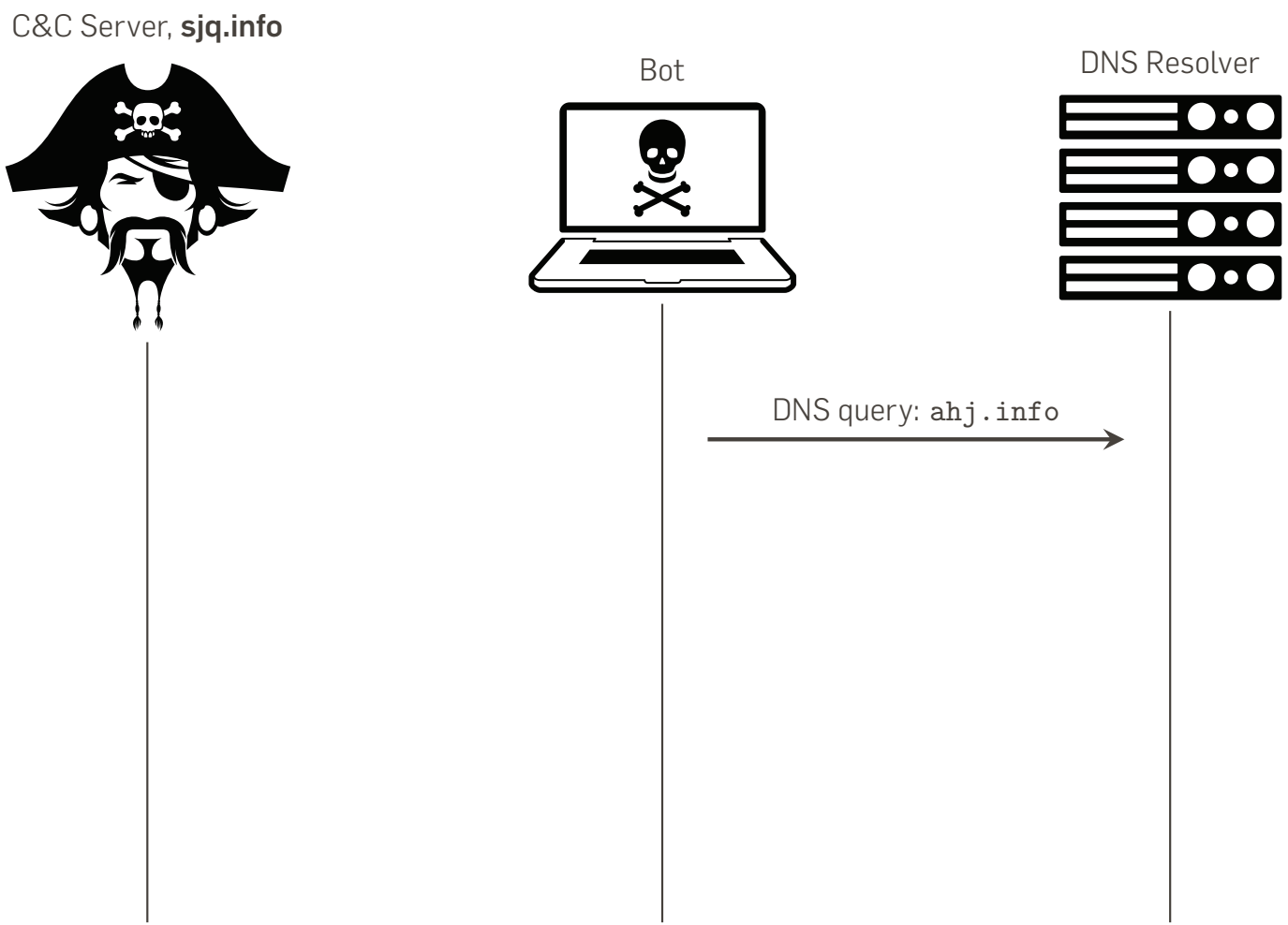
C&C Server, **sjq.info**

Bot

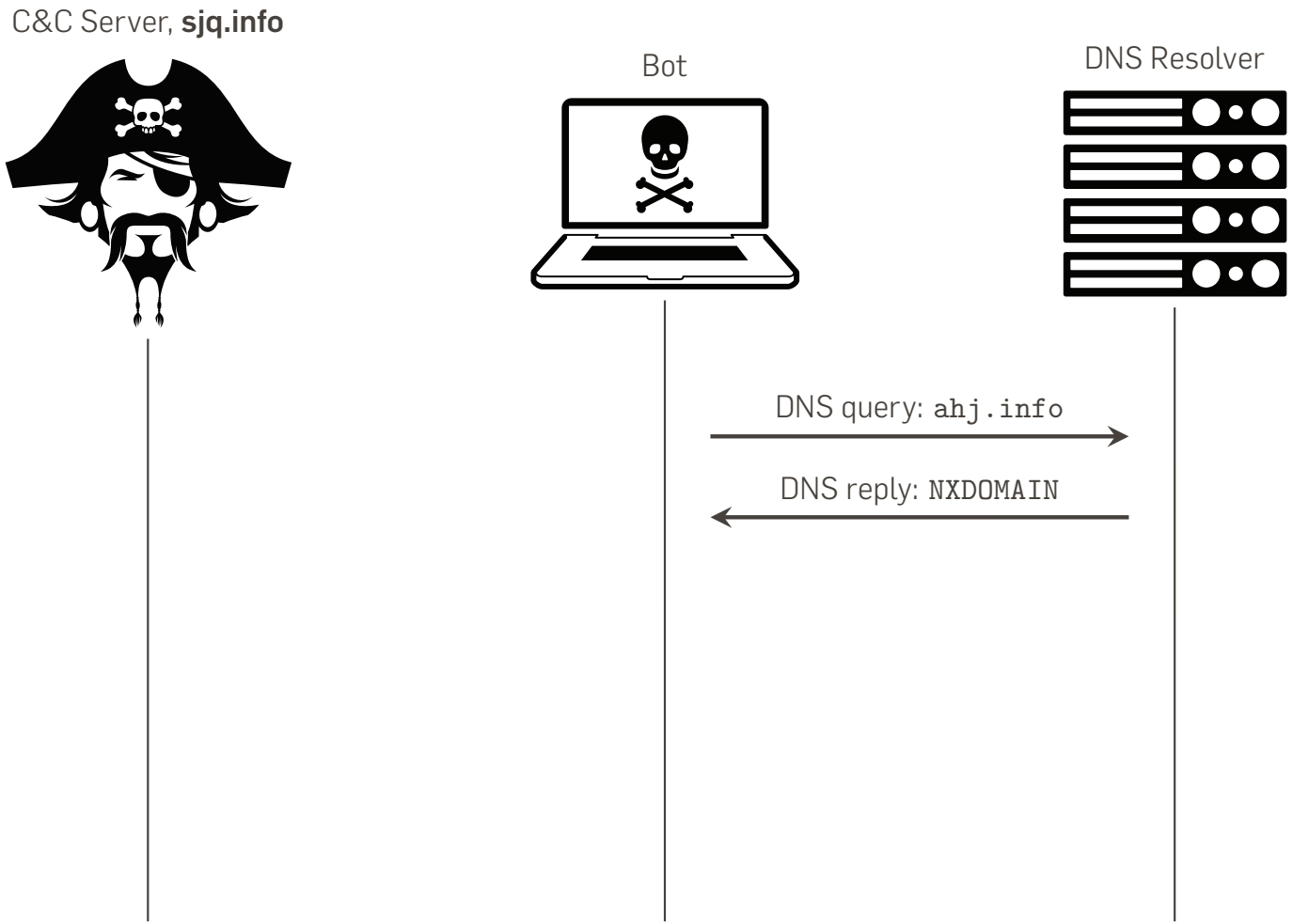
DNS Resolver



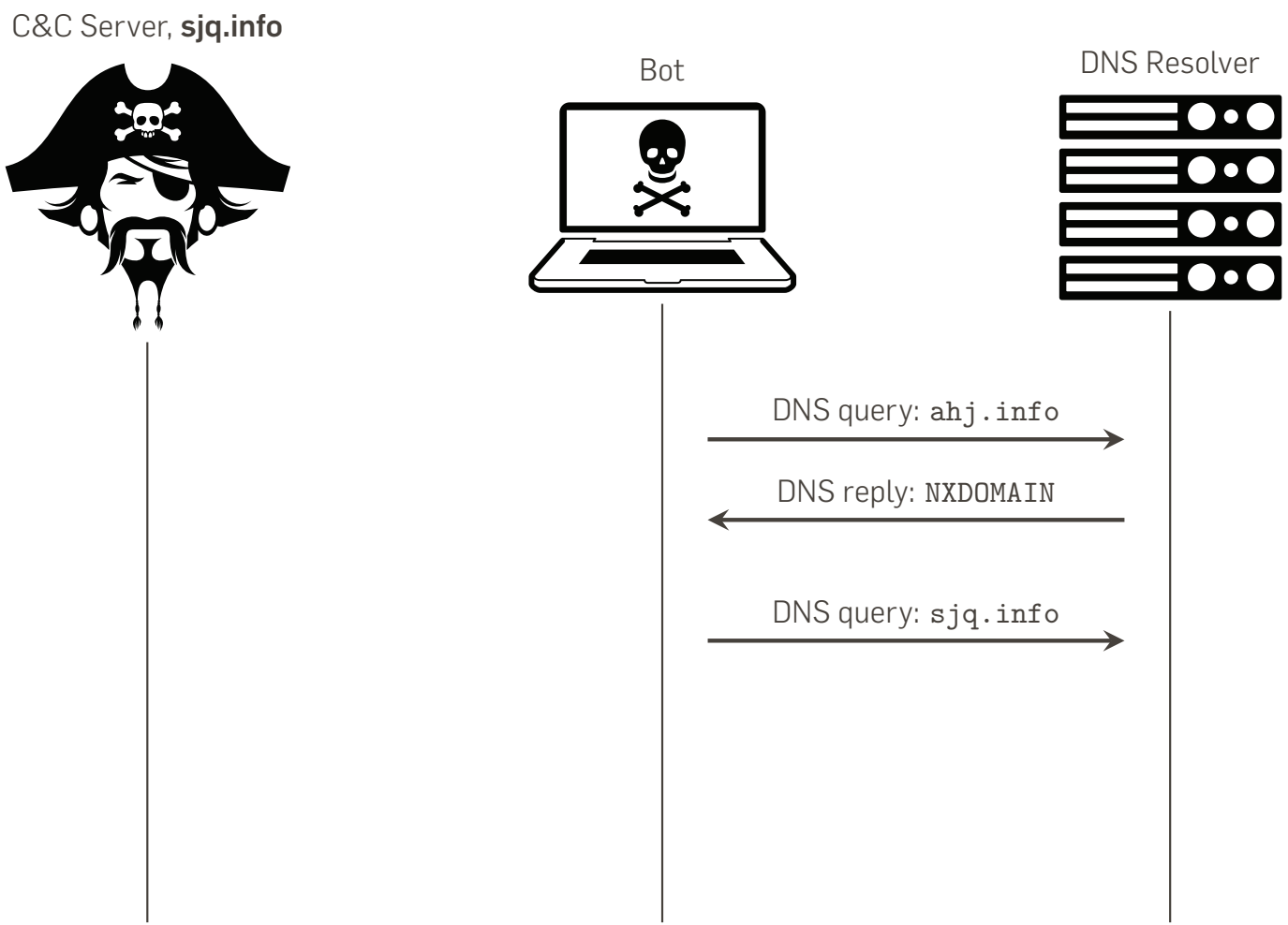
BOTNETS > DOMAIN GENERATION ALGORITHMS



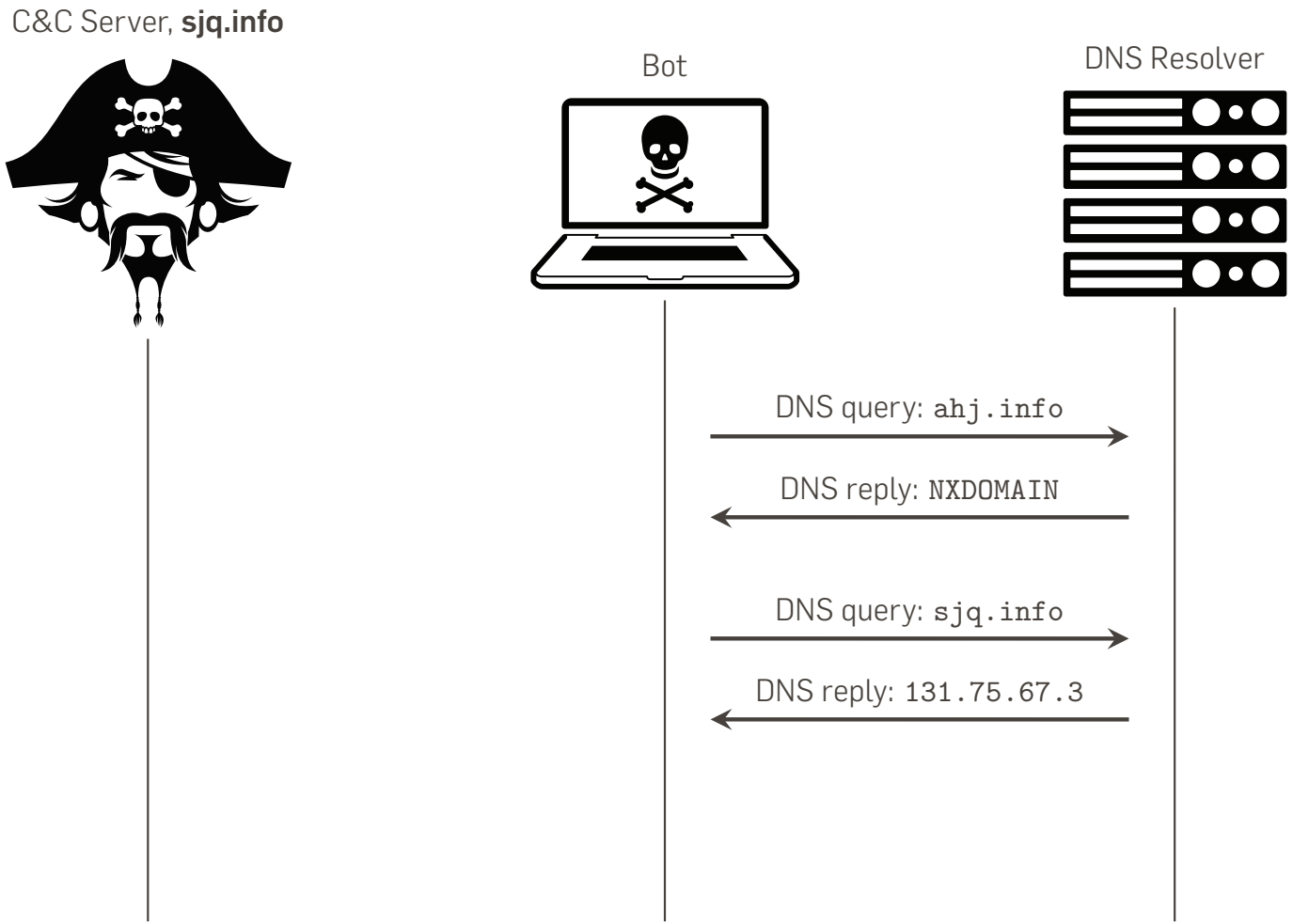
BOTNETS > DOMAIN GENERATION ALGORITHMS



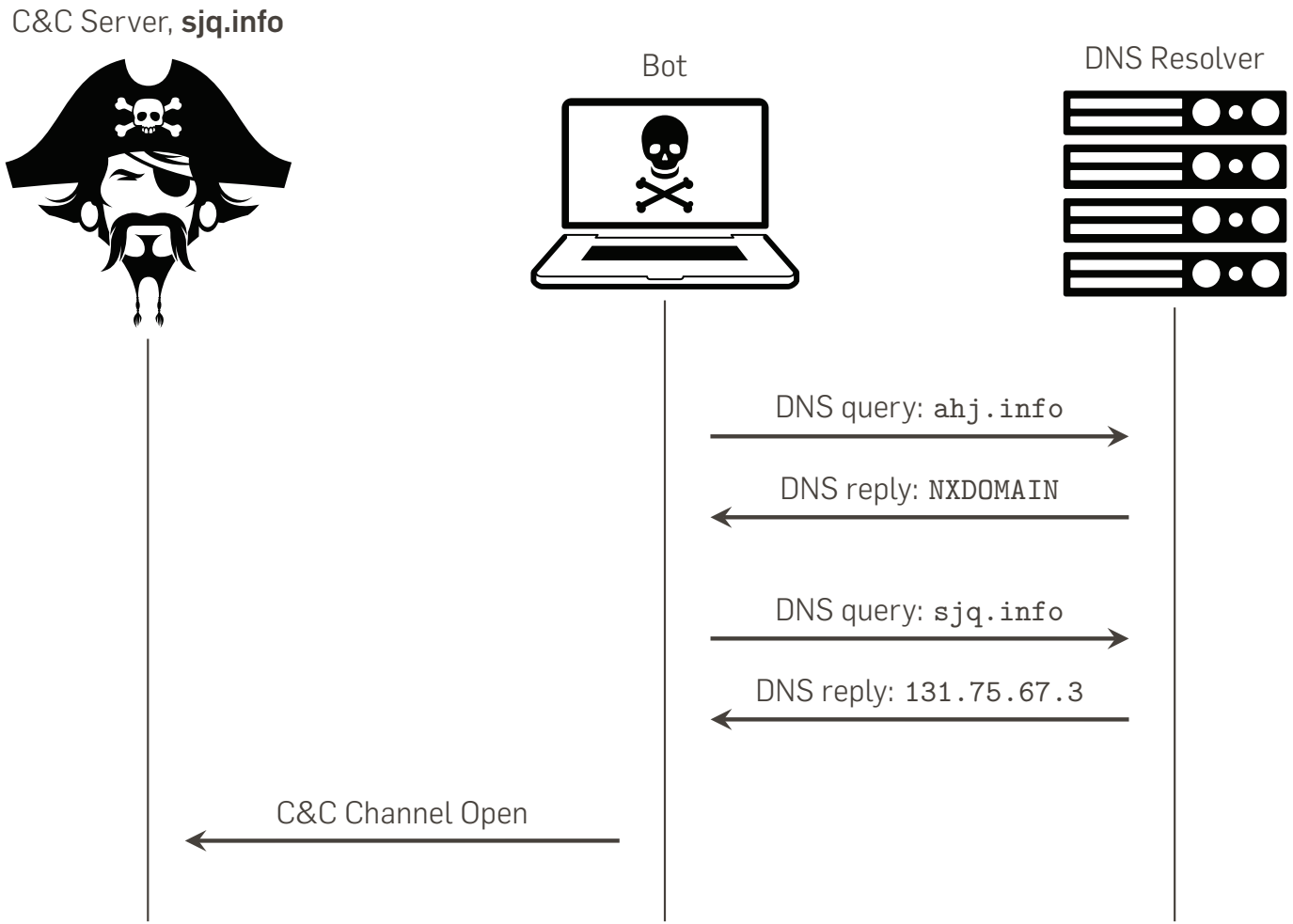
BOTNETS > DOMAIN GENERATION ALGORITHMS



BOTNETS > DOMAIN GENERATION ALGORITHMS



BOTNETS > DOMAIN GENERATION ALGORITHMS



DGA > BENEFITS FOR THE BOTMASTERS

- **Asymmetry** Botmasters Vs Defenders
  - → Thousands of domain names,
  - → only one is the right one.
- **Blacklists** do not work well

Limitations of current **research approaches**:

Limitations of current **research approaches**:

- **Supervised:** require labeled data

Limitations of current **research approaches**:

- **Supervised:** require labeled data
  - → "That domain name is known to be DGA generated",
  - → "That other domain is not".

Limitations of current **research approaches**:

- **Supervised:** require labeled data
  - → "That domain name is known to be DGA generated",
  - → "That other domain is not".
- Work at the **lower levels** of the **DNS hierarchy:**
  - → not so easy to deploy,
  - → privacy (visibility of the hosts' IP addresses).

