When threat hunting fails

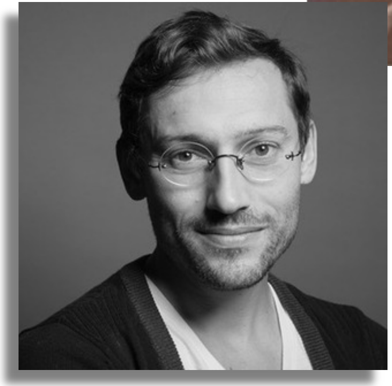
Identifying malvertising domains using lexical clustering

Tucson, January 9th, 2018

Authors

Matt Foley

kitty



David Rodriguez

Dhia Mahjoub

Agenda

Background

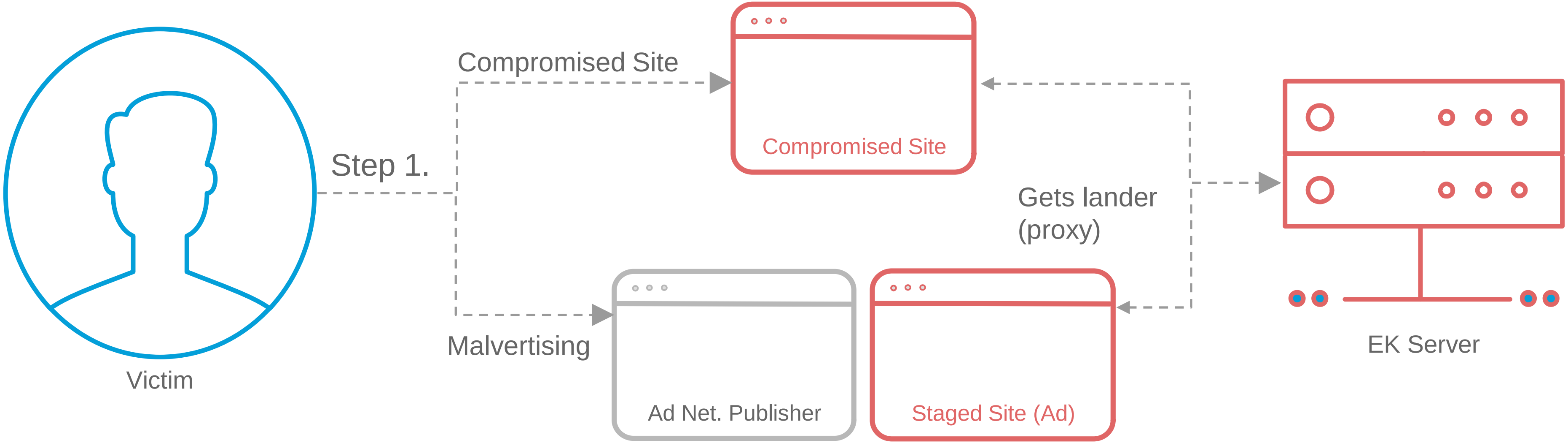
Ad Network Profiling and Filtering

Lexical Clustering

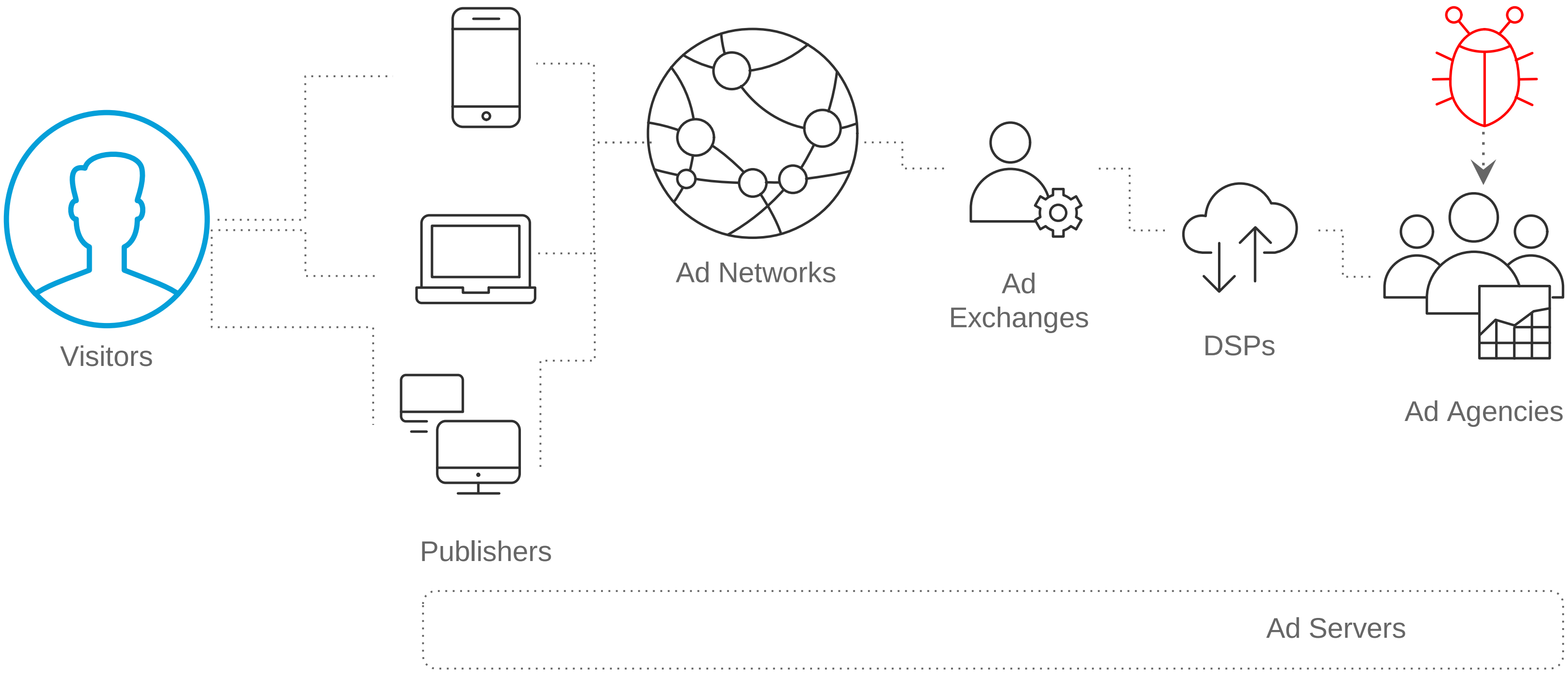
Hosting space and top talkers

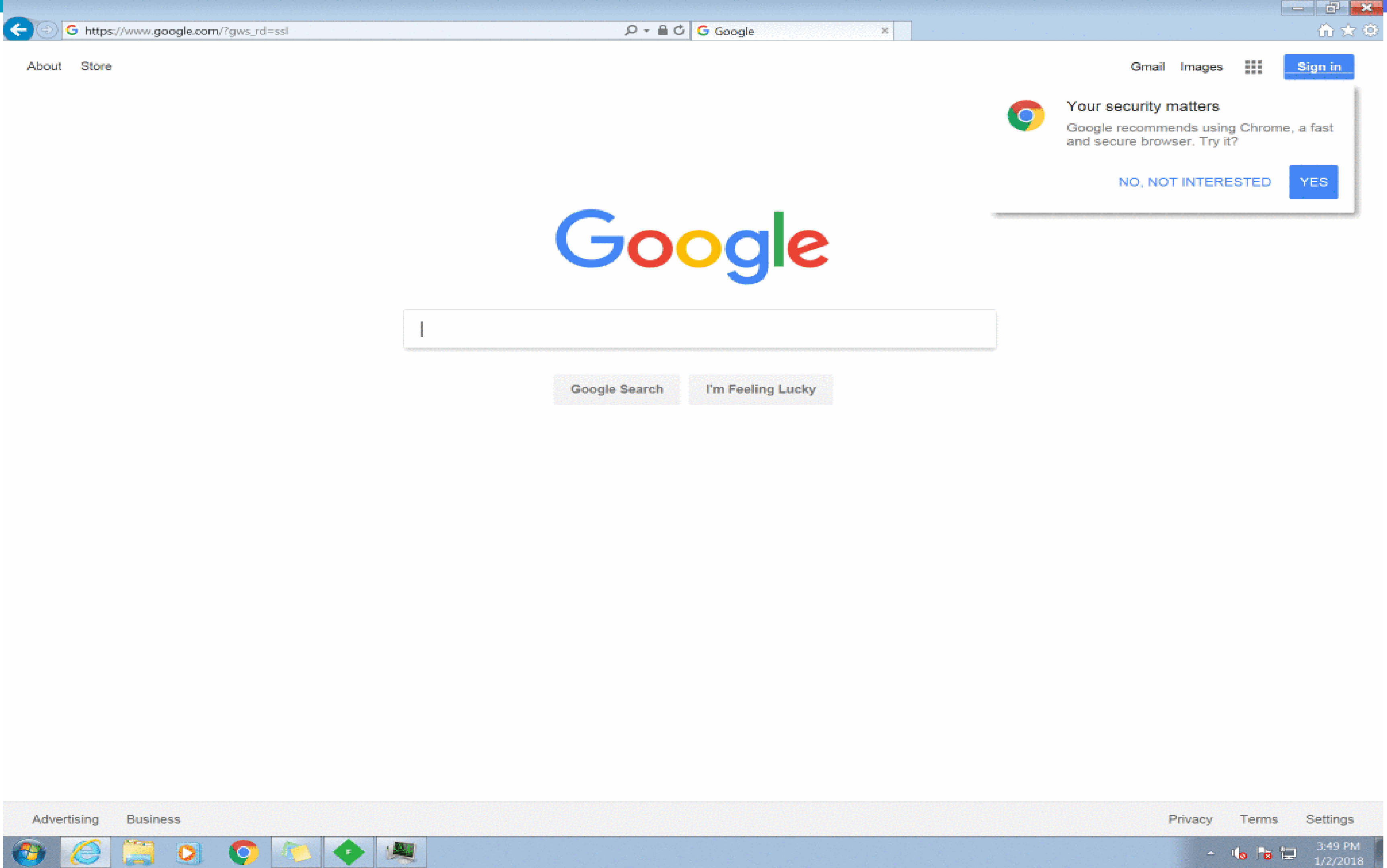
