



# When threat hunting fails

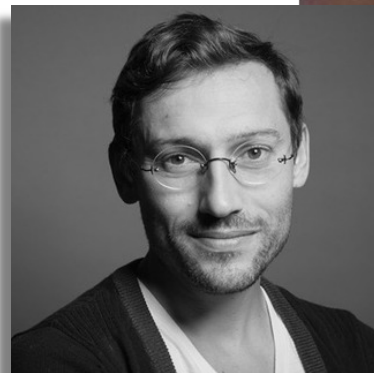
Identifying malvertising domains using lexical clustering

Tucson, January 9th, 2018

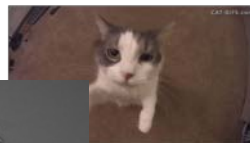
# Authors



Matt Foley



David Rodriguez



kitty



Dhia Mahjoub

# Agenda

Background

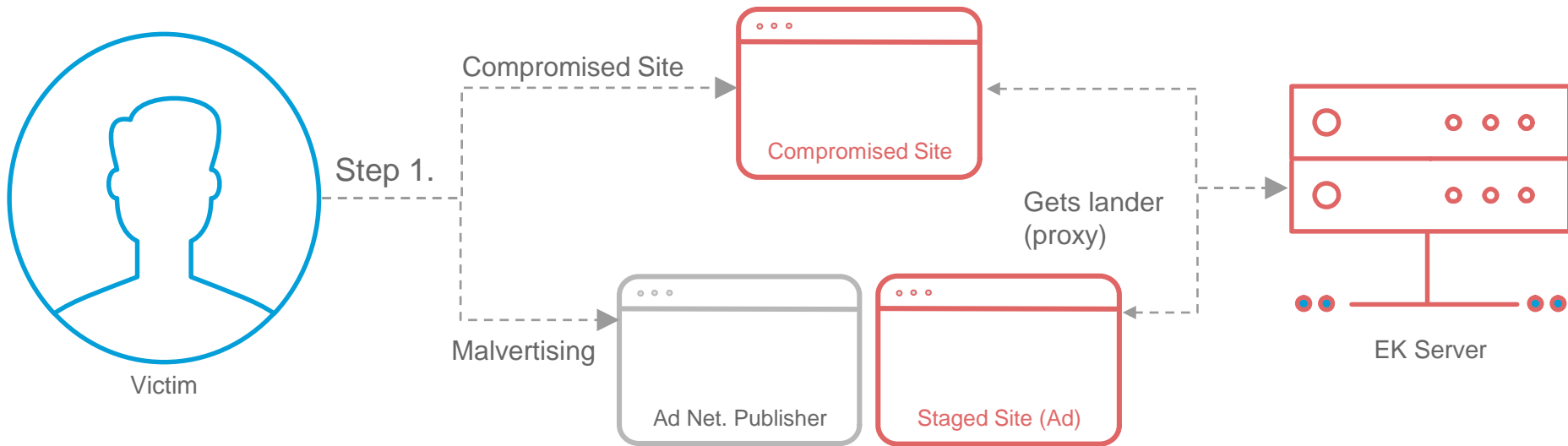
Ad Network Profiling and Filtering

Lexical Clustering

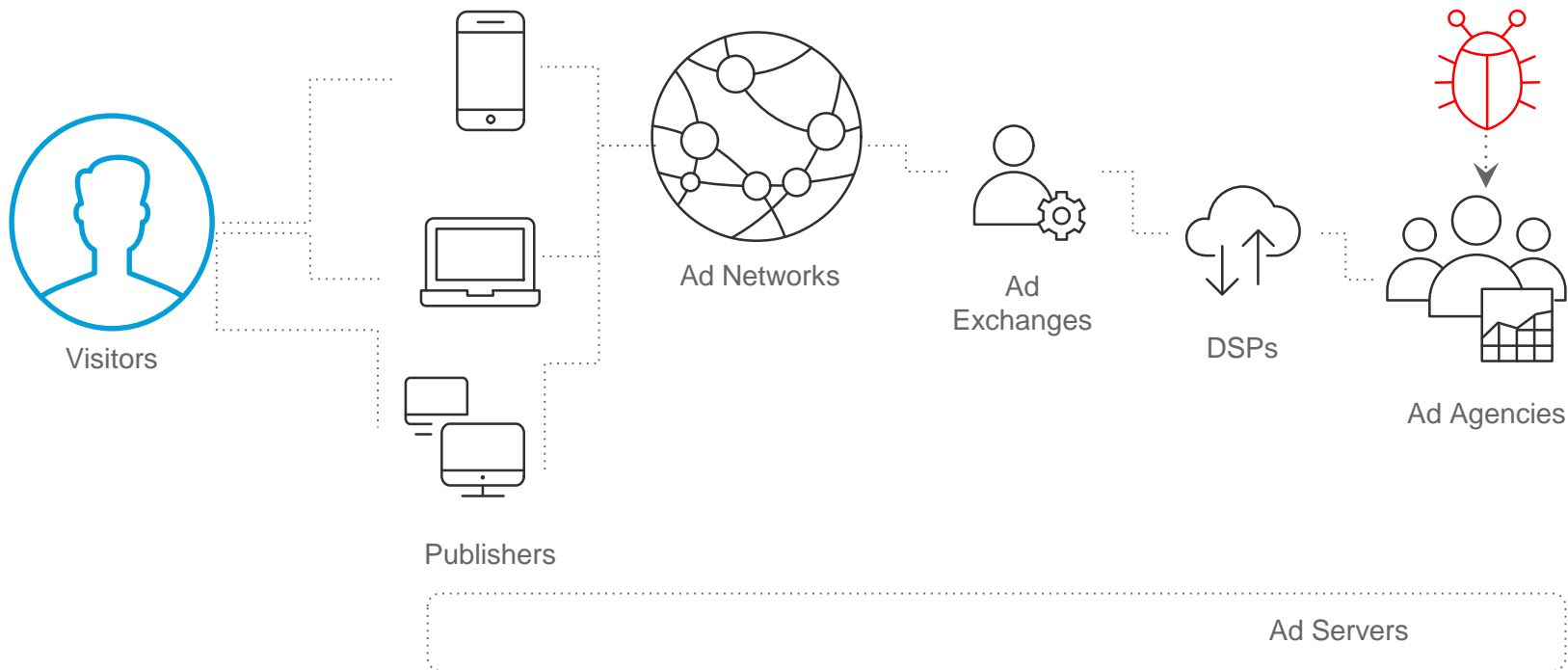
Hosting space and top talkers

# Background

# Exploit Kits



# What is Malvertising





Your security matters

Google recommends using Chrome, a fast and secure browser. Try it?