PHOENIX

Phoenix clusters DGA-generated domains from a list of of **domains known to be used by botnets**.

The core of Phoenix is its ability to **separate DGA from non-DGA** domains, using **linguistic features**.

(in a few slides)

Sources of malicious domains:

- **EXPOSURE** http://exposure.iseclab.org
- **MLD** http://www.malwaredomainlist.com
- ...and of course some **reversing** :-)

Meaningful Word Ratio (English dict)

$d = \texttt{facebook.com}$

$$R(d) = \frac{|\texttt{face}| + |\texttt{book}|}{|\texttt{facebook}|} = 1$$

likely **non-DGA generated**

$d = \texttt{pub03str.info}$

$$R(d) = \frac{|\texttt{pub}|}{|\texttt{pub03str}|} = 0.375.$$

likely **DGA generated**

N-gram Popularity (English dict)

$d =$ `facebook.com`

`fa`	`ac`	`ce`	`eb`	`bo`	`oo`	`ok`
109	343	438	29	118	114	45

mean: $S_2 = 170.8$

likely **non-DGA geneated**

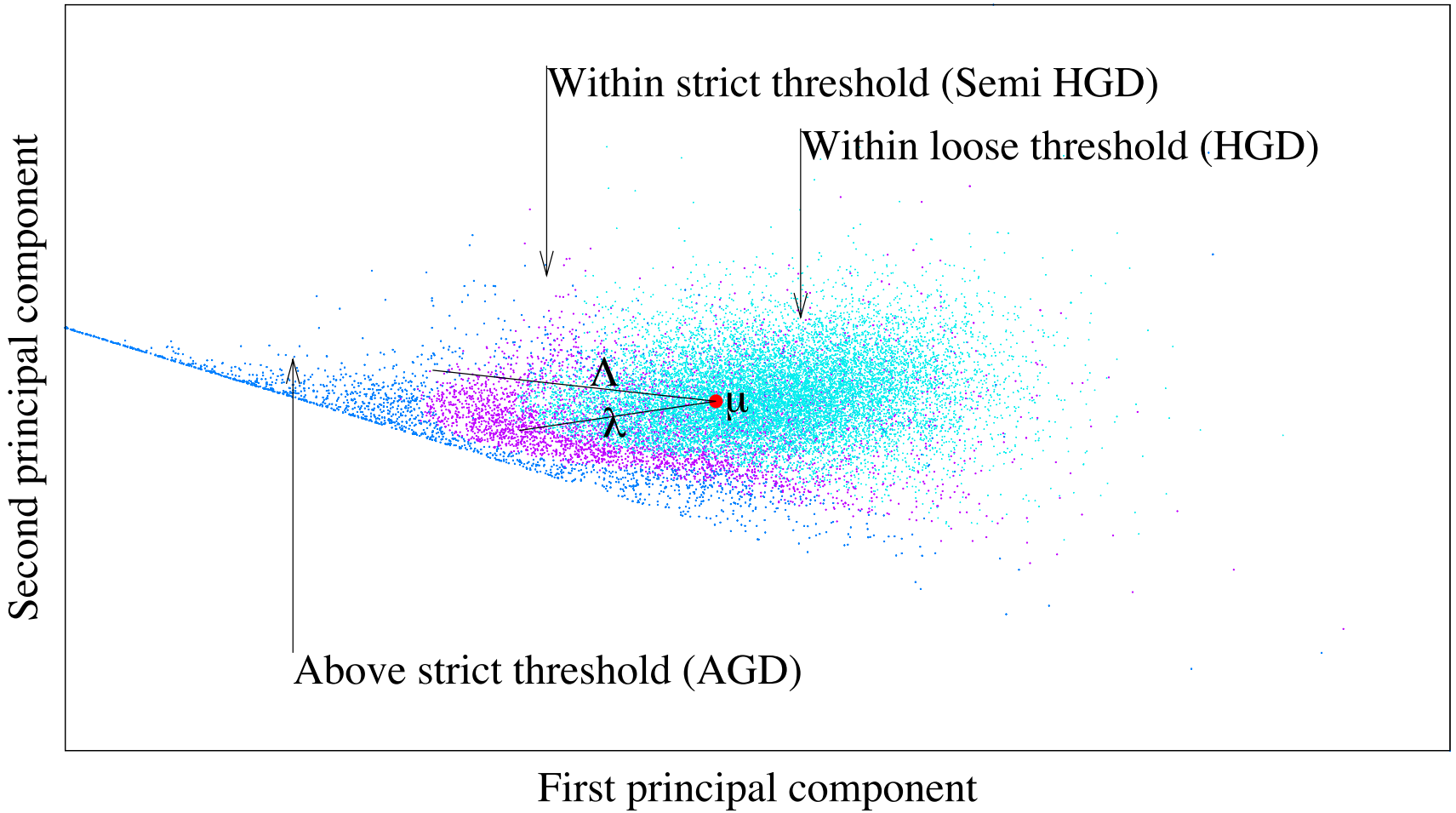
$d =$ `aawrqv.com`

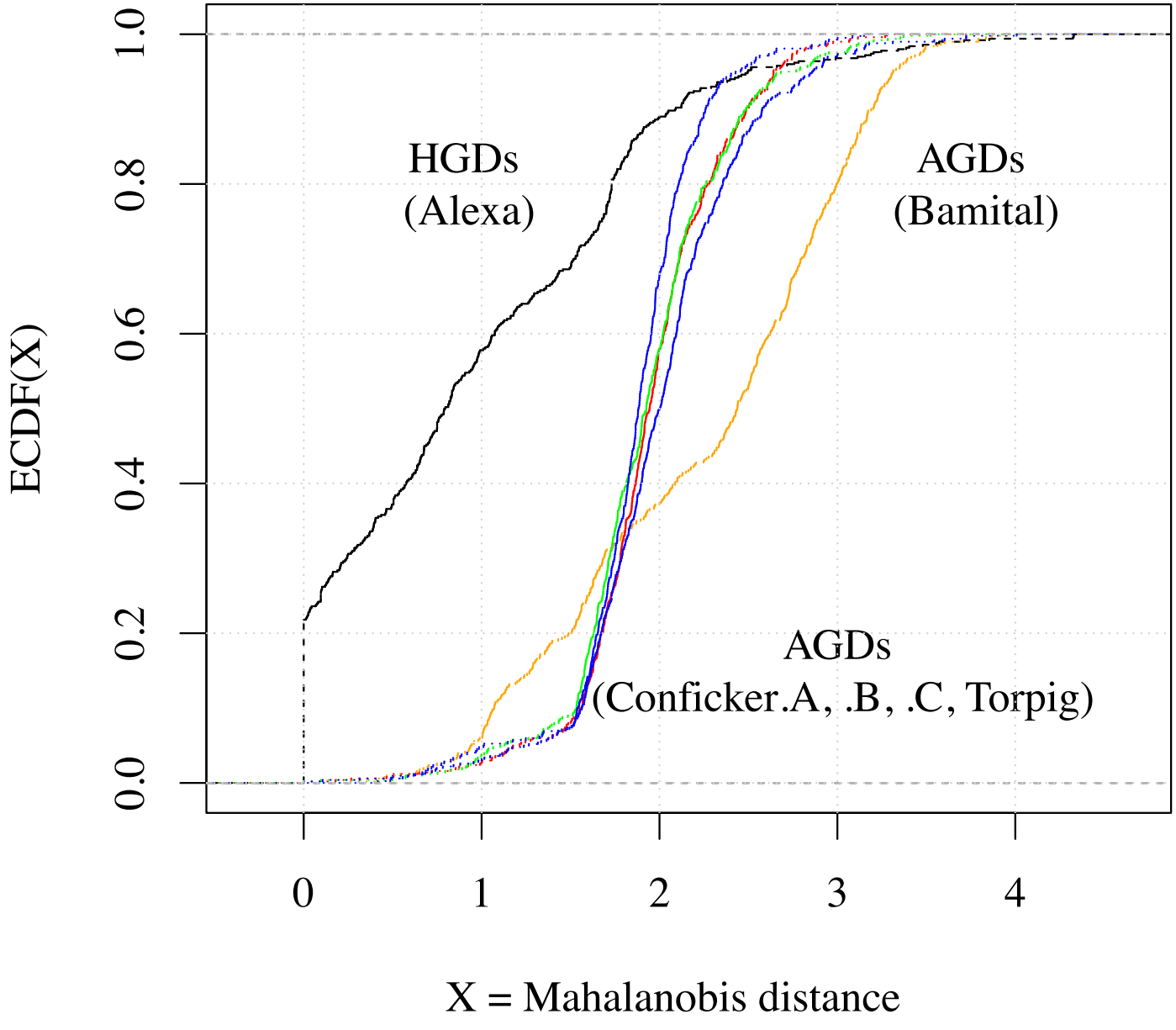
`aa`	`aw`	`wr`	`rq`	`qv`
4	45	17	0	0

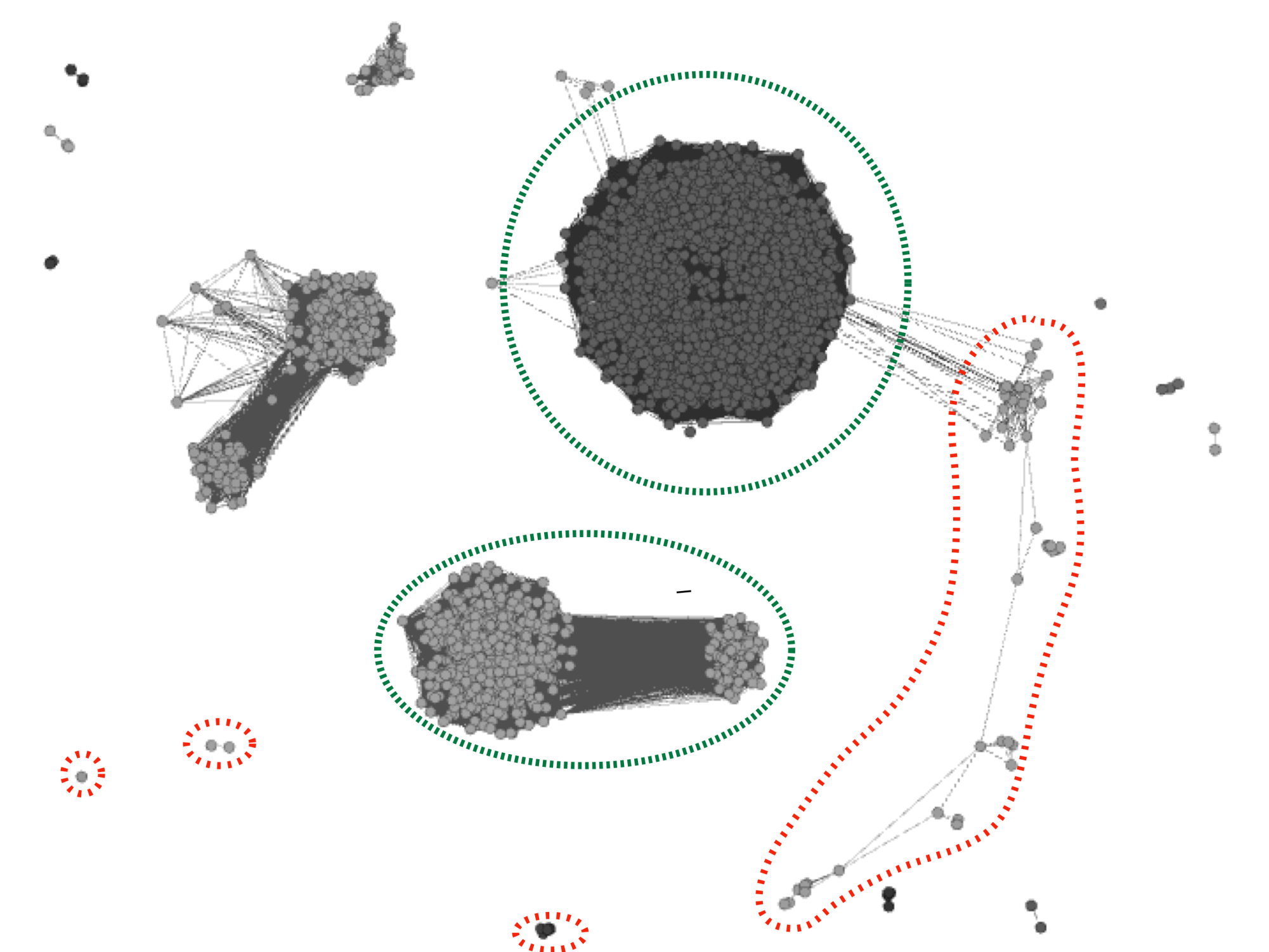
mean: $S_2 = 13.2$

likely **DGA generated**

PHOENIX > DGA VS NON-DGA




ECDF(X)
0.0
0.2
0.4
0.6
0.8
1.0
HGDs
(Alexa)
AGDs
(Bamital)
AGDs
(Conficker.A, .B, .C, Torpig)
0
1
2
3
4
X = Mahalanobis distance