Background

Exploit Kits



What is Malvertising

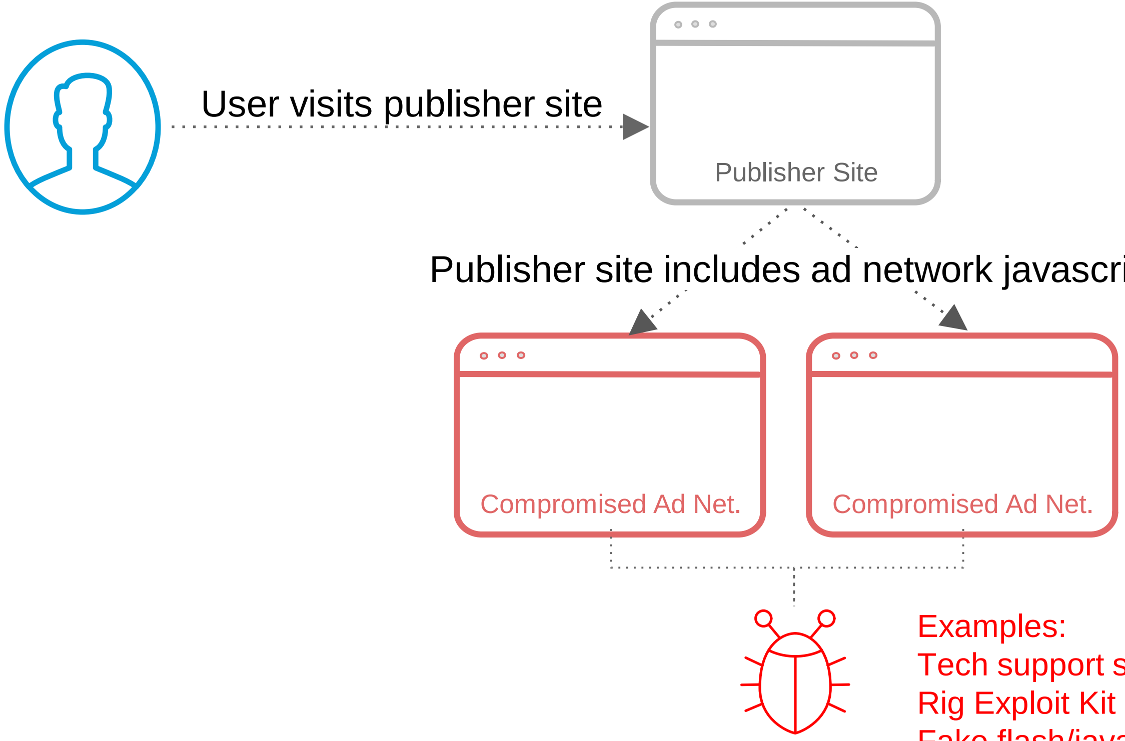




Your security matters

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

Ad Campaign Flow



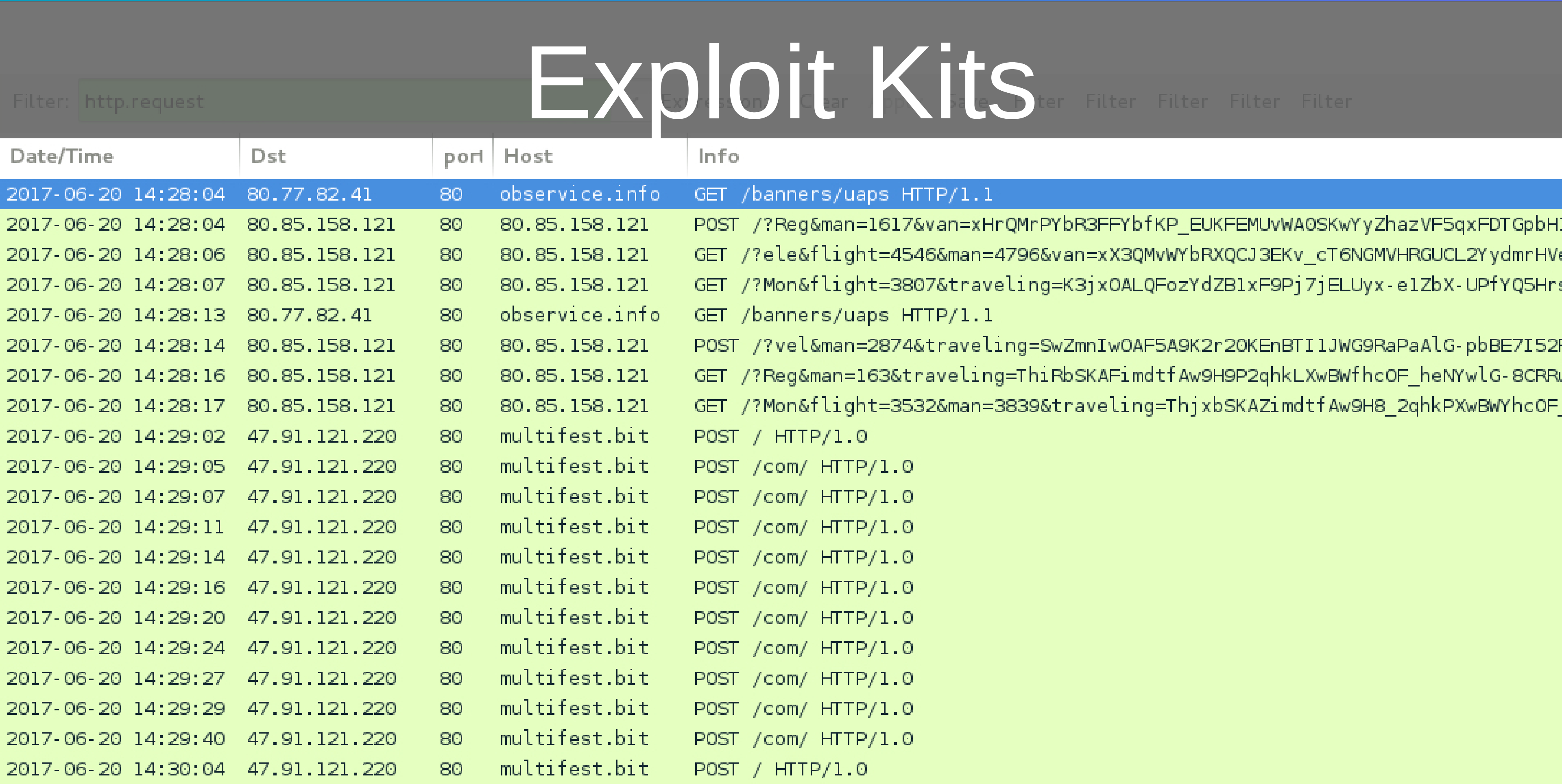
User visits publisher site

Publisher site includes ad network javascript

Examples:
Tech support scam
Rig Exploit Kit
Fake flash/java update

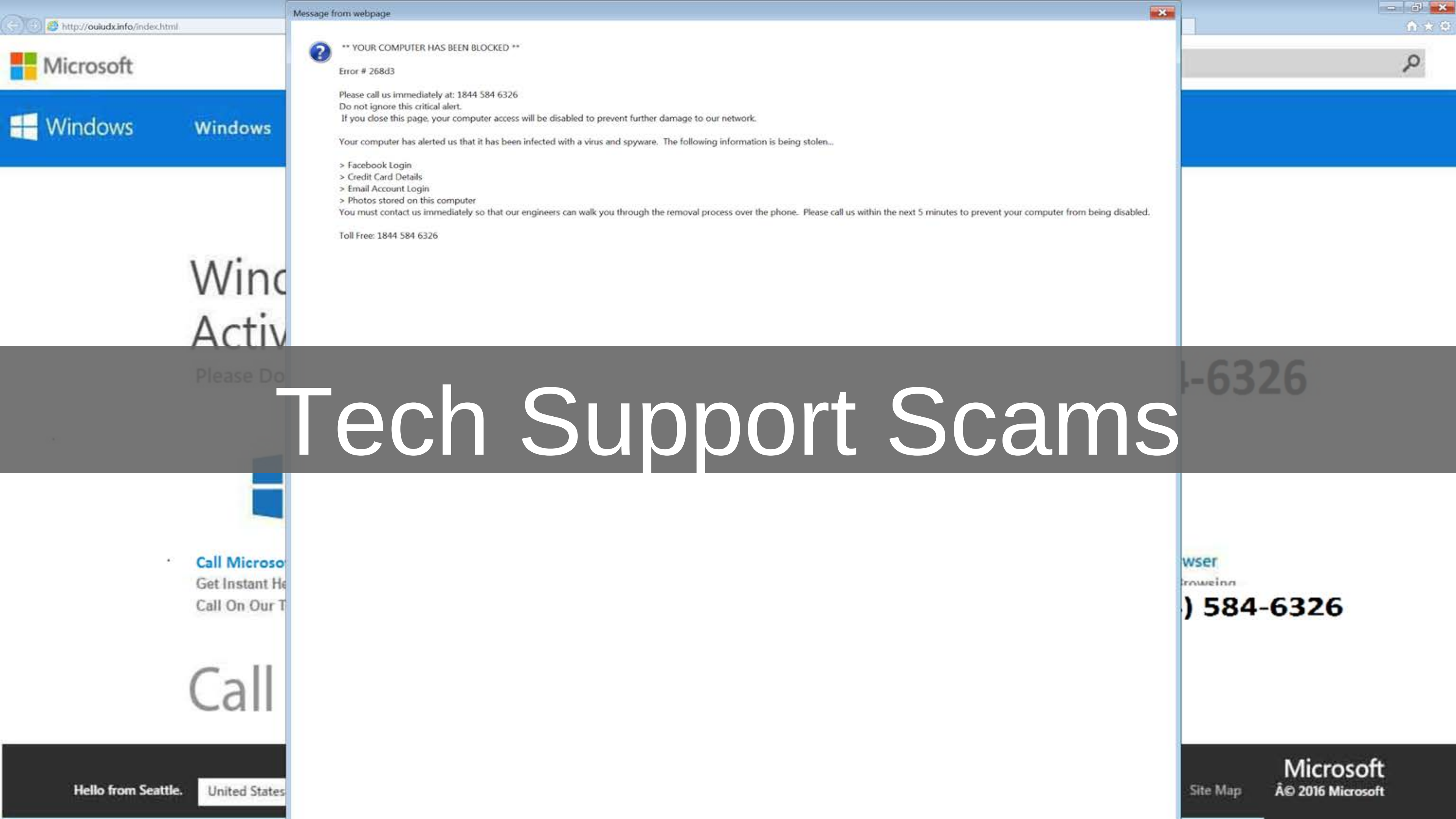