NO, NOT INTERESTED

YES

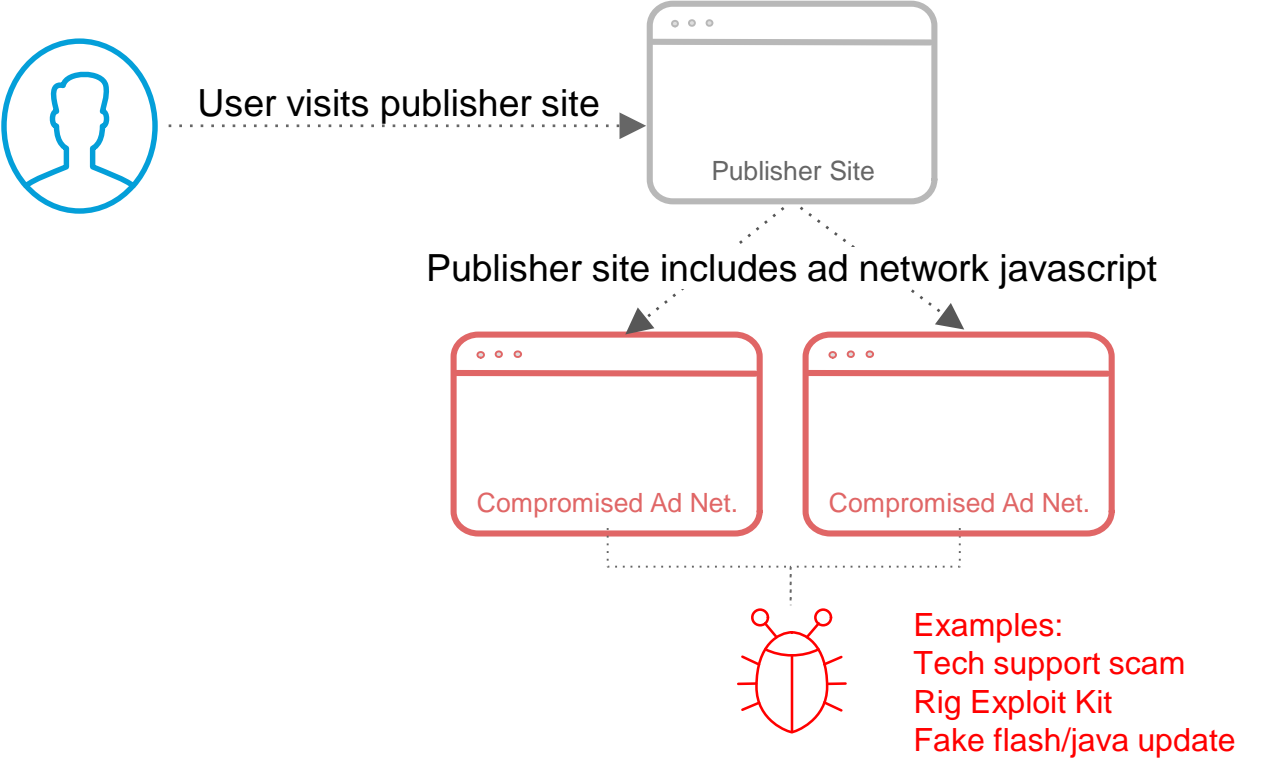
# Google

Google Search

I'm Feeling Lucky



# Ad Campaign Flow

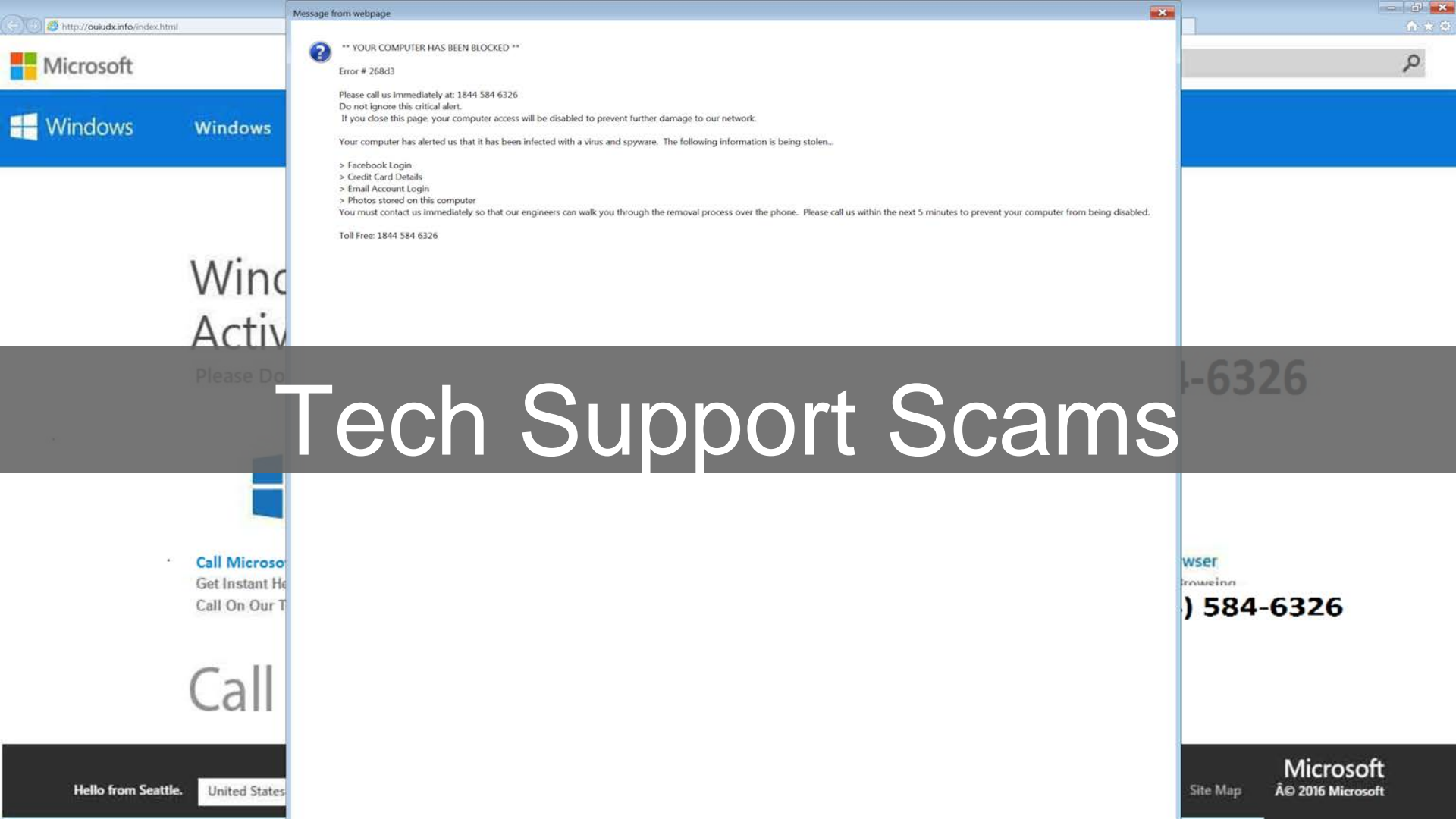




# Exploit Kits

Filter: http.request

Date/Time	Dst	port	Host	Info
2017-06-20 14:28:04	80.77.82.41	80	observice.info	GET /banners/uaps HTTP/1.1
2017-06-20 14:28:04	80.85.158.121	80	80.85.158.121	POST /?Reg&man=1617&van=xHrQMrPYbR3FFYbfKP_EUKFEMUvWA0SkwYyZhazVF5qxFDTGpbH
2017-06-20 14:28:06	80.85.158.121	80	80.85.158.121	GET /?e&flight=4546&man=4796&van=xX3QMvWyBRXQCJ3EKv_ct6NGMVHRGUCL2YydmrHV
2017-06-20 14:28:07	80.85.158.121	80	80.85.158.121	GET /?Mon&flight=3807&traveling=K3jx0ALQFozYdZB1xF9Pj7jELUyx-e1ZbX-UPfYQ5Hrs
2017-06-20 14:28:13	80.77.82.41	80	observice.info	GET /banners/uaps HTTP/1.1
2017-06-20 14:28:14	80.85.158.121	80	80.85.158.121	POST /?vel&man=2874&traveling=SwZmnIw0AF5A9K2r20KEBTI1JWG9RaPaAlG-pbBE7I52F
2017-06-20 14:28:16	80.85.158.121	80	80.85.158.121	GET /?Reg&man=163&traveling=ThiRbSKAFimdtfAw9H9P2qhkLXwBwfhcOF_heNYwLG-8CRR
2017-06-20 14:28:17	80.85.158.121	80	80.85.158.121	GET /?Mon&flight=3532&man=3839&traveling=ThjxbSKAZimdtfAw9H8_2qhkPXwBWyhcOF
2017-06-20 14:29:02	47.91.121.220	80	multifest.bit	POST / HTTP/1.0
2017-06-20 14:29:05	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:07	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:11	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:14	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:16	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:20	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:24	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:27	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:29	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:29:40	47.91.121.220	80	multifest.bit	POST /com/ HTTP/1.0
2017-06-20 14:30:04	47.91.121.220	80	multifest.bit	POST / HTTP/1.0