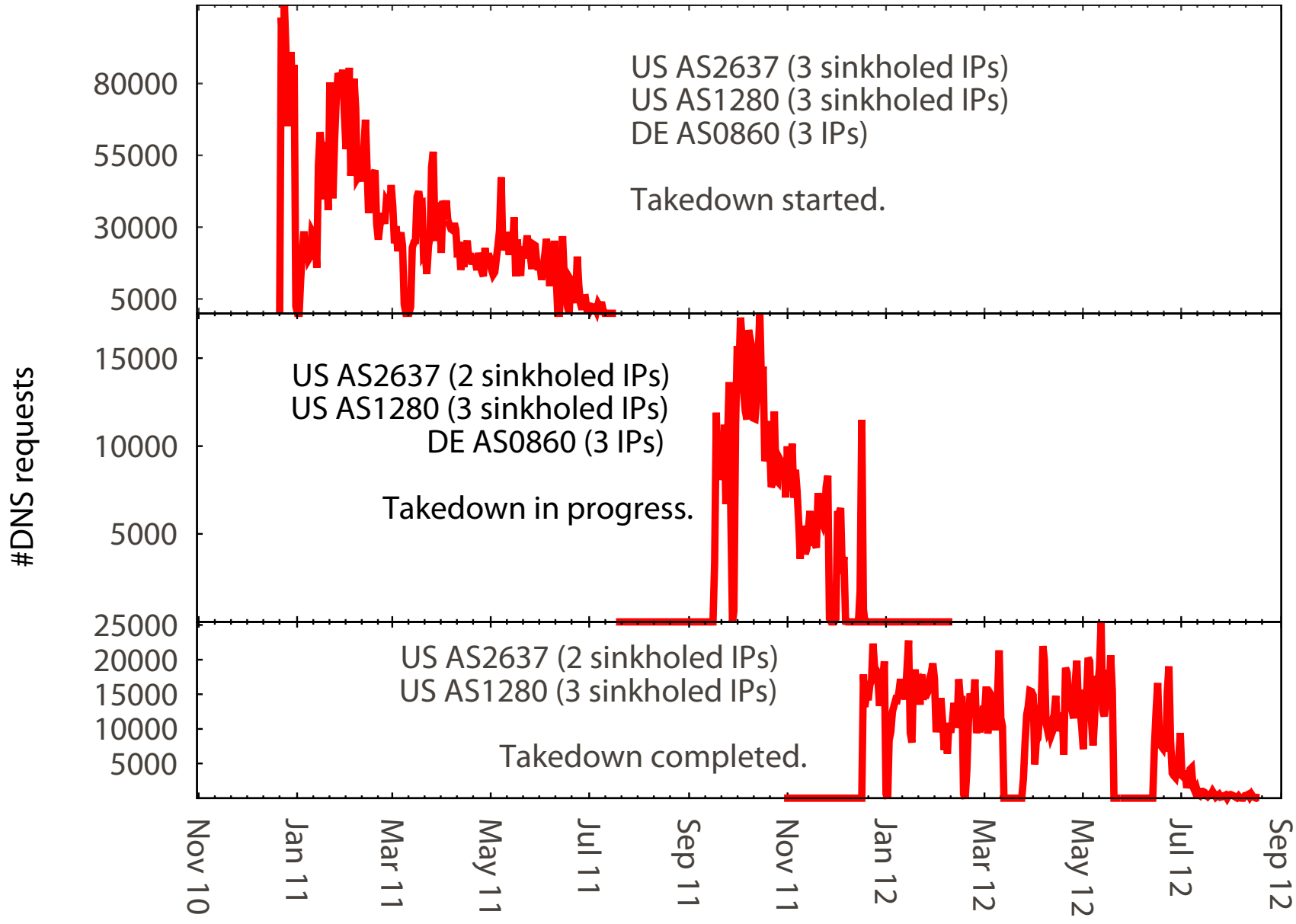


PHOENIX > RESULTS (1 WEEK)

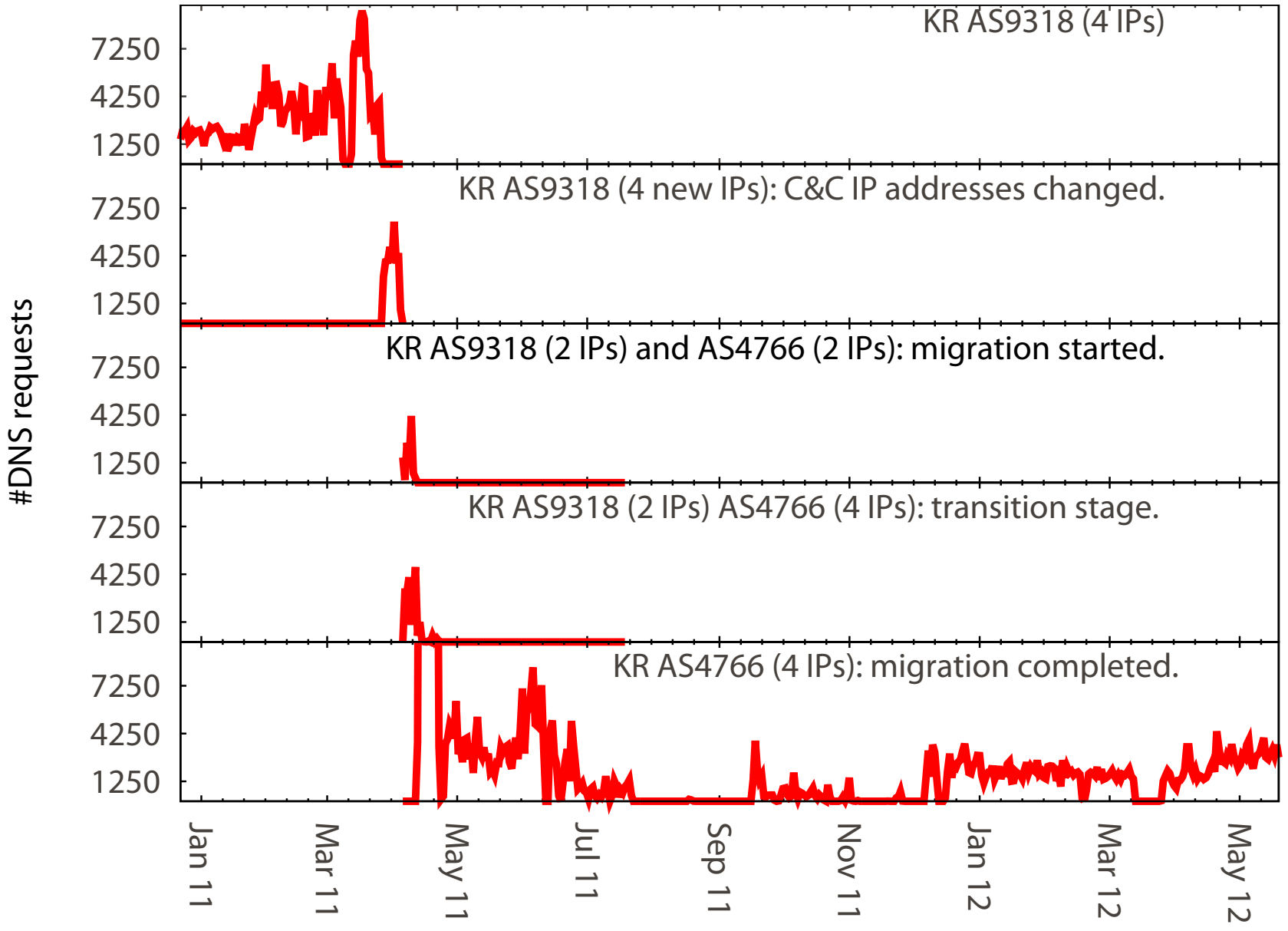
Cluster f105c	
IPs:	176.74.176.175
	208.87.35.107
Domains:	cvq.com
	epu.org
	bwn.org
(Botnet:	Palevo)

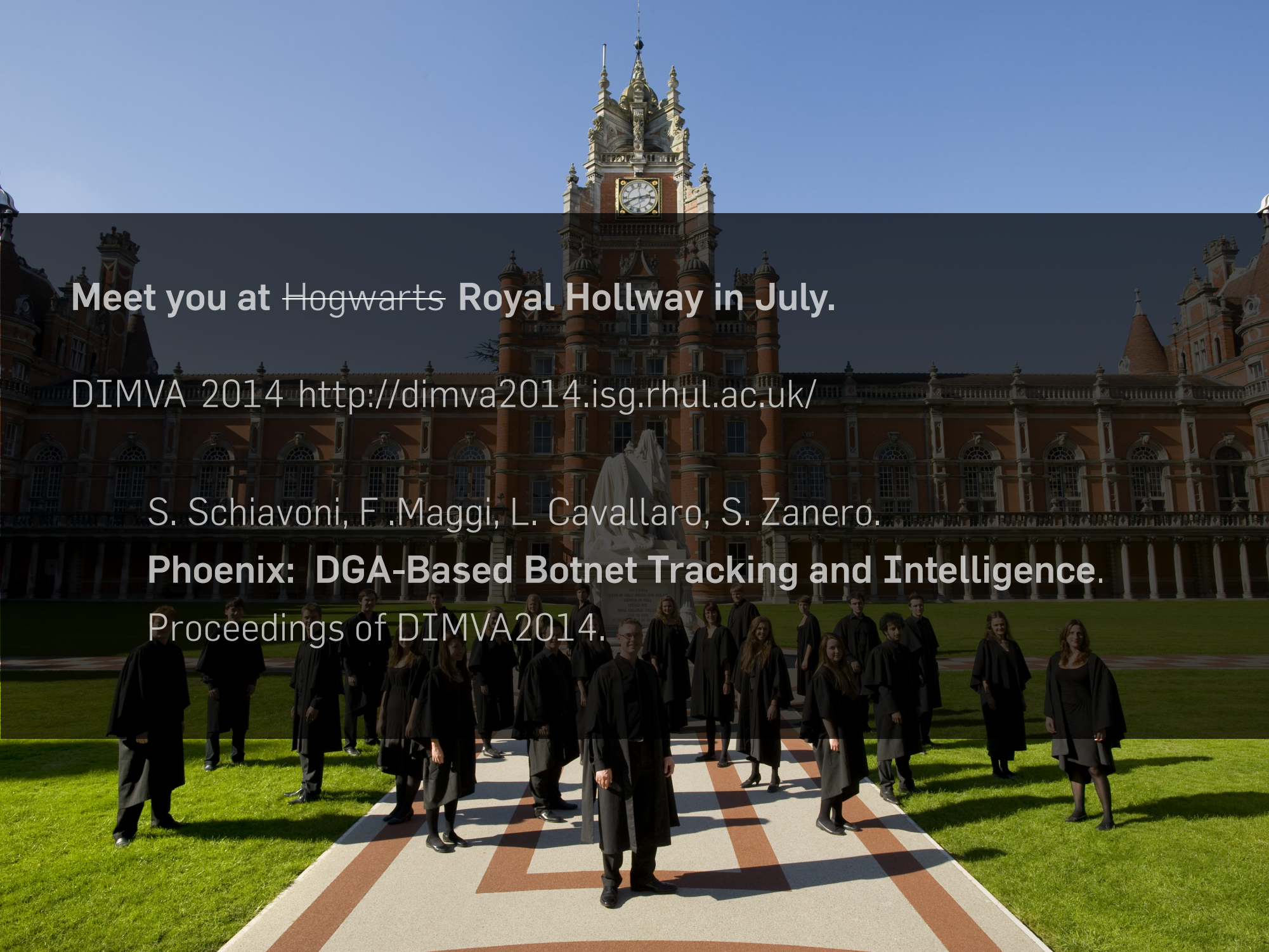
Cluster 0f468	
IPs:	217.119.57.22
	91.215.158.57
	178.162.164.24
	94.103.151.195
Domains:	jhhfghf7.tk
	faukiijjj25.tk
	pvgvy.tk
(Botnet:	Sality)

PHOENIX > TRACKING MIGRATIONS



PHOENIX > TRACKING MIGRATIONS



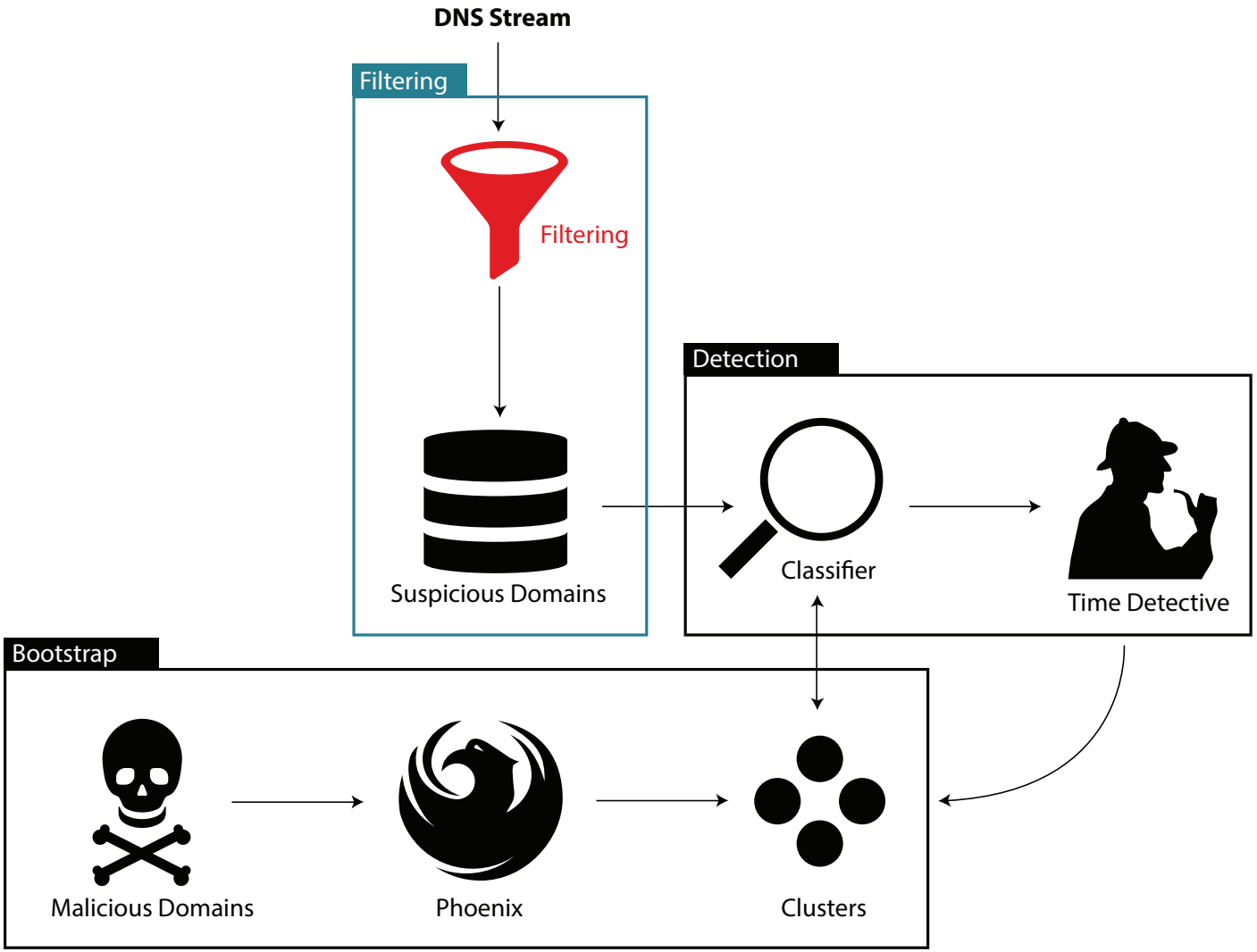

Meet you at ~~Hogwarts~~ Royal Hollway in July.
DIMVA 2014 http://dimva2014.isg.rhul.ac.uk/
S. Schiavoni, F .Maggi, L. Cavallaro, S. Zanero.
Phoenix: DGA-Based Botnet Tracking and Intelligence.
Proceedings of DIMVA2014.

Leverages historical DNS data:

- **Unable** to deal with **new DGAs**
- Unseen "domain→IP" mapping are simply **discarded.**

CERBERUS

CERBERUS > FILTERING



Insight a malicious domain automatically generated will not become popular.

Alexa Top 1M Whitelist

We whitelist the domains that appear in the Alexa Top 1M.

Insight a malicious domain automatically generated will not belong to a CDN `r4---sn-a5m7lnes.example.com`.

CDN Whitelist

We whitelist the domains that belong to the most popular CDN networks (e.g., YouTube, Google, etc.) and advertisement services.

Insight an attacker will register a domain with a TLD that does not require clearance.

TLD Whitelist

We whitelist the domains featuring a Top Level Domain that requires authorization by a third party authority before registration (e.g. `.gov`, `.edu`, `.mil`).

Insight How fast is fast?

- 2-3 years ago: TTL < 100.
- Nowadays: 80 < TTL < 300 seconds.

Why? To save money :-) See BH-US 2013 talk[4].

TTL

We filter out all those domains featuring a Time To Live outside these bounds.

[4] https://media.blackhat.com/us-13/US-13-Xu-New-Trends-in-FastFlux-Networks-Slides.pdf

Insight we are looking for DGA-generated domains.

Phoenix's DGA Filter

We filter out domains likely to be generated by humans.

Insight the attacker will register the domain just a few days before the communication will take place.

Whois

We query the Whois server and discard the domains that were registered more than Δ days before the DNS query.