Ad network fingerprints and sends user to malvertisement

Exploit Kits



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 /?ele&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-UPfYQ5Hr
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=SwZmnIw0AF5A9K2r20KEnBTI1JWG9RaPaAlG-pbBE7I52
2017-06-20 14:28:16	80.85.158.121	80	80.85.158.121	GET /?Reg&man=163&traveling=ThiRbSKAFimdtfAw9H9P2qhkLXwBWfhc0F_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_2qhkPXwBWYhc0F_
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

Tech Support Scams



Message from webpage

** YOUR COMPUTER HAS BEEN BLOCKED **

Error # 268d3

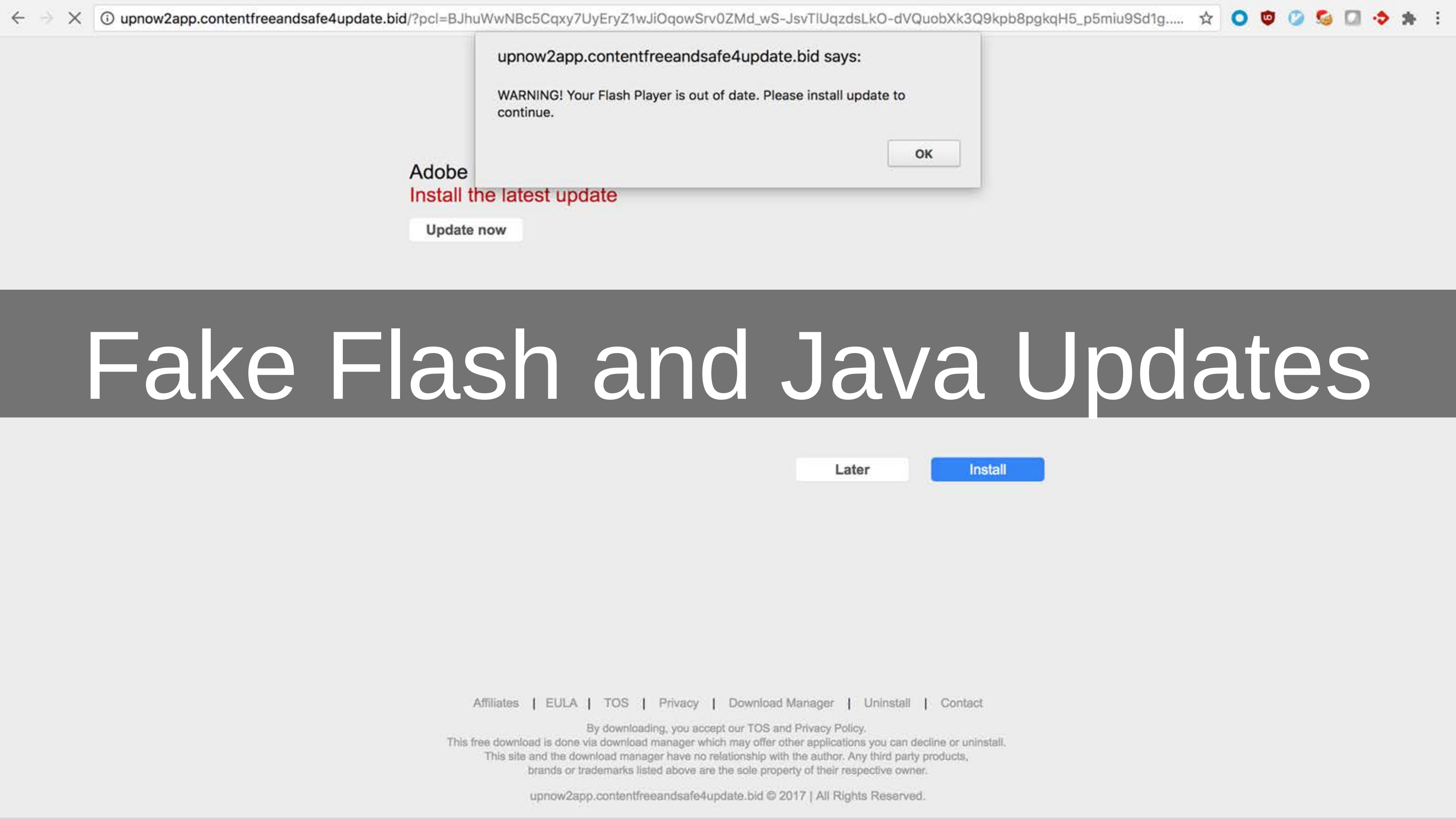
Please call us immediately at: 1844 584 6326
Do not ignore this critical alert.
If you close this page, your computer access will be disabled to prevent further damage to our network.