Message from webpage



**\*\* YOUR COMPUTER HAS BEEN BLOCKED \*\***

Error # 268d3

Please call us immediately at: 1844 584 6326

Do not ignore this critical alert.

If you close this page, your computer access will be disabled to prevent further damage to our network.

Your computer has alerted us that it has been infected with a virus and spyware. The following information is being stolen...

- > Facebook Login
- > Credit Card Details
- > Email Account Login
- > Photos stored on this computer

You must contact us immediately so that our engineers can walk you through the removal process over the phone. Please call us within the next 5 minutes to prevent your computer from being disabled.

Toll Free: 1844 584 6326

# Tech Support Scams

user

rowinn

) 584-6326

upnow2app.contentfreeandsafe4update.bid says:

WARNING! Your Flash Player is out of date. Please install update to continue.

OK

Adobe

Install the latest update

Update now

# Fake Flash and Java Updates

Later

Install

[Affiliates](#) | [EULA](#) | [TOS](#) | [Privacy](#) | [Download Manager](#) | [Uninstall](#) | [Contact](#)

By downloading, you accept our [TOS](#) and [Privacy Policy](#).  
This free download is done via download manager which may offer other applications you can decline or uninstall.  
This site and the download manager have no relationship with the author. Any third party products, brands or trademarks listed above are the sole property of their respective owner.

# Ad Network Profiling and Filtering



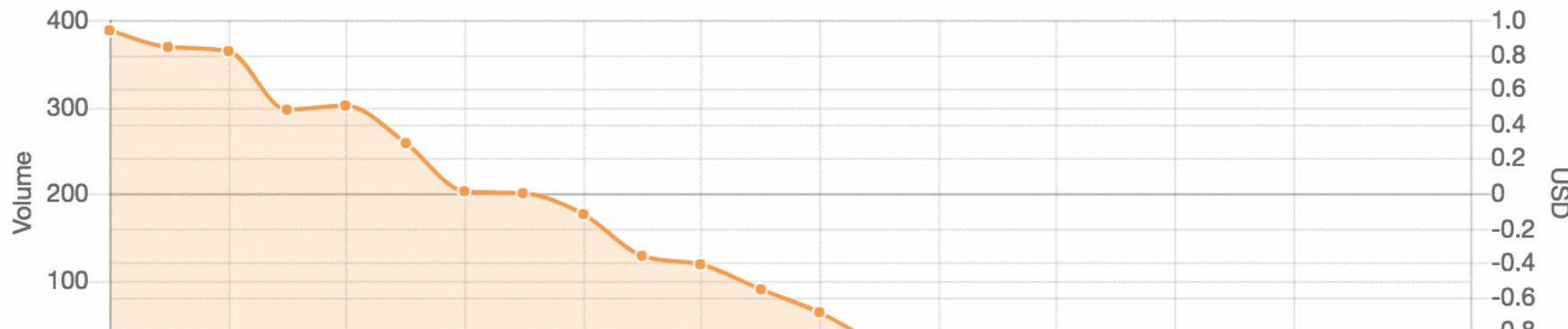
# Dashboard

Today

All Campaigns



Clicks	Unique Clicks	Conversions	Cost	Revenue	Profit/Loss	ROI
2,975	10	41	\$0	\$10.20	\$10.20	0%