RECAP ON FILTERING

Starting with 50,000 domains:

20,000 **TTL > 300** seconds;

19,000 **not** in the **Alexa Top 1M** list;

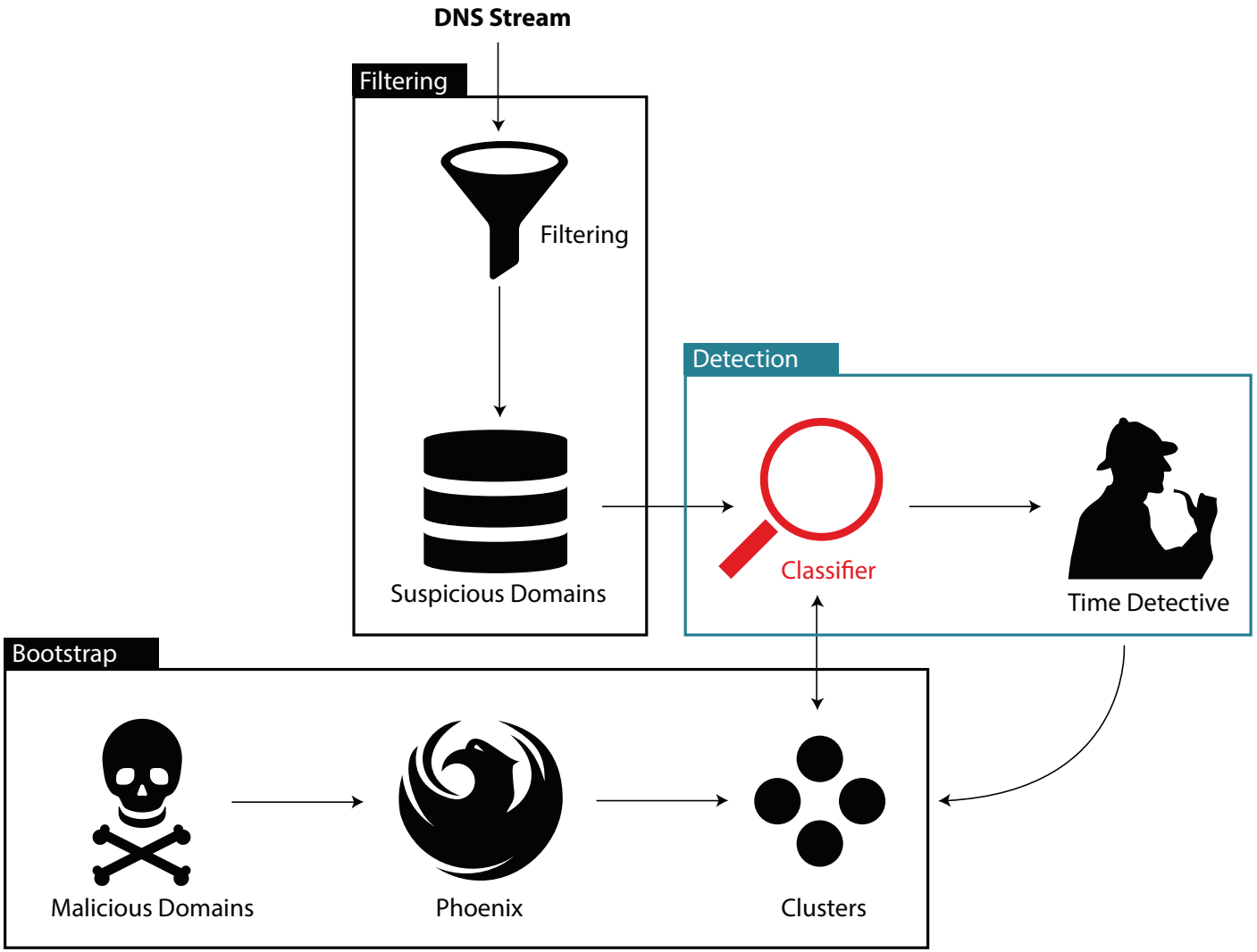
15,000 **not** in the most popular **CDNs**;

800 **likely** to be **DGA generated**;

700 **no** previous **authorization**;

300 **younger than** Δ days $\longleftarrow$ suspicious.

CERBERUS > FILTERING



CLASSIFIER > CLASSIFICATION

576.wap517.net
69.43.161.180

Cluster A

69.43.161.180

379.ns4000wip.com
418.ns4000wip.com
285.ns4000wip.com

Cluster B

69.43.161.180

391.wap517.net
251.wap517.net
340.wap517.net

Cluster C

...

CLASSIFIER > CLASSIFICATION

Cluster A

69.43.161.180

379.ns4000wip.com
418.ns4000wip.com
285.ns4000wip.com

Cluster B

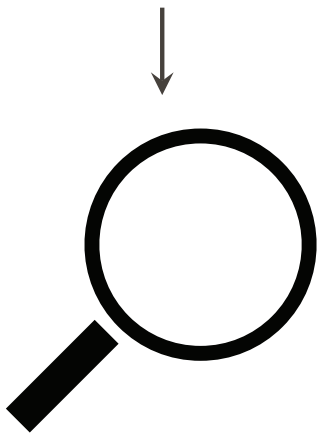
69.43.161.180

391.wap517.net
251.wap517.net
340.wap517.net

Cluster C

...

576.wap517.net
69.43.161.180



CLASSIFIER > CLASSIFICATION

Cluster A

69.43.161.180

379.ns4000wip.com
418.ns4000wip.com
285.ns4000wip.com

Cluster B

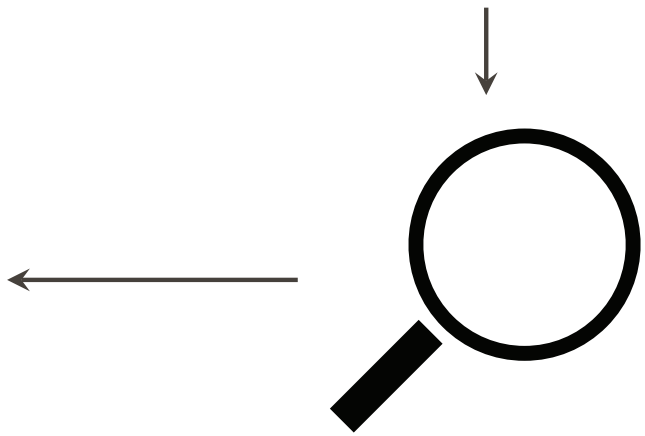
69.43.161.180

391.wap517.net
251.wap517.net
340.wap517.net

Cluster C

...

576.wap517.net
69.43.161.180



CLASSIFIER > CLASSIFICATION

Cluster A

69.43.161.180

379.ns4000wip.com
418.ns4000wip.com
285.ns4000wip.com

Cluster B

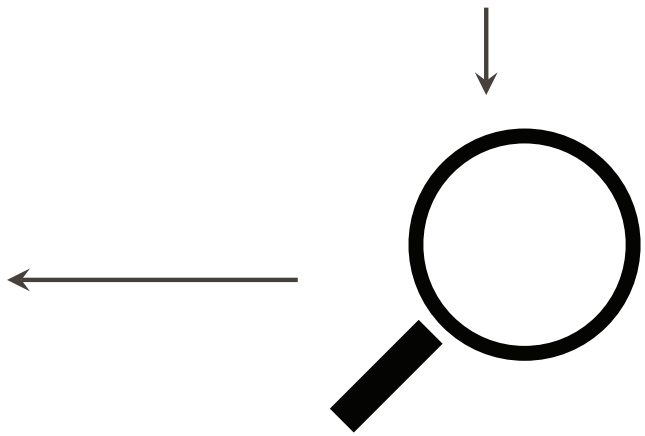
69.43.161.180

391.wap517.net
251.wap517.net
340.wap517.net

Cluster C

...

576.wap517.net
69.43.161.180



CLASSIFIER > CLASSIFICATION

Cluster A

69.43.161.180

379.ns4000wip.com
418.ns4000wip.com
285.ns4000wip.com

Cluster B

69.43.161.180

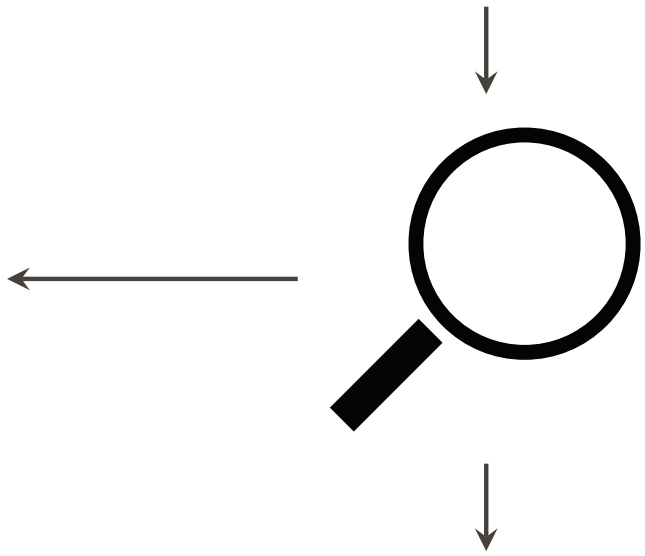
391.wap517.net
251.wap517.net
340.wap517.net

Cluster C

...

576.wap517.net

69.43.161.180



Train the Classifier on A, B

CLASSIFIER > CLASSIFICATION

Cluster A

`69.43.161.180`

`379.ns4000wip.com`
`418.ns4000wip.com`
`285.ns4000wip.com`

Cluster B

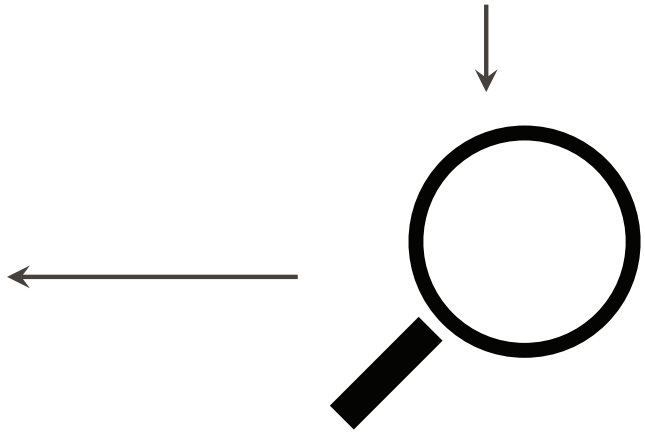
`69.43.161.180`

`391.wap517.net`
`251.wap517.net`
`340.wap517.net`

Cluster C

...

`576.wap517.net`
`69.43.161.180`



Train the Classifier on A, B

Assign `576.wap517.net` to B

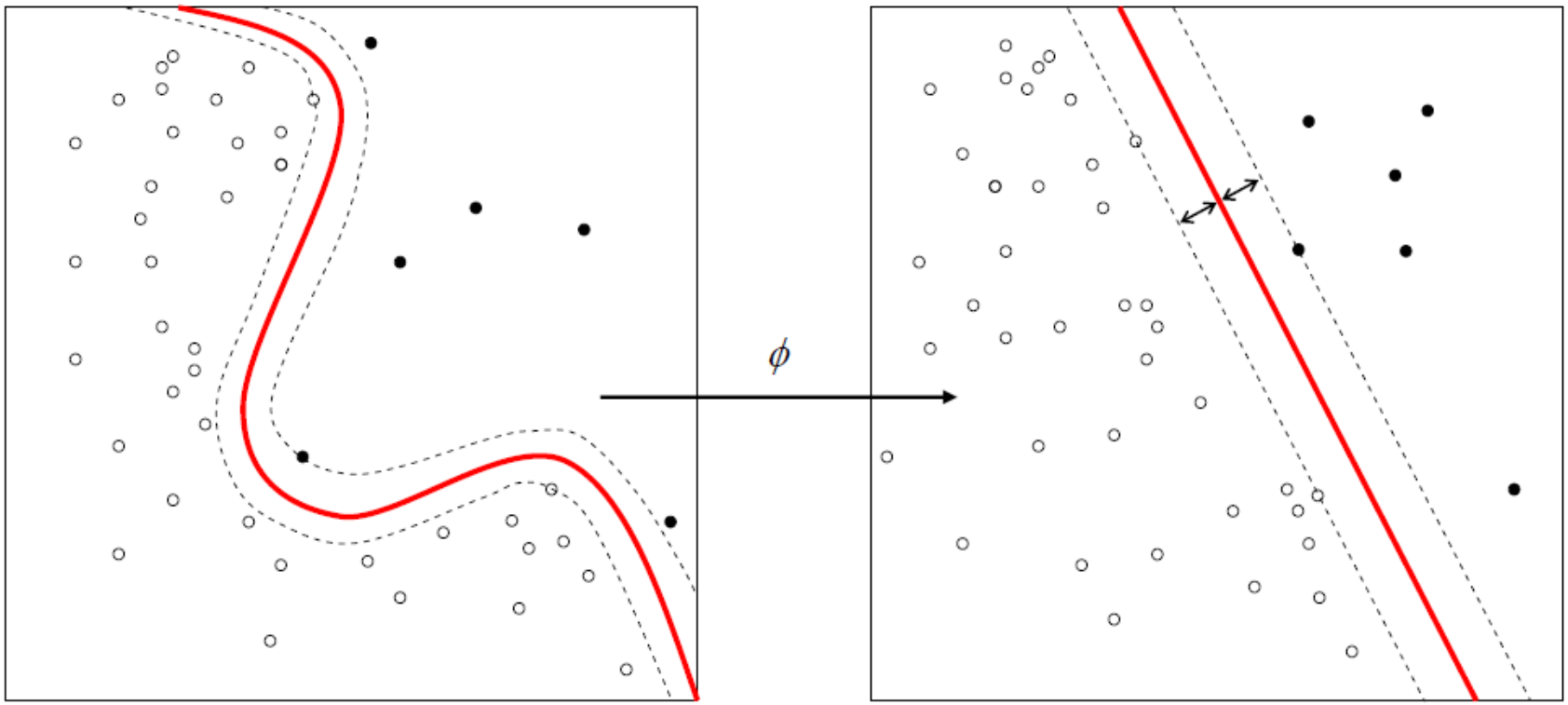
Developed at Royal Holloway in 2002, by Lodhi et al.

	c-a	c-t	a-t	c-r	a-r
ϕ(cat)	λ^2	λ^3	λ^2	0	0
ϕ(car)	λ^2	0	0	λ^3	λ^2

How many substrings of size $k = 2$?

$$\ker(car, cat) = \lambda^4$$
$$\ker(car, car) = \ker(cat, cat) = 2\lambda^4 + \lambda^6$$
$$\ker_n(car, cat) = \frac{\lambda^4}{(2\lambda^4 + \lambda^6)} = \frac{1}{(2 + \lambda^2)} \in [0, 1]$$

SVM: find one hyperplane or a set of them that has the largest distance to the nearest training data point of any class