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

Fake Flash and Java Updates



upnow2app.contentfreeandsafe4update.bid/?pcl=BJhuWwNBc5Cqxy7UyEryZ1wJiOqowSrv0ZMd_wS-JsvTlUqzdsLkO-dVQuobXk3Q9kpb8pgkqH5_p5miu9Sd1g.....

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

Later

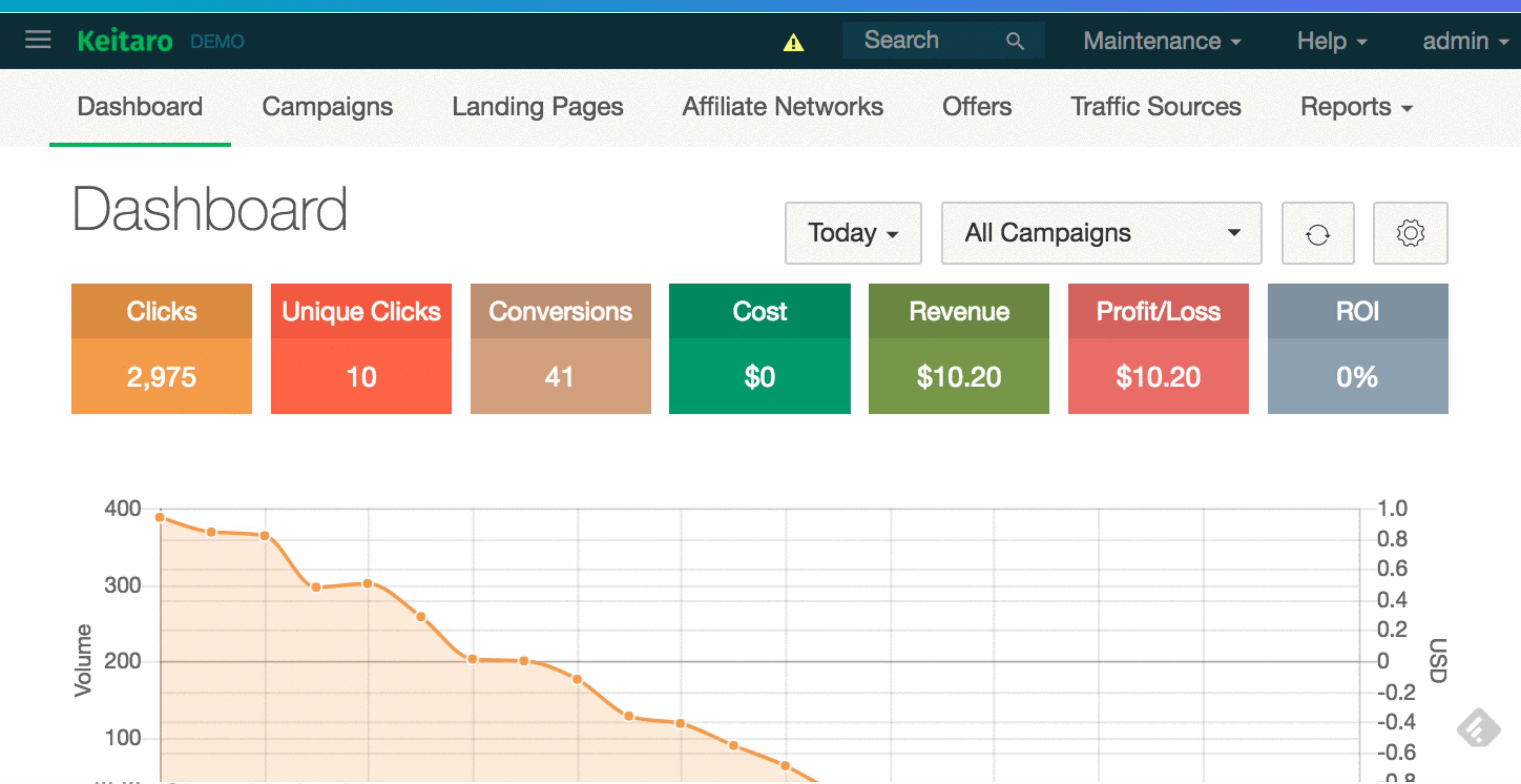
Install

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.

upnow2app.contentfreeandsafe4update.bid © 2017 | All Rights Reserved.