Cisco Umbrella



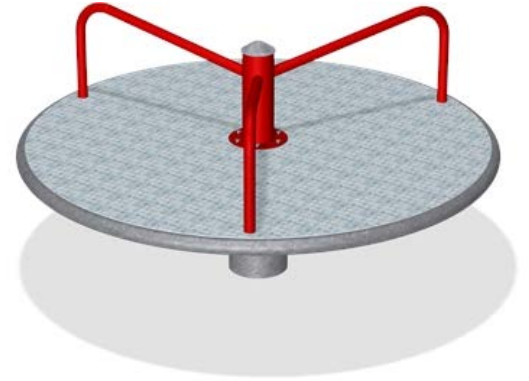
# Proxy Network



Squid Proxy



Choice of region

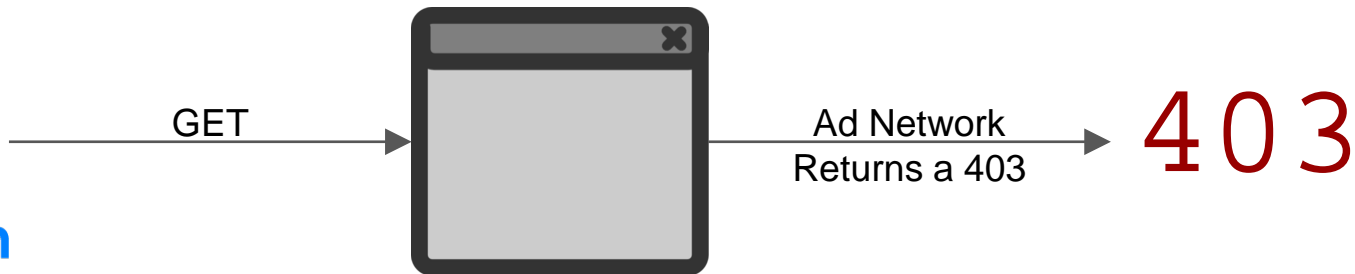


Rotating IPs

# Filtering on non-residential IP Address

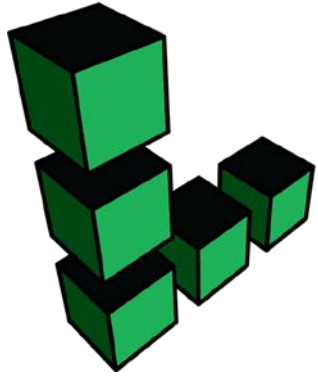


Browsing with  
DigitalOcean  
proxy



Ad Network



# Attempts with other VPS providers



linode

VULTR



	Server	OS	Location	Charges	Status
	<b>squidproxy</b> 512 MB Server - 104.238.137.26		 Miami	\$1.33	<span style="color: green;">●</span> Running <span style="float: right;">⋮</span>

## Advanced Proxy Check



The following lists several of the test results that we perform to attempt to detect a proxy server. Some tests may result in a false positive for situations where there the IP being tested is a network sharing device. In some situations a proxy server is the normal circumstance (AOL users and users in some countries).

Thank you for participating in our test of detecting proxy servers. This proxy detector is constantly being updated. If you are using a proxy server and it was not detected please check back in a few days and see if we are able to detect the proxy server.

To test a different IP address please use the [IP lookup](#) tool.

**VPN leaking your REAL IP address? Try our [VPN Leak test](#).**

Proxy server not detected.

IP	104.238.137.26
rDNS	FALSE
WIMIA Test	FALSE
Tor Test	FALSE
Loc Test	FALSE
Header Test	FALSE
DNSBL Test	FALSE

← →  http://www.clicksgear.com/watch?key=596a48e9d70877a41de91c667376547d&kw=%5B%22le%22%2C%22n%C2%B01%22%2C%22des%22%2C%22sites%22%2C%22de

[Anonymous Proxy detected, click here.](#)

# A



## ADSTERRA ADVERTISEMENT

YOU ARE SEEING THIS PAGE, BECAUSE YOU ARE USING AN ANONYMOUS PROXY.

IN ORDER TO PROTECT OUR ADVERTISERS, WE DO NOT DISPLAY REAL ADS TO THE WEBSITES OWNERS. THE REAL ADS WILL BE DISPLAYED TO YOUR REGULAR TRAFFIC AND WILL BE PAID.

# Lexical Clustering

# Attention to Details

upnow2app.contentfreeandsafe4update.bid says:

WARNING! Your Flash Player is out of date. Please install update to continue.

OK

Adobe

Install the latest update

Update now

# Fake Flash and Java Updates

Later

Install

[Affiliates](#) | [EULA](#) | [TOS](#) | [Privacy](#) | [Download Manager](#) | [Uninstall](#) | [Contact](#)

By downloading, you accept our [TOS](#) and [Privacy Policy](#).  
This free download is done via download manager which may offer other applications you can decline or uninstall.  
This site and the download manager have no relationship with the author. Any third party products, brands or trademarks listed above are the sole property of their respective owner.



SEARCH

PATTERN SEARCH

BULK EDIT



INVESTIGATE

Constrain RegEx search to

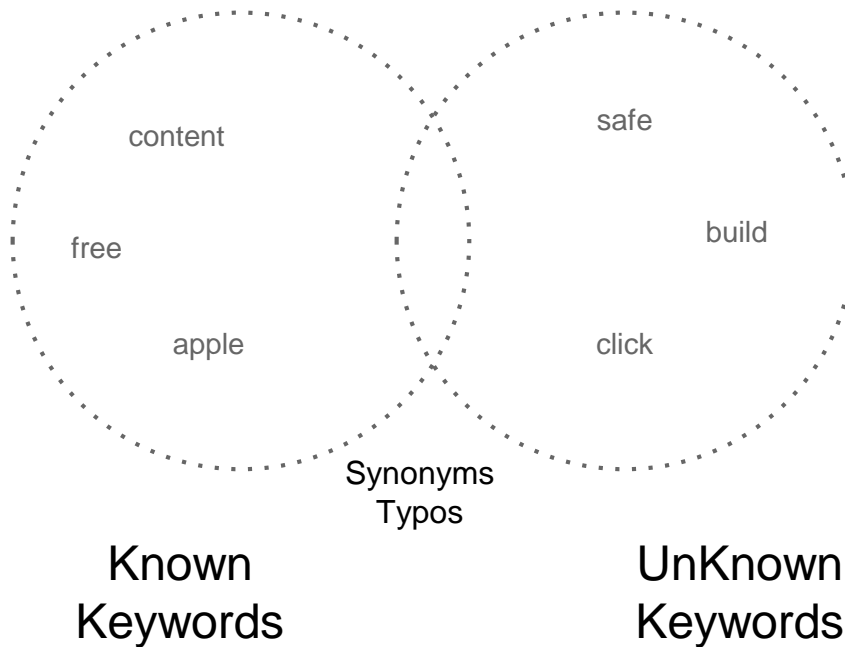
Showing 460 results for **contentfreeandsafe.\***