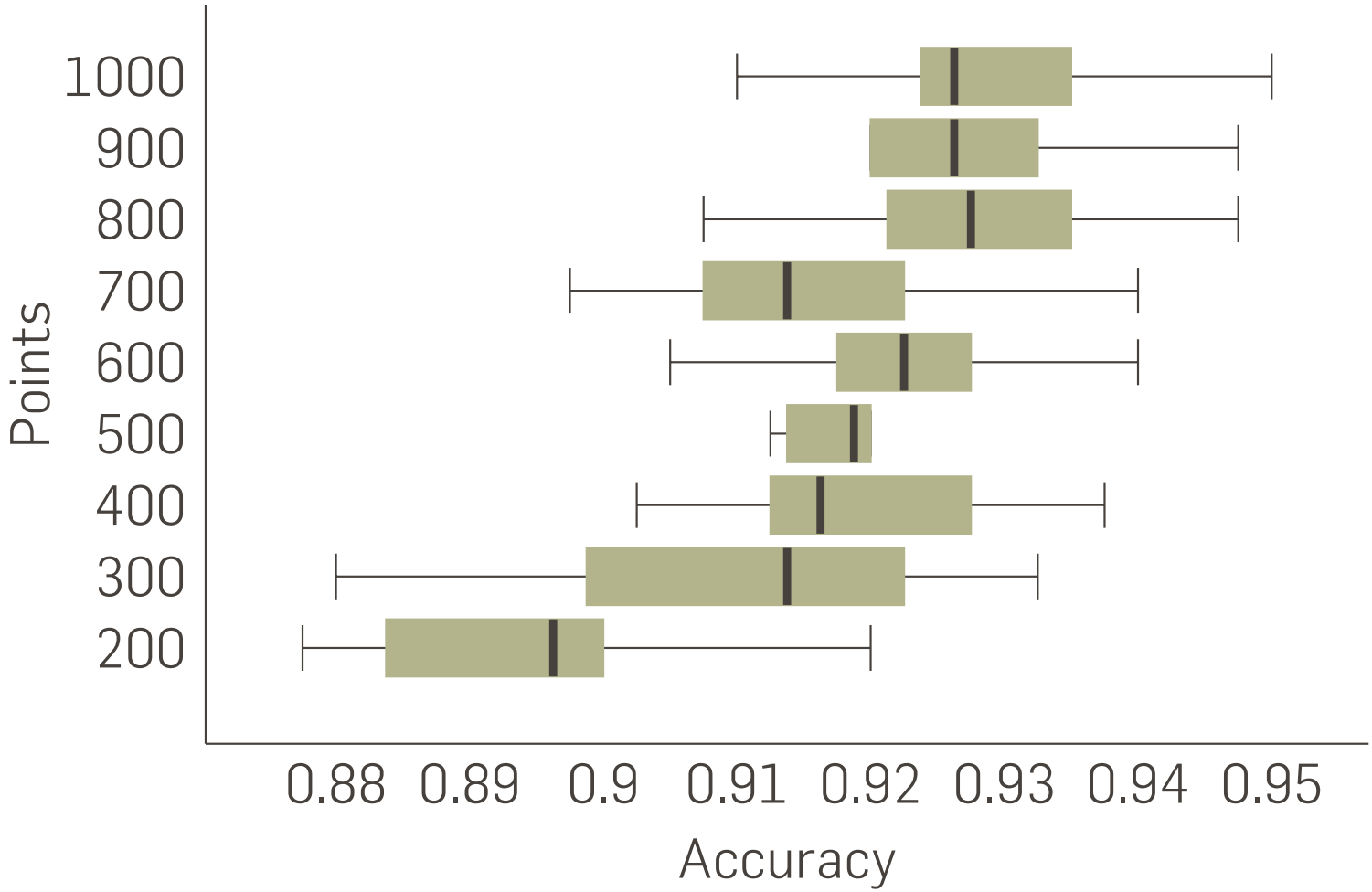


RESULTS

on passive DNS data from

https://farsightsecurity.com/Services/SIE/

RESULTS > CLASSIFIER



Training 1000, Testing 100
Overall Accuracy ≃ 0.95

	a	b	c	d
a	100	0	0	0
b	1	92	6	1
c	2	0	98	0
d	3	0	6	91

a

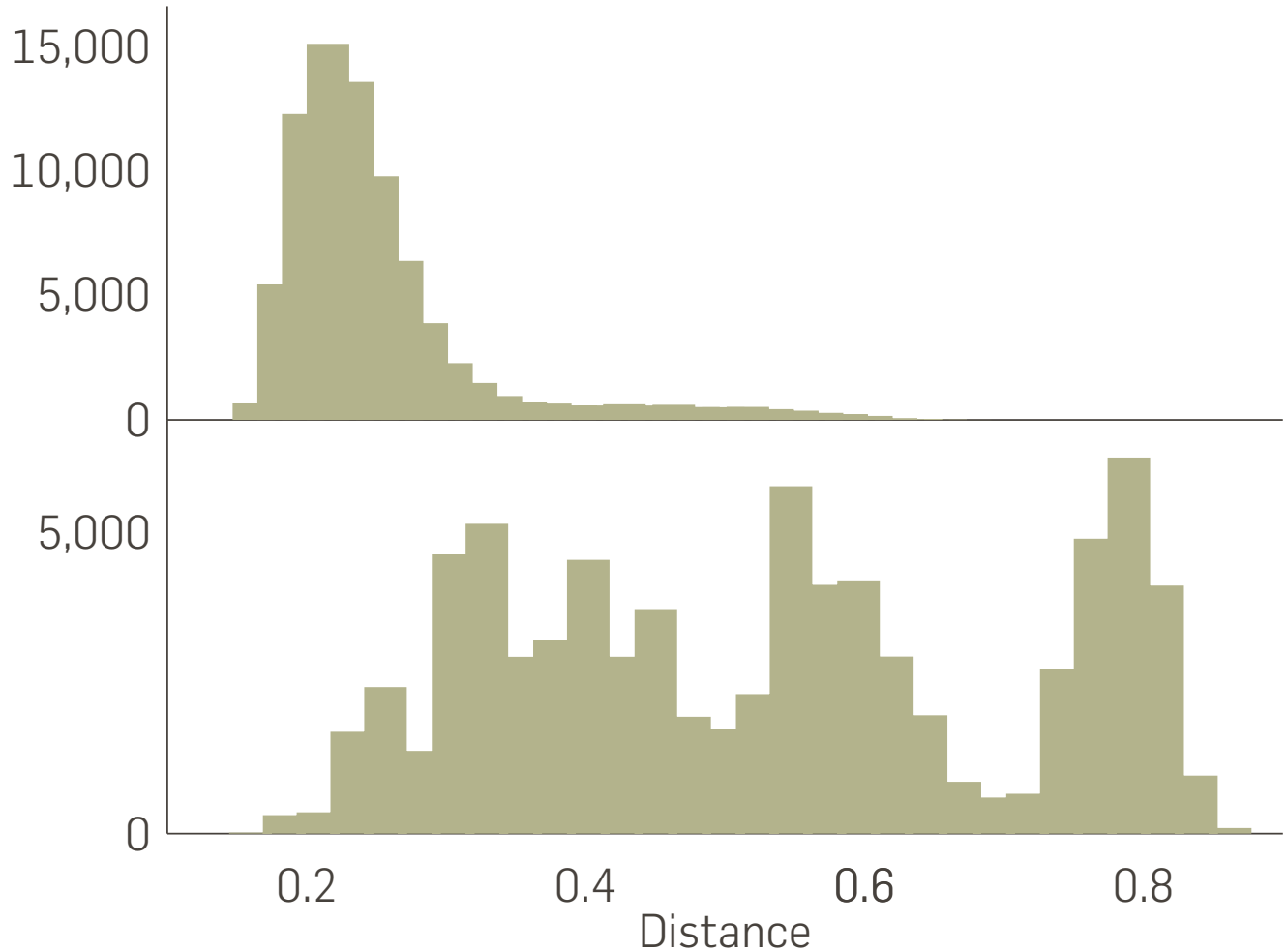
caaa89e...d4ca925b3e2.co.cc
f1e01ac...51b64079d86.co.cc
b

kdnvfyc.biz
wapzzwvpwq.info
c

jhhfghf7.tk
faukiijjj25.tk
d

cvq.com
epu.org

CLASSIFICATION > PAIRWISE DISTANCES



The **Time Detective** discovers new botnets.

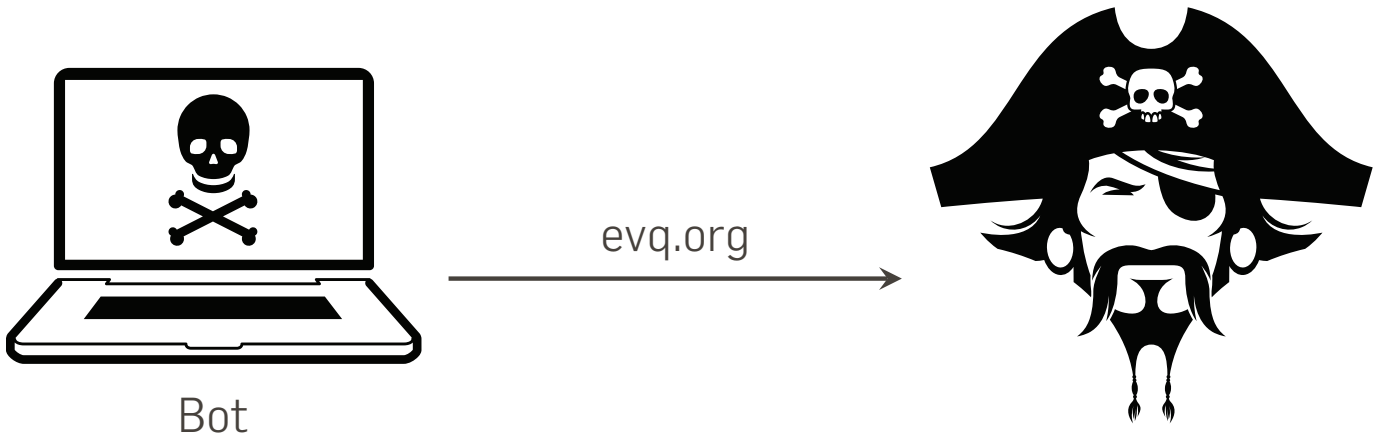
Every Δ the bots **contact** the C&C Server, on a **new domain**.

Bot

Botmaster
`131.175.65.1`

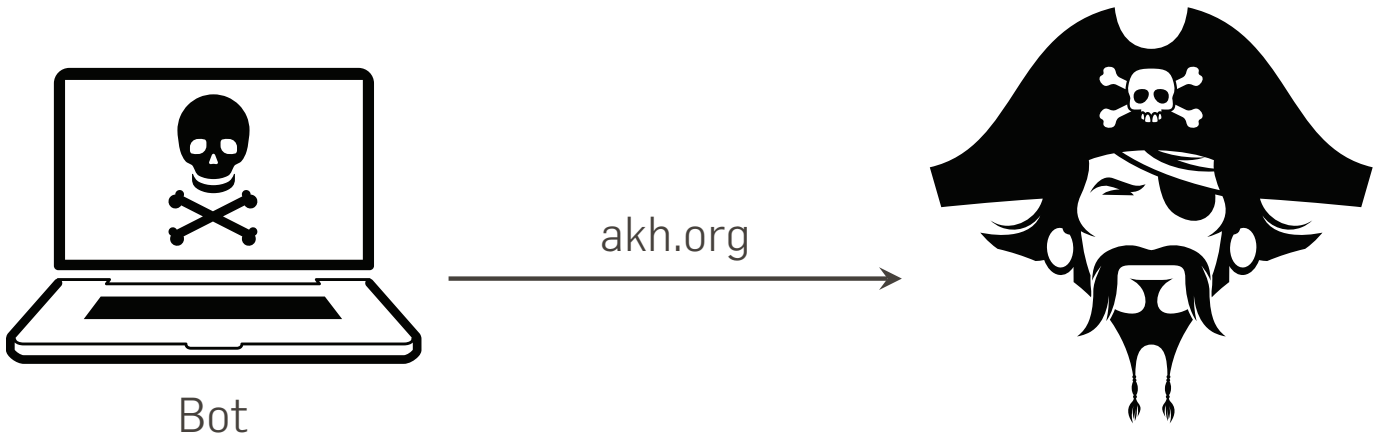
`131.175.65.1: [ ]`

Every Δ the bots **contact** the C&C Server, on a **new domain**.



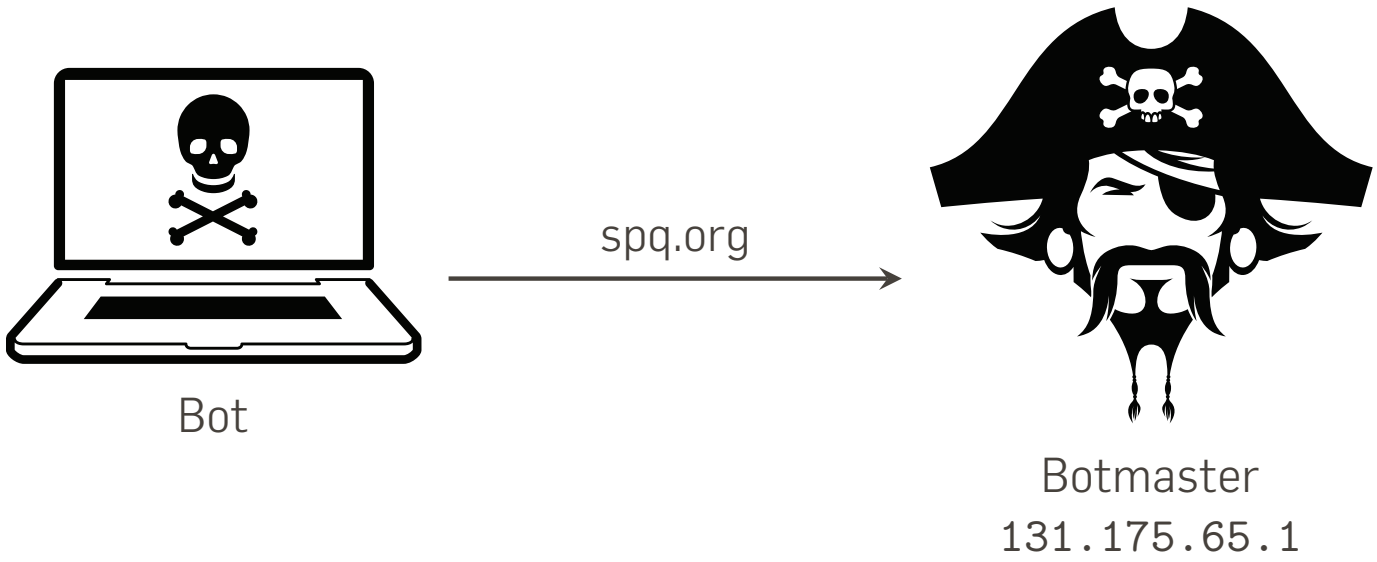
`131.175.65.1`: { evq.org }

Every Δ the bots **contact** the C&C Server, on a **new domain**.



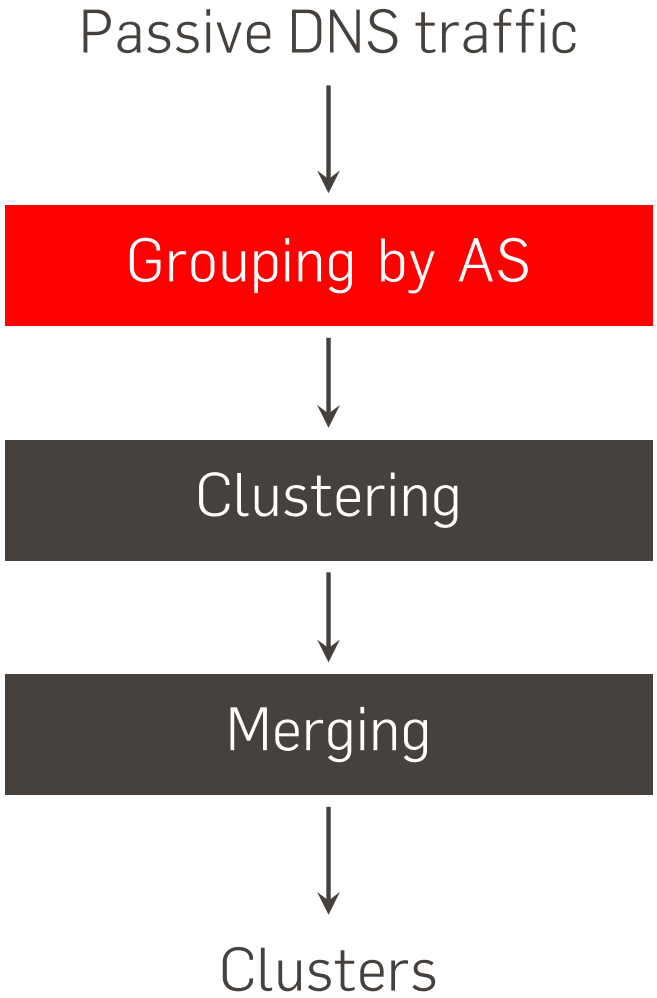
`131.175.65.1`: { evq.org , akh.org }

Every Δ the bots **contact** the C&C Server, on a **new domain**.