Ad Network Profiling and Filtering



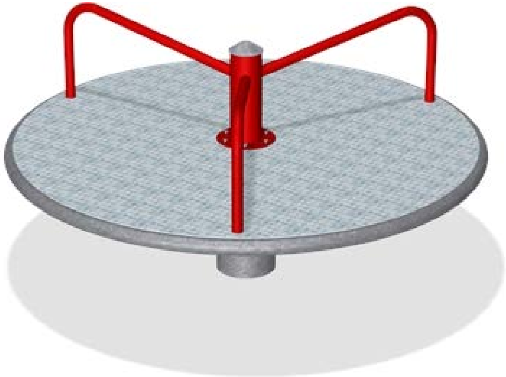
Dashboard

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

Proxy Network

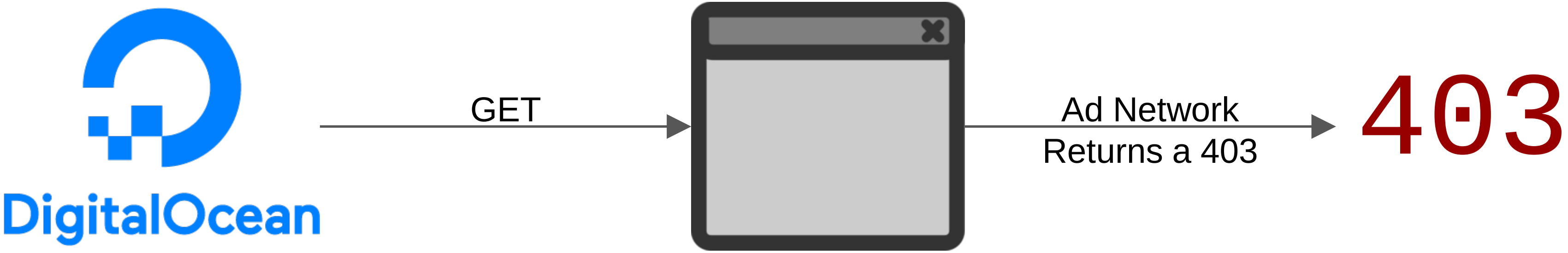
Squid Proxy

Choice of region



Rotating IPs

Filtering on non-residential IP Address



GET

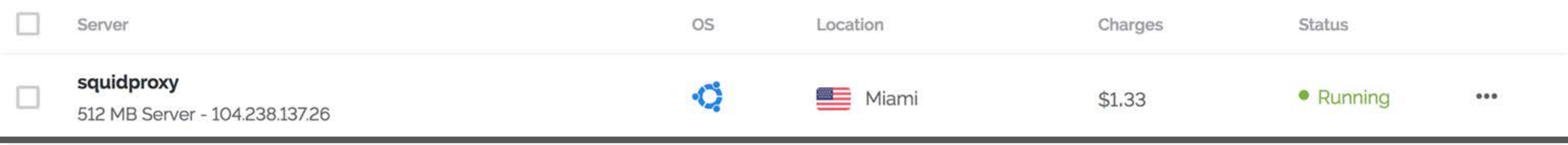
Ad Network Returns a 403

403

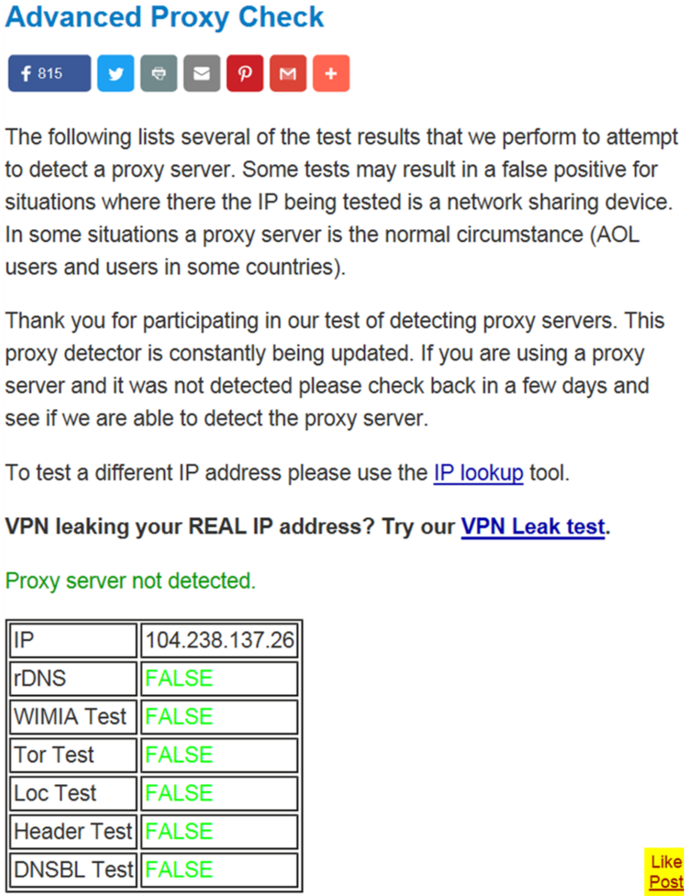
Browsing with DigitalOcean proxy

Ad Network

Attempts with other VPS providers



□	Server	OS	Location	Charges	Status	
□	**squidproxy** 512 MB Server - 104.238.137.26		Miami	$1.33	● Running	•••



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).