Domain Name	Security Categ...	First Seen
<a href="#">contentfreeandsafe2updating.stream</a>	Newly Seen Do...	December 13, 2017, 3:17pm
<a href="#">contentfreeandsafetoupdating.review</a>	Newly Seen Do...	December 13, 2017, 3:09pm
<a href="#">contentfreeandsafe4updating.date</a>	Newly Seen Do...	December 13, 2017, 3:00pm
<a href="#">contentfreeandsafeupdatesgreat.win</a>	Newly Seen Do...	December 13, 2017, 2:18pm
<a href="#">contentfreeandsafeupdatingnew.win</a>		December 13, 2017, 11:27am
<a href="#">contentfreeandsafetoupgrade.stream</a>		December 13, 2017, 11:16am
<a href="#">contentfreeandsafe4upgrading.download</a>		December 13, 2017, 10:39am

# More or Less Traveled Roads



# Consider **the** almighty RegeX Keywords



# Consider the almighty RegeX

~~grep "\*.fake.\*"~~

# Traffic Pattern of Fake Update Sites

contentfreeandsafe4update.bid

INVESTIGATE

BACK TO TOP

DNS queries



contentfreeandsafe2update.date

INVESTIGATE

BACK TO TOP

DNS queries

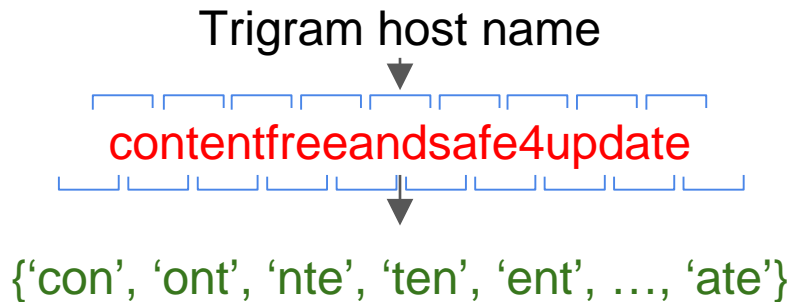


# Traffic Pattern of Fake Update Sites

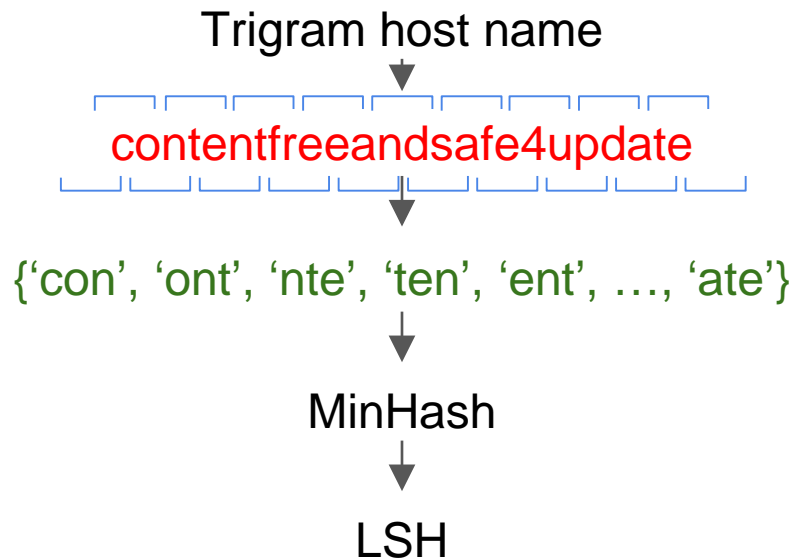
~~Look for burst in traffic~~

For one word, many

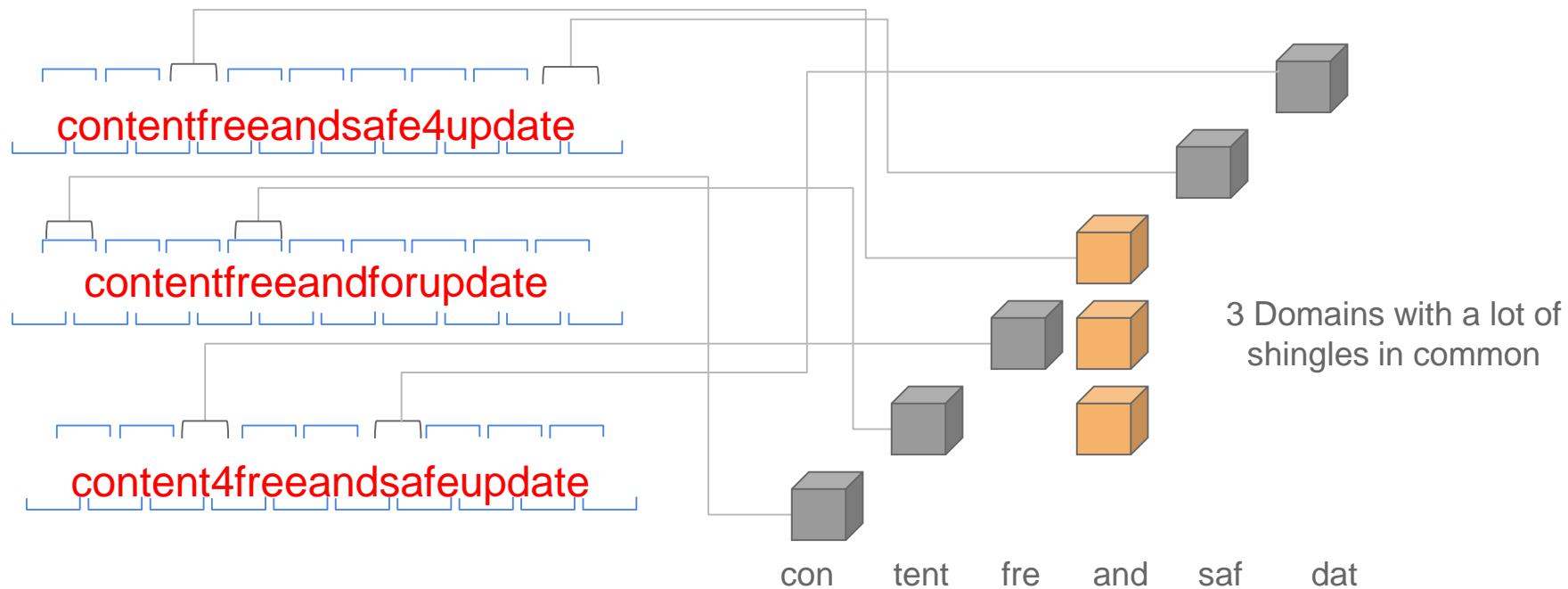
# Shingling Fake Flash and Java Update



# Shingling Fake Flash and Java Update



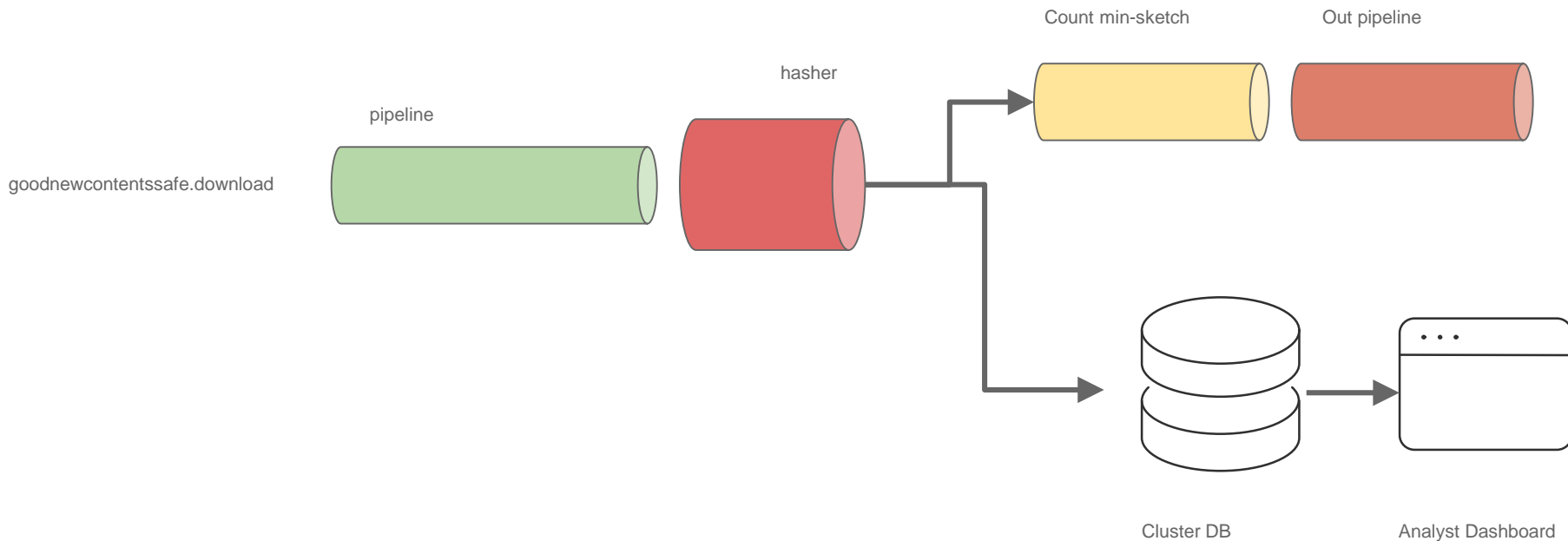
# Locality Sensitive Hashing Fake Flash





# On to production

# Clustering Pipeline Realtime/Batch



# Payday

# Fake Flash and Java Update Lexical Clustering

cluster\_1:

goodnewcontentssafe.download  
goodnewfreecontentsload.date  
goodnewfreecontentall.trade