`131.175.65.1`: { evq.org , akh.org , spq.org }

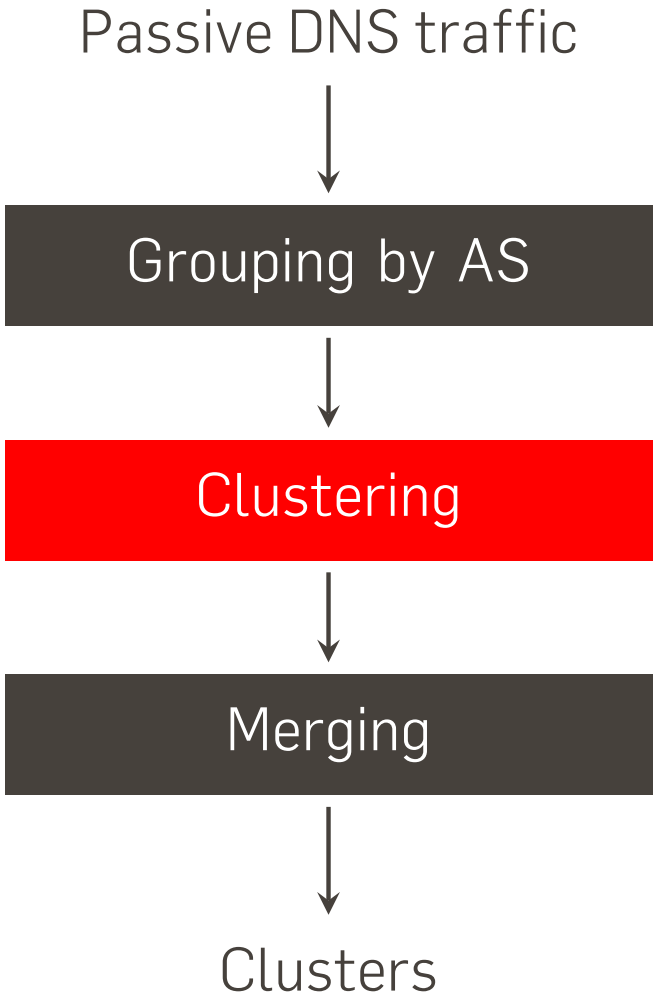
TIME DETECTIVE > STEPS



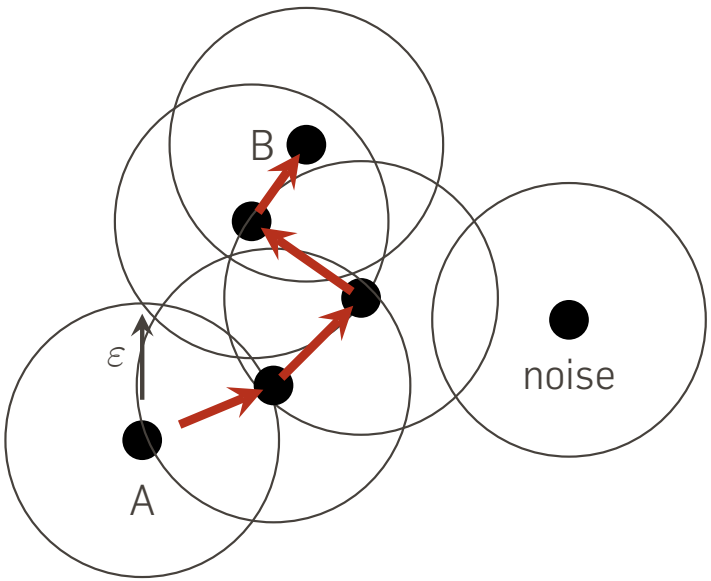
We assume a **lazy attacker** behavior: If (s)he finds an obliging AS, (s)he will buy a few IPs in there.

We group together the domains that point to IPs within the **same AS**.

TIME DETECTIVE > STEPS



DBSCAN



SSK as the distance

automatic tuning:

- $minPts$ domains per cluster,
- ε distance threshold.

$minPts = 7$ domains per cluster

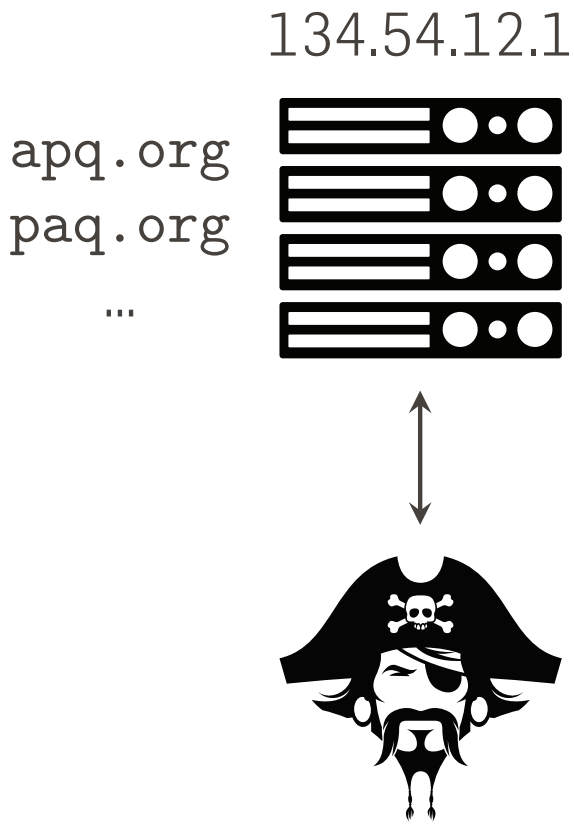
Observation period in days.

Rationale: the bots will contact the C&C server at least **once a day**.

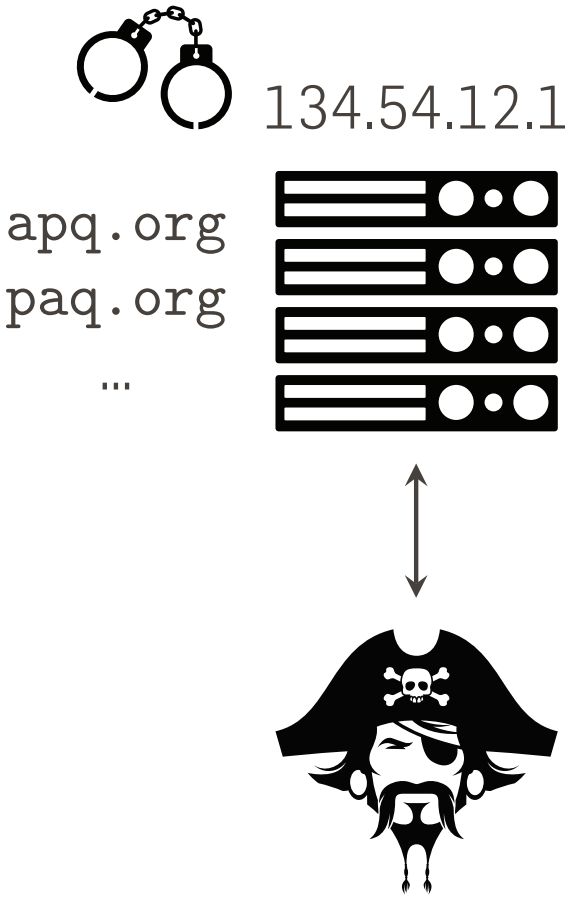
$$\frac{\text{intra-cluster distances}}{\text{inter-cluster distances}} \rightarrow 0 \text{ (minimize)}$$

What if a new cluster is actually a **known botnet** that **migrated** the C&C server somewhere else?

TIME DETECTIVE > MERGING

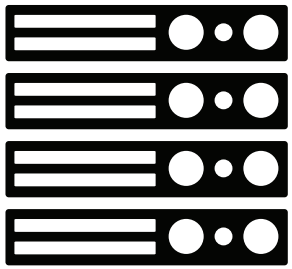


TIME DETECTIVE > MERGING



TIME DETECTIVE > MERGING

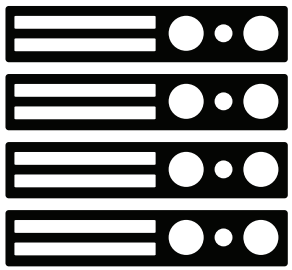
134.54.12.1



What t' h3ck!

TIME DETECTIVE > MERGING

134.54.12.1

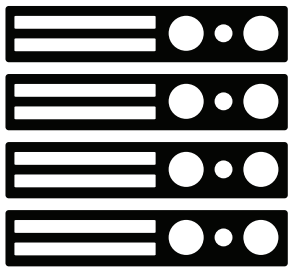


Migration

What t' h3ck!