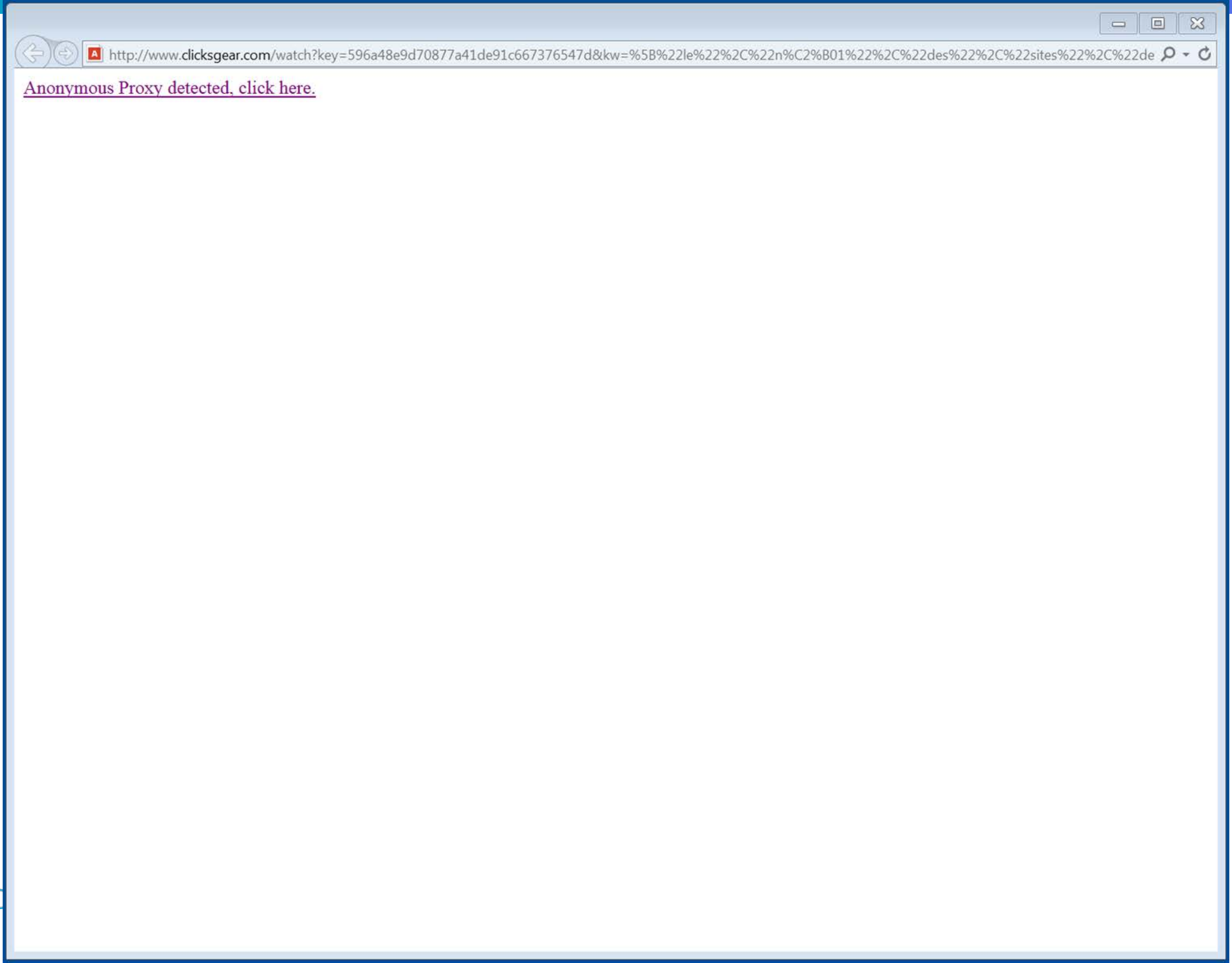
Thank you for participating in our test of detecting proxy servers. This proxy detector is constantly being updated. If you are using a proxy server and it was not detected please check back in a few days and see if we are able to detect the proxy server.

To test a different IP address please use the IP lookup tool.

VPN leaking your REAL IP address? Try our VPN Leak test.

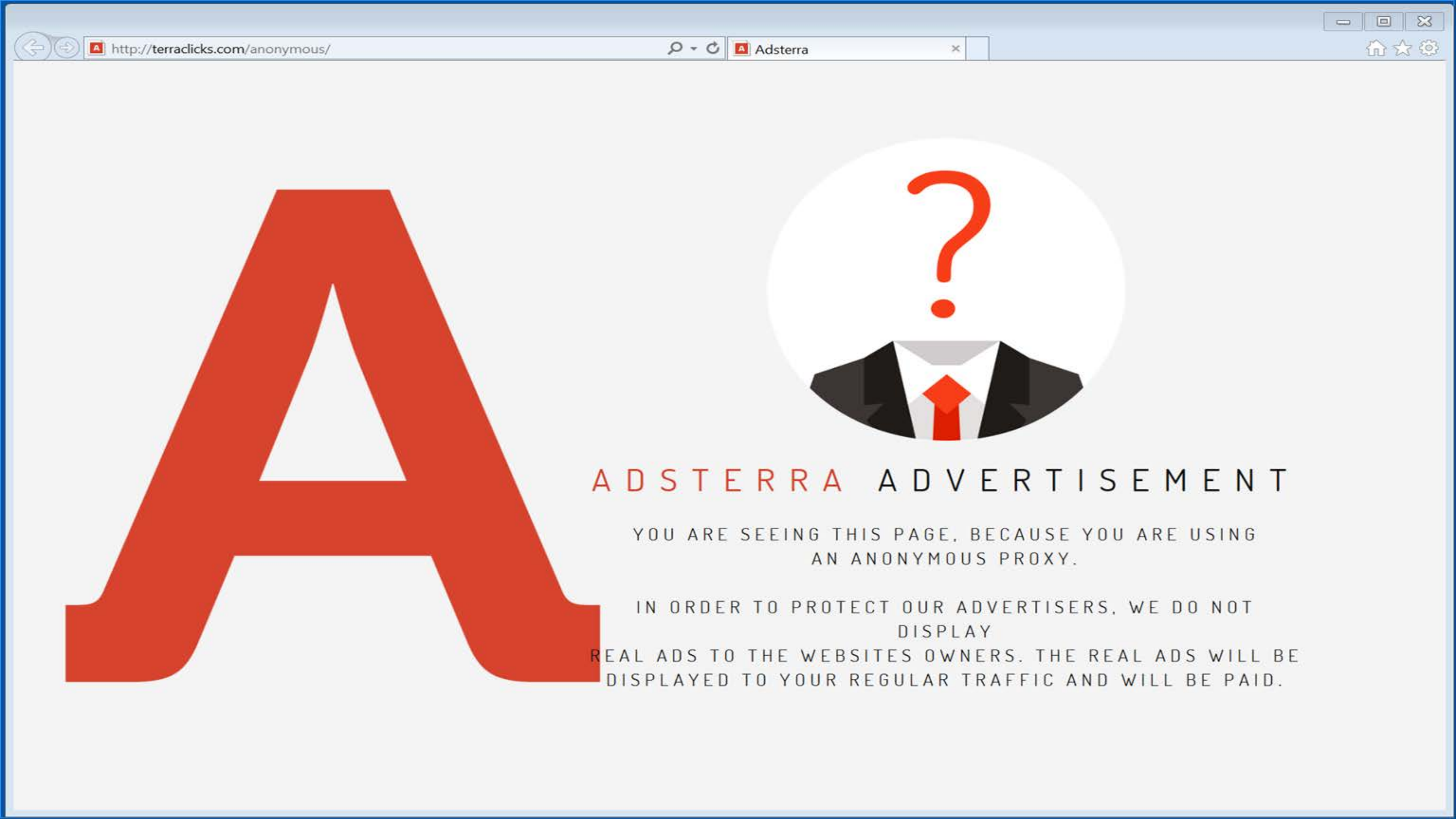
Proxy server not detected.