...

cluster\_3:

artificialintelligencesweden.se  
artificialintelligencechip.com  
artificialintelligence.net.cm

...

cluster\_2:

call-microsoftnw-err81711102.win  
call-microsoftnw-err99817109.win  
call-microsoftnw-err81711101.win

...

cluster\_4:

mkto-sj220048.com  
mkto-sj220146.com  
mkto-sj220162.com

...

# We need help

# Simple Flask App Dashboard

Cluster ID	Preview	Type
1513796401_0	2goodcalling118121234567890.tk 2goodcalling1181212345.tk	Fake Update Tech Support Scam Suspicious

```
30
31     for j, domain in enumerate(entry['domains']):
32         entry['domains'][j] = {'domain': domain, 'timestamp': entry['ti
33
34     entry.pop('timestamps')
35
36     for i, idx in enumerate(date_changes):
37         n = date_changes[i+1] if i < len(date_changes) - 1 else None
38         r[idx:n] = sorted(r[idx:n], key=lambda x: x['c_num'])
39
40 @app.route("/clusters/attribution", methods=['POST'])
41 def attribution():
42     if not request.json:
43         return "Error!"
44     resp = {}
45     for cluster_id in request.json:
46         attr = request.json[cluster_id]
47         ret = add_attribution(cluster_id, attr)
48         resp[cluster_id] = ret
49
50     if ret == 'success' and BLOCKING:
51         domains = m.get_cluster_domains(cluster_id)['domains']
52         block_description = "Domain showed similarities to {0} malverti
53         print "Blocking domains: {0}".format(", ".join(domains))
54         block(domains, block_description=block_description)
55
56     return jsonify(resp)
57
58 @app.route("/clusters/attribution/<string:cluster_id>")
59 def get_attribution(cluster_id):
60     return jsonify(m.get_attribution(cluster_id))
61
62 @app.route("/clusters/uncategorized")
63 def get_uncategorized():
64     r = [entry for entry in m.get_uncategorized()]
65
66     if not r or len(r) == 1:
67         return jsonify(results=r)
68
```

# Hosting space and top talkers

# Where are these hosted? Any patterns?

- Take 1 week's worth of detections and their hosting space; Jan 1-7
- Some hosters are consistently abused

AS12876, FR

AS14618 Amazon AWS and more

Some IPs are actively hosting thousands of domains for



- Some hosters are highly infested with shady, toxic content; dedicated?  
AS202023, LLHOST, RO; phishing, tech support scams, fake updates, porn



# Who is querying these domains?

- Take 1 week's worth of detections; Jan 1-7 and user IPs
- 10 busiest hours

20000+ user IPs querying 2000+ malvertising domains

- Some top talker clusters emerge

Security companies owned ranges querying hundreds of domains

Some rogue networks querying hundreds of domains

# Summary

# Dashboard

Today All Campaigns

Clicks	Unique Clicks	Conversions	Cost	Revenue	Profit/Loss	ROI
2,975	10	41	\$0	\$10.20	\$10.20	0%



SEARCH PATTERN SEARCH BULK EDIT

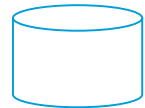
contentfreeandsafe.\* INVESTIGATE Constrain RegEx search to Last 30 days

Showing 460 results for contentfreeandsafe.\*

Domain Name	Security Categ...	First Seen
contentfreeandsafe2updating.stream	Newly Seen Do...	December 13, 2017, 3:17pm
contentfreeandsafetoupdating.review	Newly Seen Do...	December 13, 2017, 3:09pm
contentfreeandsafe4updating.date	Newly Seen Do...	December 13, 2017, 3:00pm
contentfreeandsafeupdatesgreat.win	Newly Seen Do...	December 13, 2017, 2:18pm
contentfreeandsafetoupdatingnew.win		December 13, 2017, 11:27am
contentfreeandsafetoupgrade.stream		December 13, 2017, 11:16am
contentfreeandsafe4upgrading.download		December 13, 2017, 10:39am

~~grep ".\*fake.\*"~~

~~Look for bursts in traffic~~



user IPs



hosting IPs

# Current and Future Work

NLP on misspellings and common typos

Models to categorize clusters

Identifying malicious file hosts using belief propagation

# Thank you

# Questions?

# We are hiring

Matt Foley, [matfoley@cisco.com](mailto:matfoley@cisco.com)

David Rodriguez, [davrodr3@cisco.com](mailto:davrodr3@cisco.com)

Dhia Mahjoub, [dmahjoub@cisco.com](mailto:dmahjoub@cisco.com)