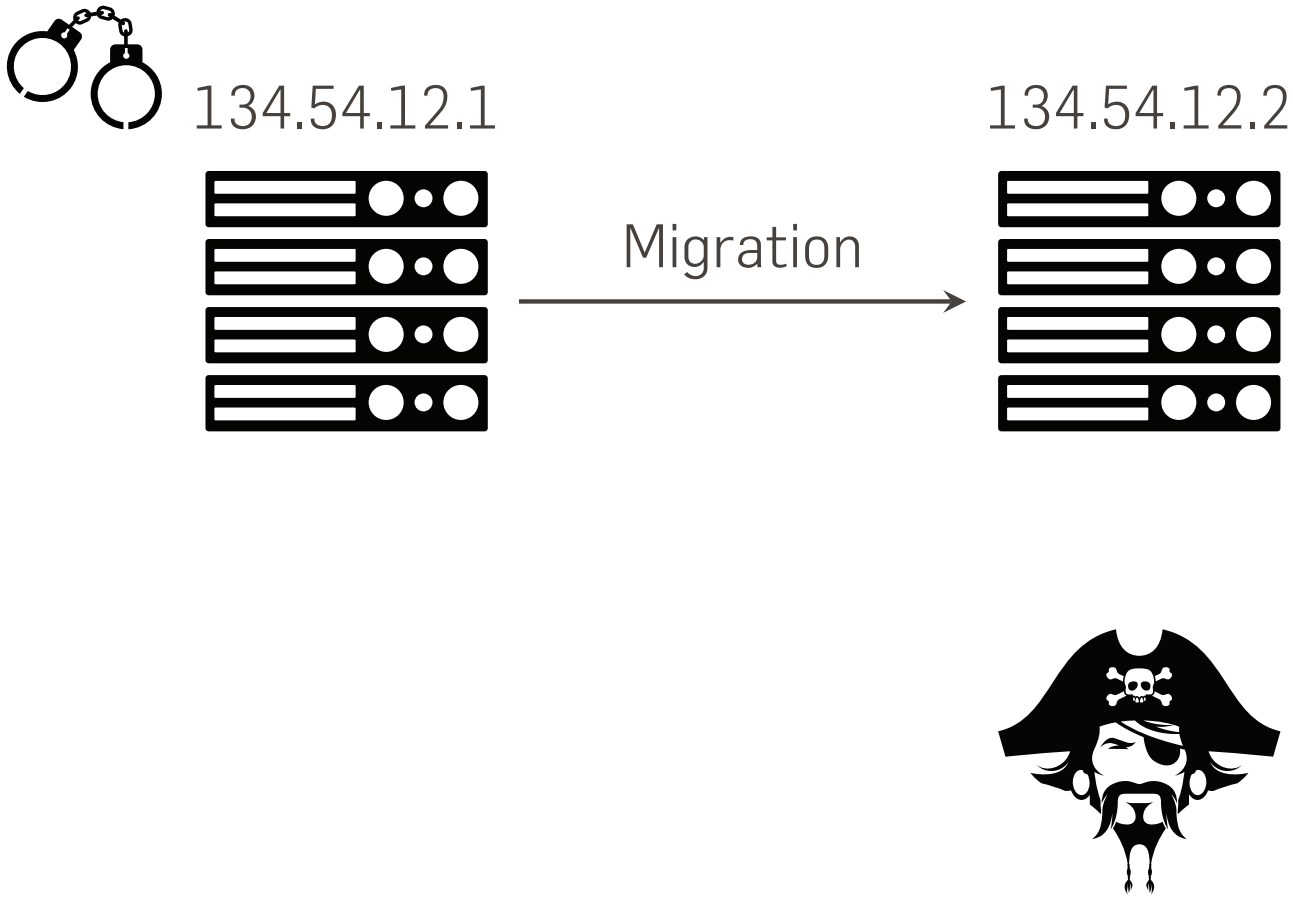
TIME DETECTIVE > MERGING

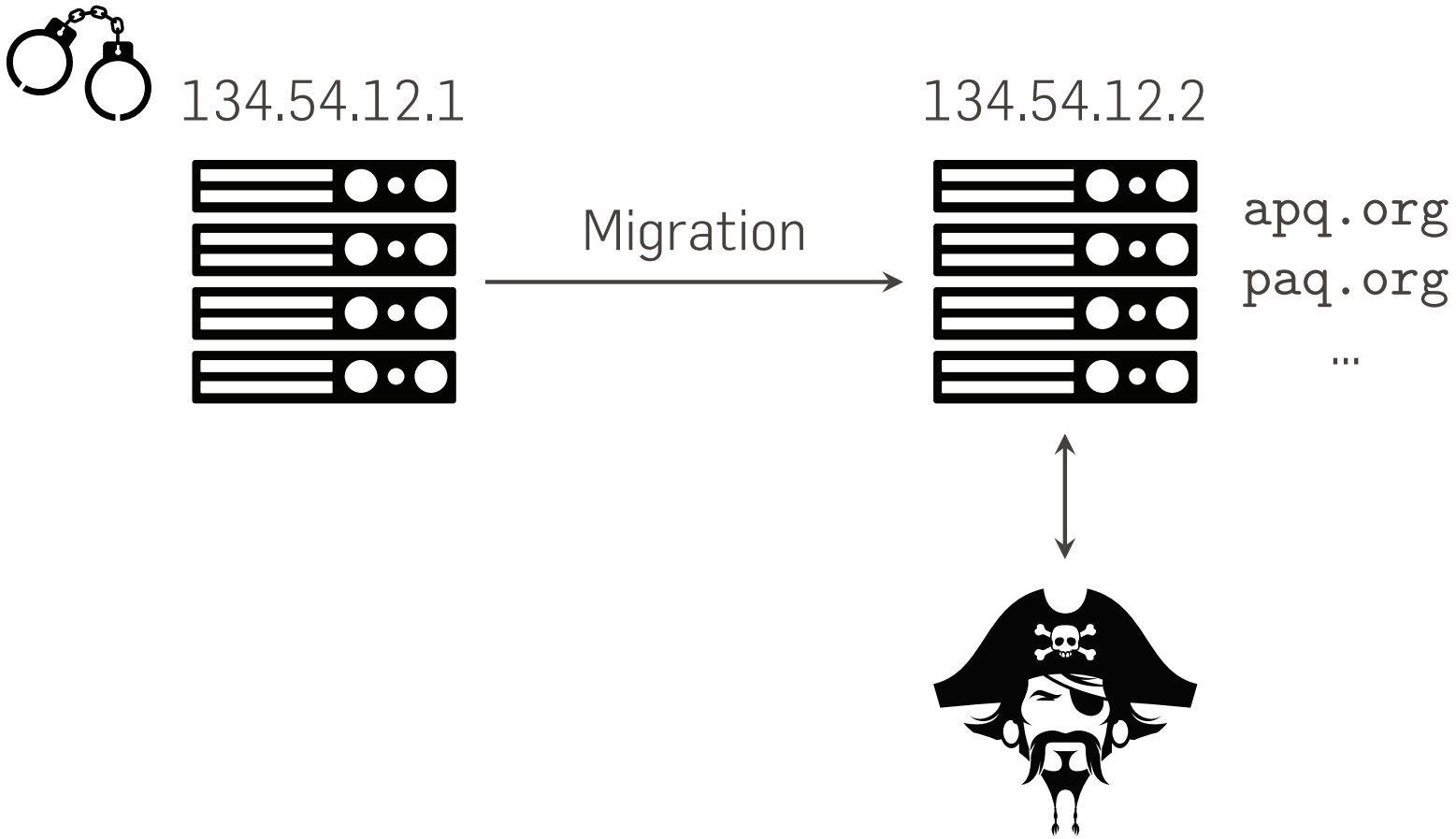
134.54.12.1

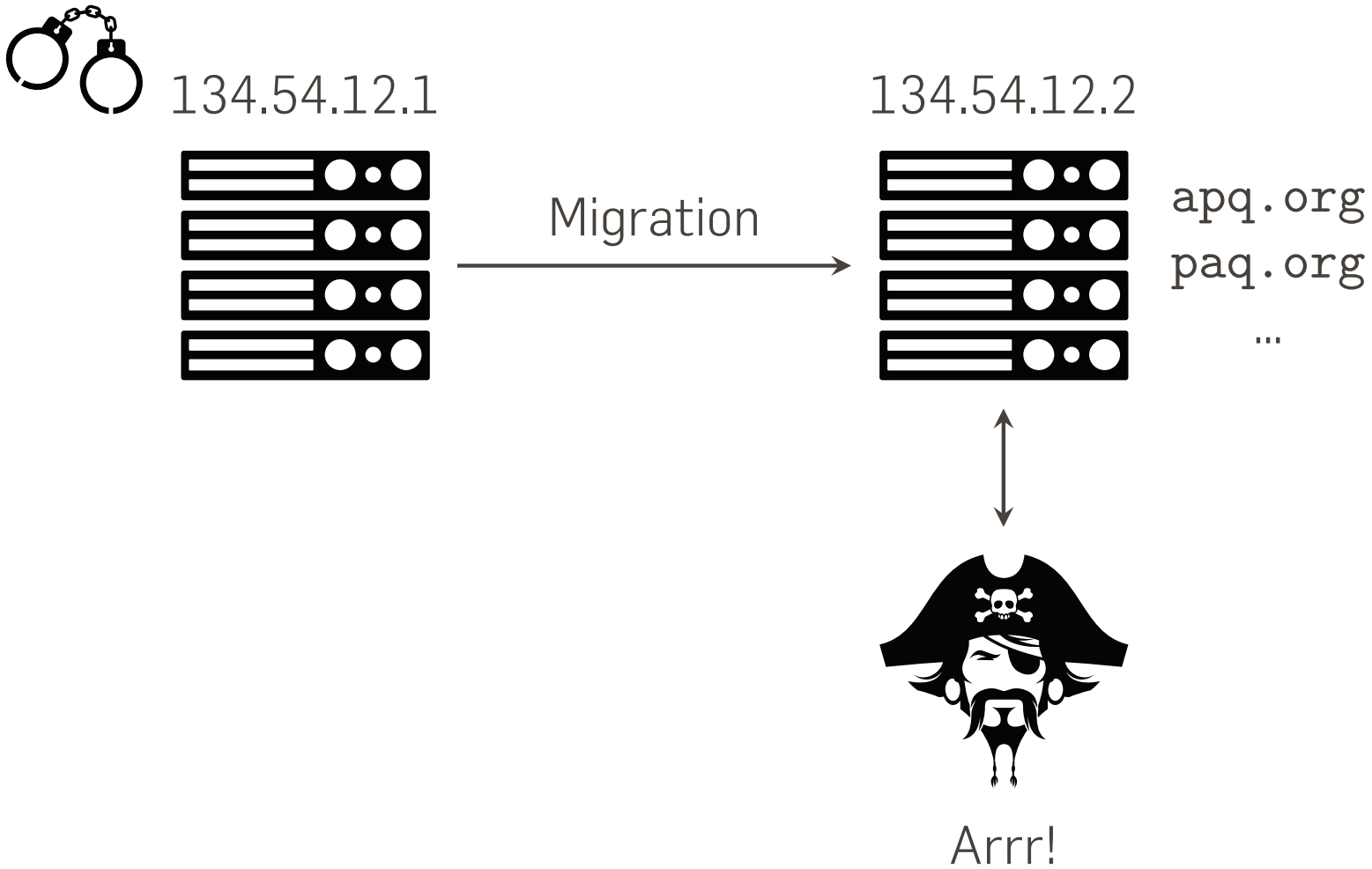


Migration

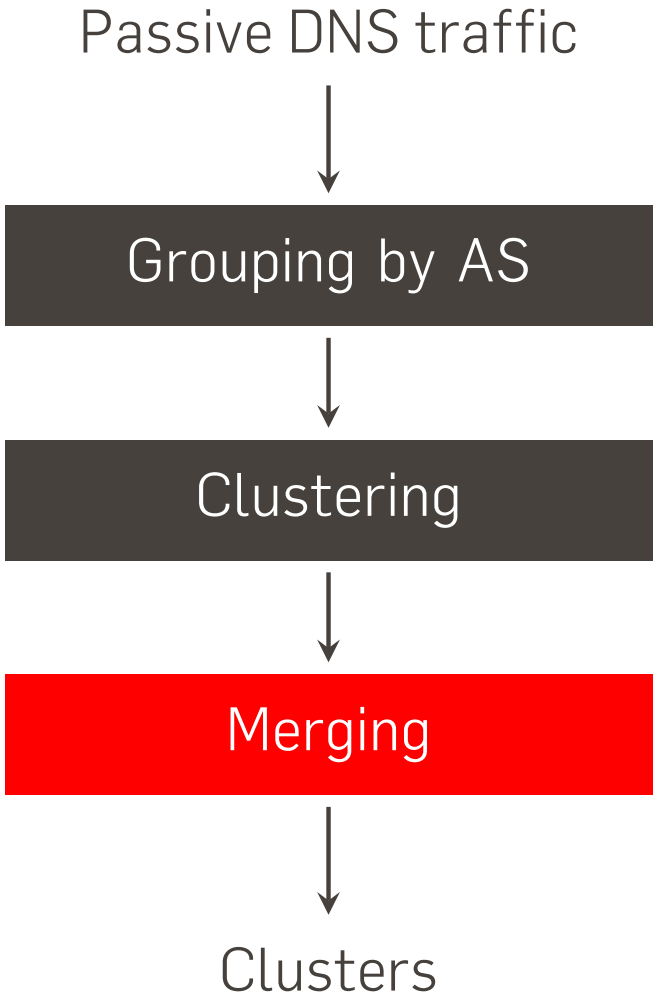
What t' h3ck!


134.54.12.1
134.54.12.2
Migration


134.54.12.1
Migration
134.54.12.2
apq.org
paq.org
...


134.54.12.1
Migration
134.54.12.2
apq.org
paq.org
...
Arrr!

TIME DETECTIVE > STEPS



Suppose you have cluster A and B.

Suppose you have cluster A and B.

$$A = \begin{array}{c} \\ \texttt{dom}_1 \\ \texttt{dom}_2 \\ \vdots \\ \texttt{dom}_m \end{array} \begin{array}{c} \begin{array}{ccc} \texttt{dom}_1 & \cdots & \texttt{dom}_m \end{array} \\ \begin{pmatrix} d_{1,1} & \cdots & d_{1,m} \\ d_{2,1} & \cdots & d_{2,m} \\ \vdots & \ddots & \vdots \\ d_{m,1} & \cdots & d_{m,m} \end{pmatrix} \end{array}$$

Suppose you have cluster A and B.

$$A = \begin{array}{c} \\ \texttt{dom}_1 \\ \texttt{dom}_2 \\ \vdots \\ \texttt{dom}_m \end{array} \begin{array}{c} \begin{array}{ccc} \texttt{dom}_1 & \cdots & \texttt{dom}_m \end{array} \\ \begin{pmatrix} d_{1,1} & \cdots & d_{1,m} \\ d_{2,1} & \cdots & d_{2,m} \\ \vdots & \ddots & \vdots \\ d_{m,1} & \cdots & d_{m,m} \end{pmatrix} \end{array} \quad B = \begin{array}{c} \\ \texttt{dom}_1 \\ \texttt{dom}_2 \\ \vdots \\ \texttt{dom}_n \end{array} \begin{array}{c} \begin{array}{ccc} \texttt{dom}_1 & \cdots & \texttt{dom}_n \end{array} \\ \begin{pmatrix} d_{1,1} & \cdots & d_{1,n} \\ d_{2,1} & \cdots & d_{2,n} \\ \vdots & \ddots & \vdots \\ d_{n,1} & \cdots & d_{n,n} \end{pmatrix} \end{array}$$

Suppose you have cluster A and B.

$$A = \begin{array}{c} \\ \mathtt{dom}_1 \\ \mathtt{dom}_2 \\ \vdots \\ \mathtt{dom}_m \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{dom}_1 & \cdots & \mathtt{dom}_m \end{array} \\ \begin{pmatrix} d_{1,1} & \cdots & d_{1,m} \\ d_{2,1} & \cdots & d_{2,m} \\ \vdots & \ddots & \vdots \\ d_{m,1} & \cdots & d_{m,m} \end{pmatrix} \end{array} \quad B = \begin{array}{c} \\ \mathtt{dom}_1 \\ \mathtt{dom}_2 \\ \vdots \\ \mathtt{dom}_n \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{dom}_1 & \cdots & \mathtt{dom}_n \end{array} \\ \begin{pmatrix} d_{1,1} & \cdots & d_{1,n} \\ d_{2,1} & \cdots & d_{2,n} \\ \vdots & \ddots & \vdots \\ d_{n,1} & \cdots & d_{n,n} \end{pmatrix} \end{array}$$

$$A \sim B = \begin{array}{c} \\ \mathtt{dom}_1 \\ \mathtt{dom}_2 \\ \vdots \\ \mathtt{dom}_m \end{array} \begin{array}{c} \begin{array}{cccc} \mathtt{dom}_1 & \mathtt{dom}_2 & \cdots & \mathtt{dom}_n \end{array} \\ \begin{pmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,n} \\ d_{2,1} & d_{2,2} & \cdots & d_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m,1} & d_{m,2} & \cdots & d_{m,n} \end{pmatrix} \end{array}$$

TIME DETECTIVE > WELCH TEST

Stats to the rescue!

Stats to the rescue!

$$A = \begin{array}{c} \\ \mathtt{dom}_1 \\ \mathtt{dom}_2 \\ \vdots \\ \mathtt{dom}_m \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{dom}_1 & \cdots & \mathtt{dom}_m \end{array} \\ \begin{pmatrix} d_{1,1} & \cdots & d_{1,m} \\ d_{2,1} & \cdots & d_{2,m} \\ \vdots & \ddots & \vdots \\ d_{m,1} & \cdots & d_{m,m} \end{pmatrix} \end{array} \qquad A \sim B = \begin{array}{c} \\ \mathtt{dom}_1 \\ \mathtt{dom}_2 \\ \vdots \\ \mathtt{dom}_m \end{array} \begin{array}{c} \begin{array}{cccc} \mathtt{dom}_1 & \mathtt{dom}_2 & \cdots & \mathtt{dom}_n \end{array} \\ \begin{pmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,n} \\ d_{2,1} & d_{2,2} & \cdots & d_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m,1} & d_{m,2} & \cdots & d_{m,n} \end{pmatrix} \end{array}$$

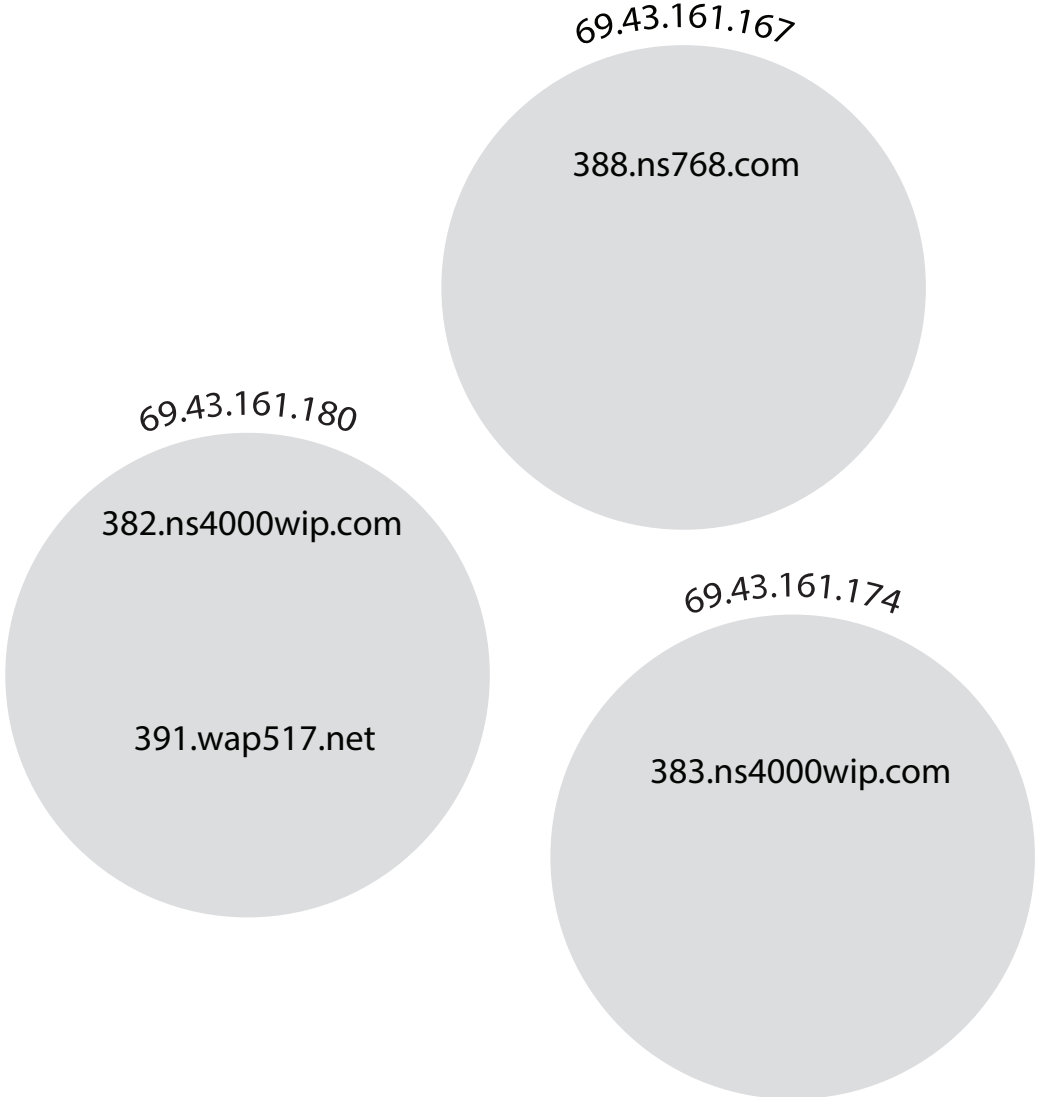
Stats to the rescue!

$$A = \begin{array}{c} \\ \mathtt{dom}_1 \\ \mathtt{dom}_2 \\ \vdots \\ \mathtt{dom}_m \end{array} \begin{array}{c} \begin{array}{ccc} \mathtt{dom}_1 & \cdots & \mathtt{dom}_m \end{array} \\ \begin{pmatrix} d_{1,1} & \cdots & d_{1,m} \\ d_{2,1} & \cdots & d_{2,m} \\ \vdots & \ddots & \vdots \\ d_{m,1} & \cdots & d_{m,m} \end{pmatrix} \end{array} \quad A \sim B = \begin{array}{c} \\ \mathtt{dom}_1 \\ \mathtt{dom}_2 \\ \vdots \\ \mathtt{dom}_m \end{array} \begin{array}{c} \begin{array}{cccc} \mathtt{dom}_1 & \mathtt{dom}_2 & \cdots & \mathtt{dom}_n \end{array} \\ \begin{pmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,n} \\ d_{2,1} & d_{2,2} & \cdots & d_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m,1} & d_{m,2} & \cdots & d_{m,n} \end{pmatrix} \end{array}$$

Welch test: do A and $A \sim B$ have different intra-cluster distance distributions?

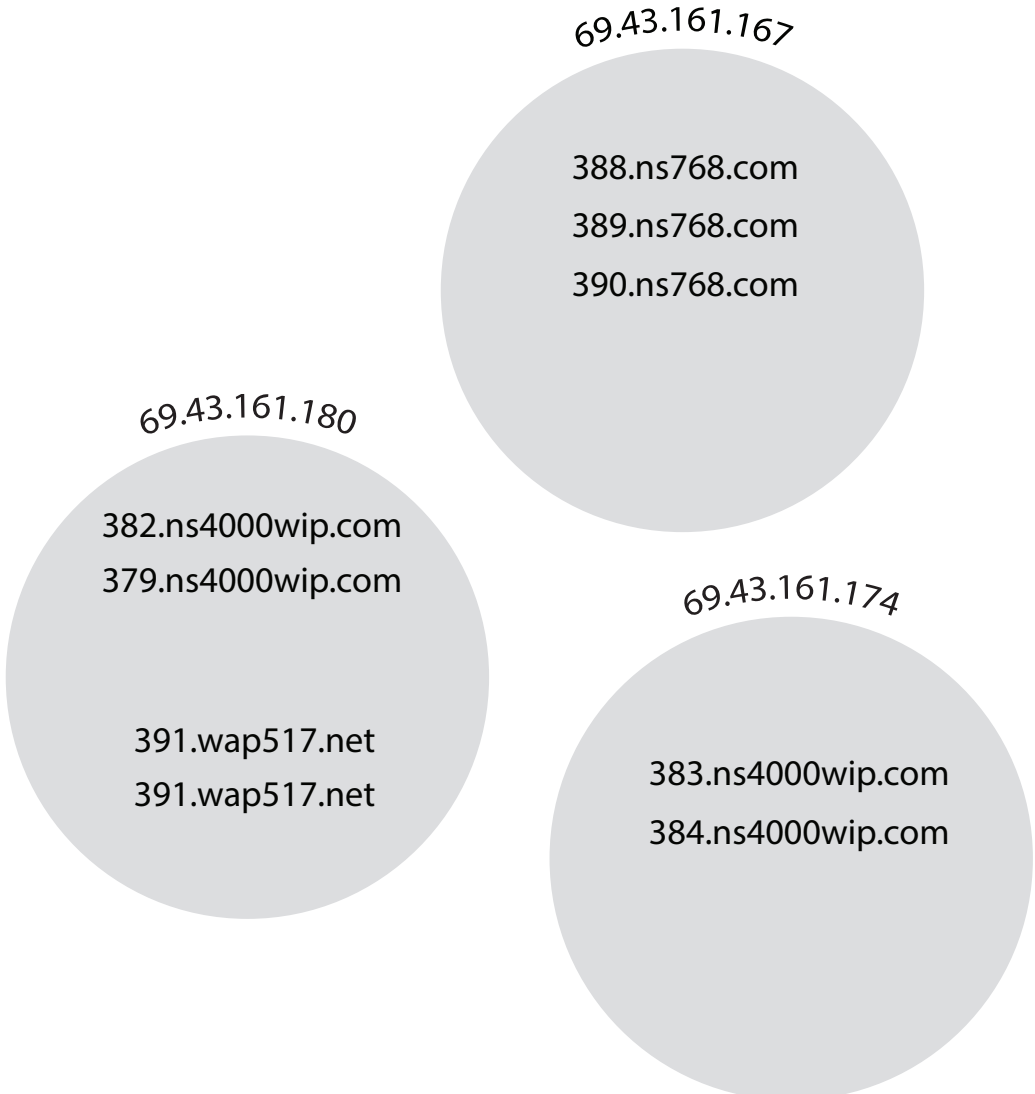
TIME DETECTIVE > EXAMPLE

Day 1



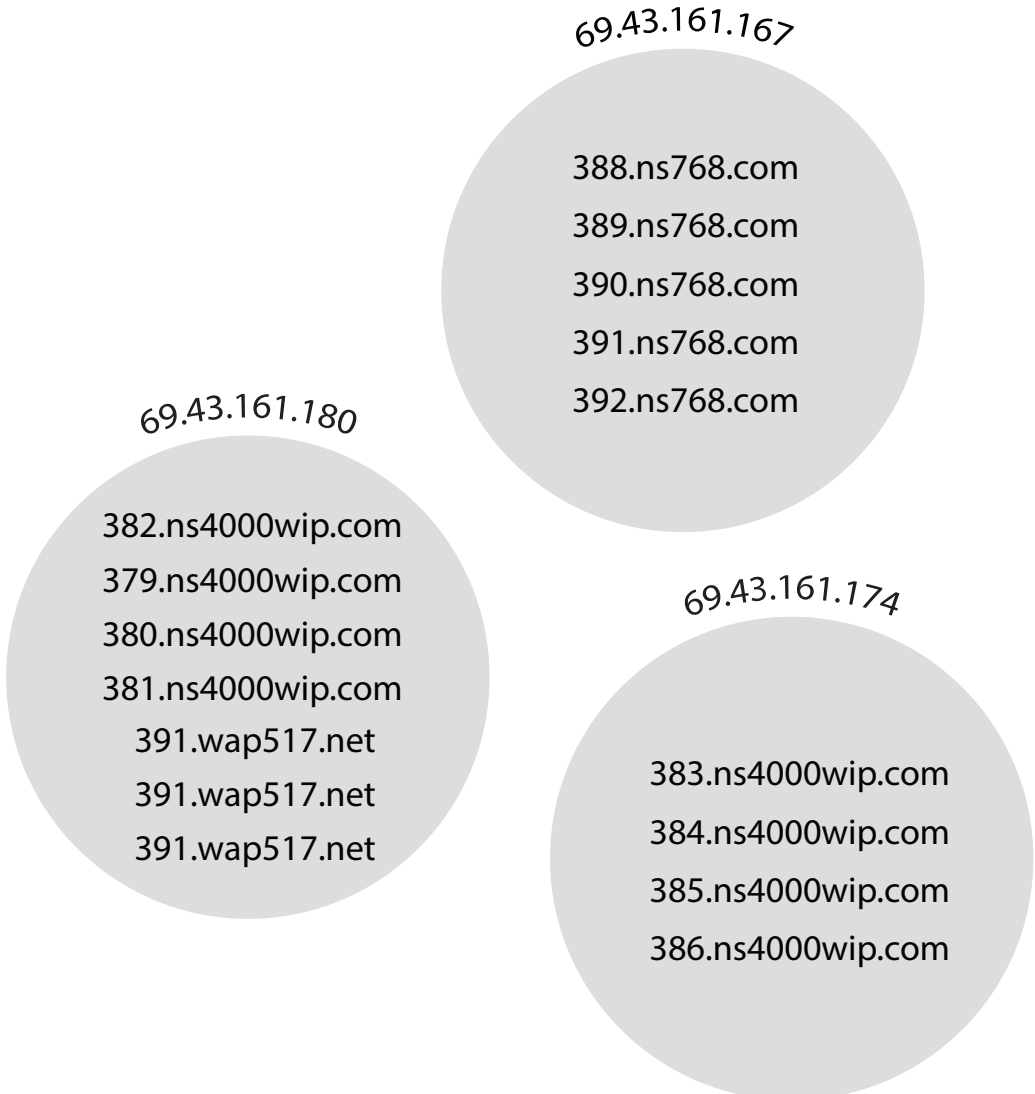
TIME DETECTIVE > EXAMPLE

Day 2



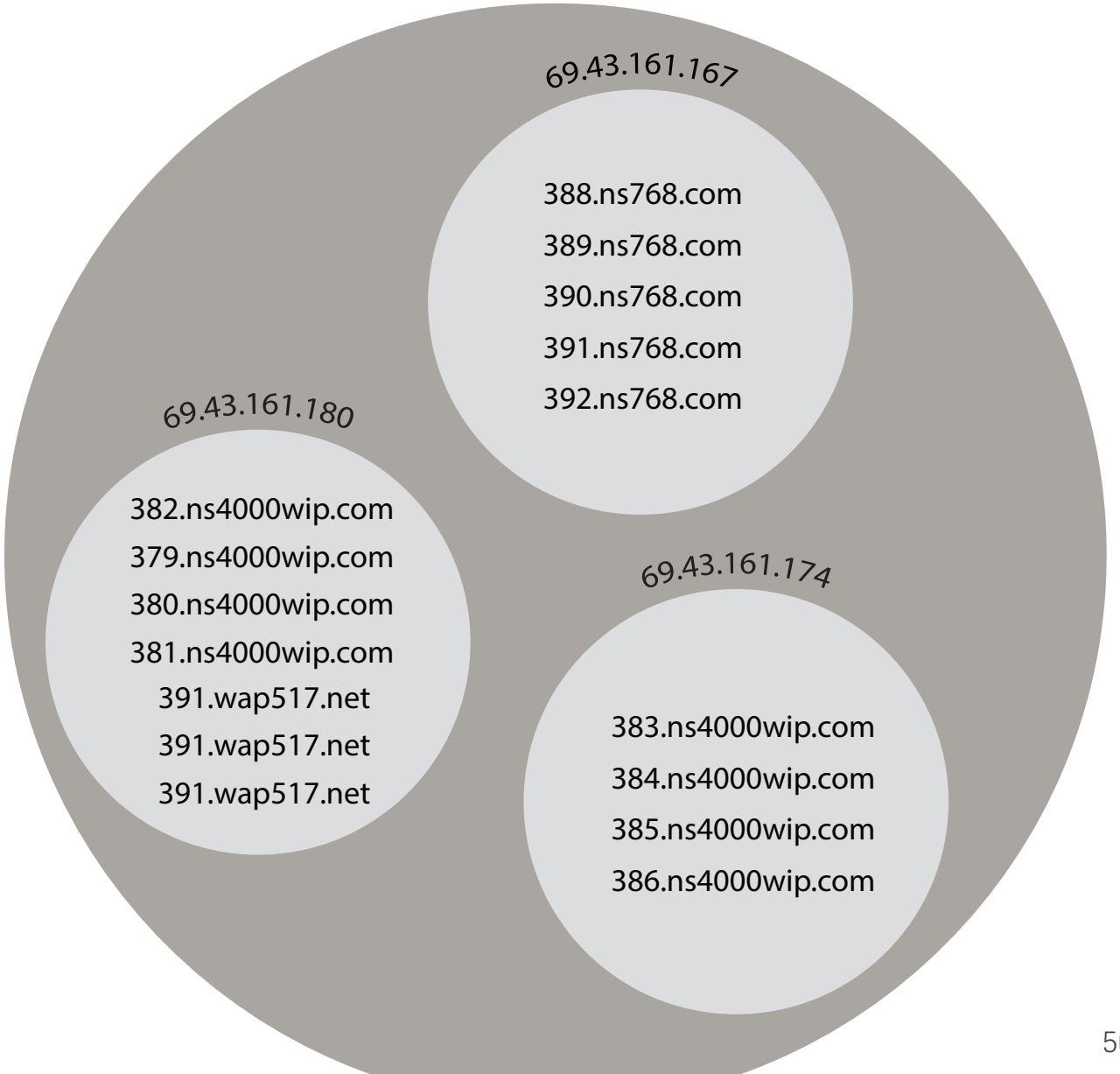
TIME DETECTIVE > EXAMPLE

Day 7



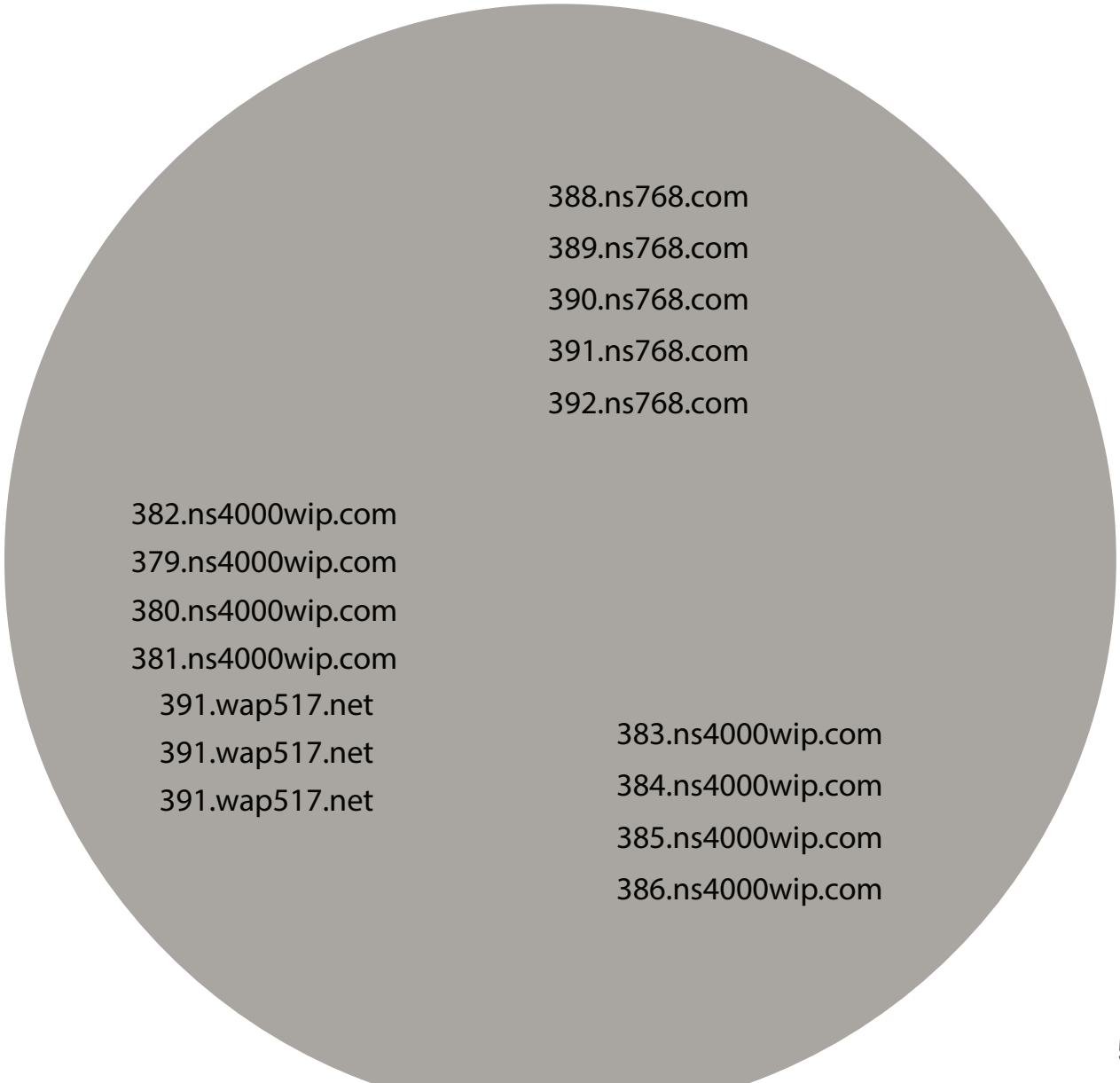
TIME DETECTIVE > EXAMPLE

AS 22489



TIME DETECTIVE > EXAMPLE

Merge



TIME DETECTIVE > EXAMPLE

Cluster

New clusters produced

Cluster 1	Cluster 2	Cluster 3
388.ns768.com	382.ns4000wip.com	391.wap517.net
389.ns768.com	379.ns4000wip.com	391.wap517.net
390.ns768.com	380.ns4000wip.com	391.wap517.net
391.ns768.com	381.ns4000wip.com	
392.ns768.com	383.ns4000wip.com	
	384.ns4000wip.com	
	385.ns4000wip.com	
	386.ns4000wip.com	

RESULTS

on passive DNS data from
https://farsightsecurity.com/Services/SIE/

187 domains classified as malicious and **labeled**.

Labeled 07e21

Botnet:	Conficker
Domains:	`hhdboqazof.biz`
	`poxqmrfj.biz`
	`hcsddszzzc.ws`
	`tnoucgrje.biz`
	`gwizoxej.biz`
	`jnmuoiki.biz`

3,576 domains were considered **suspicious** by Cerberus and **stored**, together with their IP address.

Then we ran the clustering routine to **discover new botnets**.

TIME DETECTIVE > CLUSTERING

Botnet	AS	IPs	Size
Sality	15456	62.116.181.25	26
Palevo	53665	199.59.243.118	40
Jadtre*	22489	69.43.161.180 69.43.161.174	173
Jadtre**	22489	69.43.161.180	37
Jadtre***	22489	69.43.161.167	47
Hiloti	22489	69.43.161.167	24
Palevo	47846	82.98.86.171 82.98.86.176 82.98.86.175	142
Jusabli	30069	69.58.188.49	73
Generic Trojan	12306	82.98.86.169 82.98.86.162 82.98.86.178 82.98.86.163	57

Cluster	IP	Sample Domains
Jadtre*	69.43.161.180	379.ns4000wip.com
	69.43.161.174	418.ns4000wip.com
		285.ns4000wip.com
Jadtre**	69.43.161.180	391.wap517.net
		251.wap517.net
		340.wap517.net
Jadtre***	69.43.161.167	388.ns768.com
		353.ns768.com
		296.ns768.com

TIME DETECTIVE > MERGING

Cluster a (Old)	
IPs:	176.74.76.175
	208.87.35.107
Domains	cvq.com
	epu.org
	bwn.org
	lxx.net

Cluster b (New)	
IPs:	82.98.86.171
	82.98.86.176
	82.98.86.175
	82.98.86.167
	82.98.86.168
	82.98.86.165
Domains	knw.info
	rrg.info
	nhy.org
	ydt.info

Cluster a (Old)	
IPs:	`176.74.76.175`
	`208.87.35.107`
Domains	`cvq.com`
	`epu.org`
	`bwn.org`
	`lxx.net`

Cluster b (New)	
IPs:	`82.98.86.171`
	`82.98.86.176`
	`82.98.86.175`
	`82.98.86.167`
	`82.98.86.168`
	`82.98.86.165`
Domains	`knw.info`
	`rrg.info`
	`nhy.org`
	`ydt.info`

Both belonging to the **Palevo botnet**.

- **187** malicious domains **detected and labeled**

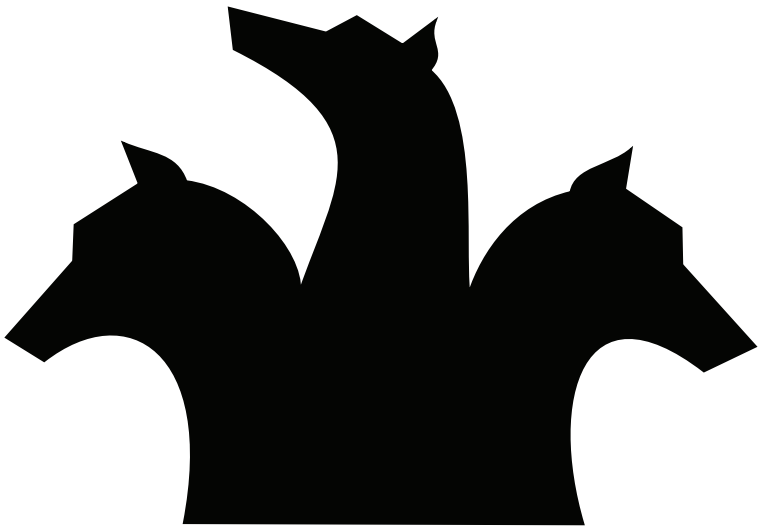
- **187** malicious domains **detected and labeled**
- **3,576 suspicious** domains collected

- **187** malicious domains **detected and labeled**
- **3,576 suspicious** domains collected
- **47 clusters** of DGA-generated domains **discovered**

- **187** malicious domains **detected and labeled**
- **3,576 suspicious** domains collected
- **47 clusters** of DGA-generated domains **discovered**
- **319** new domains **detected in the next 24 hours**

CONCLUSIONS & FUTURE WORK

CONCLUSIONS



CERBERUS

- discovers and characterizes unknown DGA-based activity,
- unsupervised,
- easy to deploy,
- privacy preserving.

FUTURE WORK

FUTURE WORK

this-is-an-easy-way-to-evade-the-linguistic-filter.com

FUTURE WORK

Release **Cerberus** as a web service. Hopefully!

THANK YOU

federico.maggi@polimi.it
@phretor