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.



ADSTERRA ADVERTISEMENT

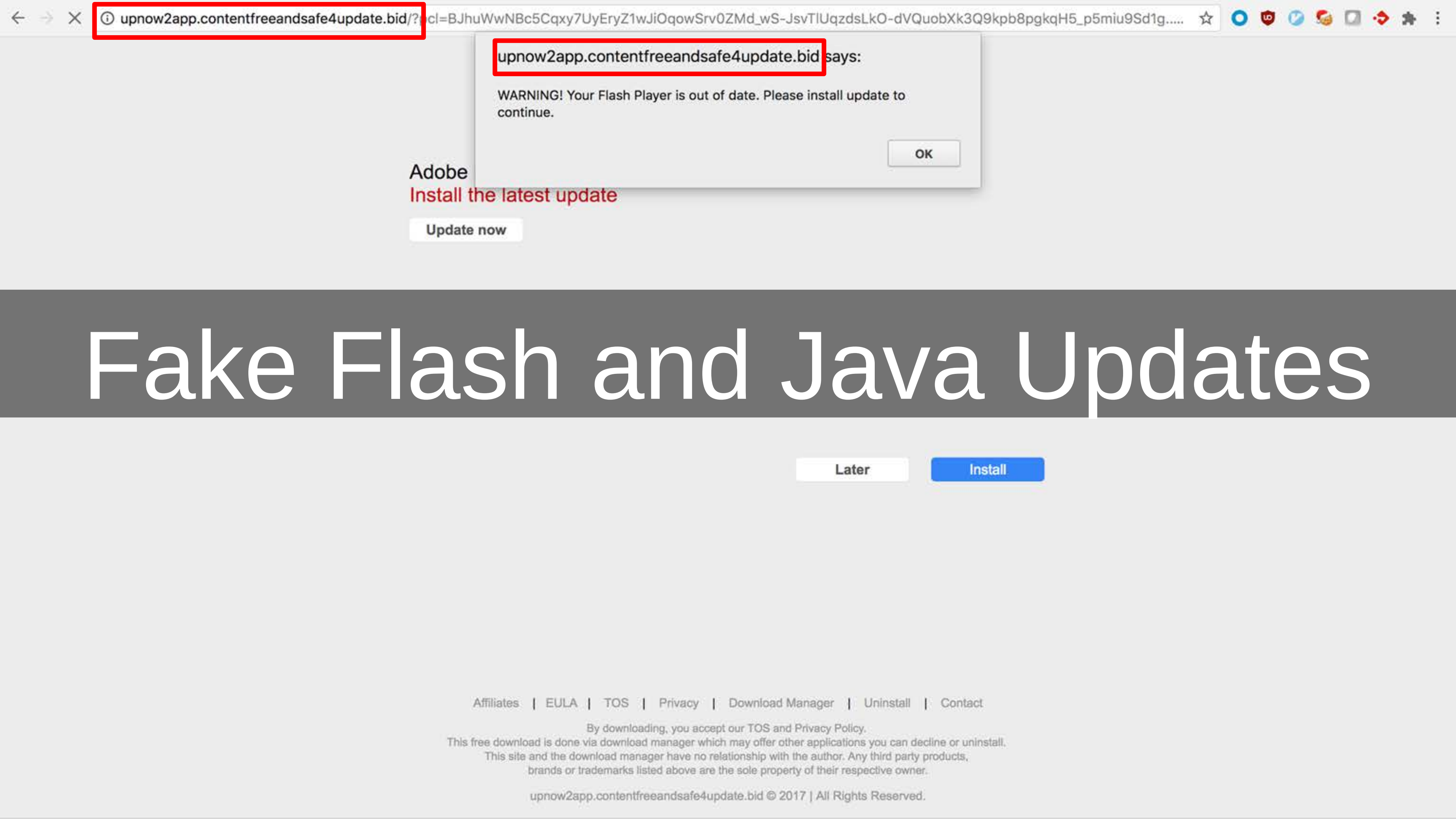
YOU ARE SEEING THIS PAGE, BECAUSE YOU ARE USING AN ANONYMOUS PROXY.

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

Fake Flash and Java Updates



upnow2app.contentfreeandsafe4update.bid/?pcl=BJhuWwNBc5Cqxy7UyEryZ1wJiOqowSrv0ZMd_wS-JsvTlUqzdsLkO-dVQuobXk3Q9kpb8pgkqH5_p5miu9Sd1g.....

upnow2app.contentfreeandsafe4update.bid says:

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

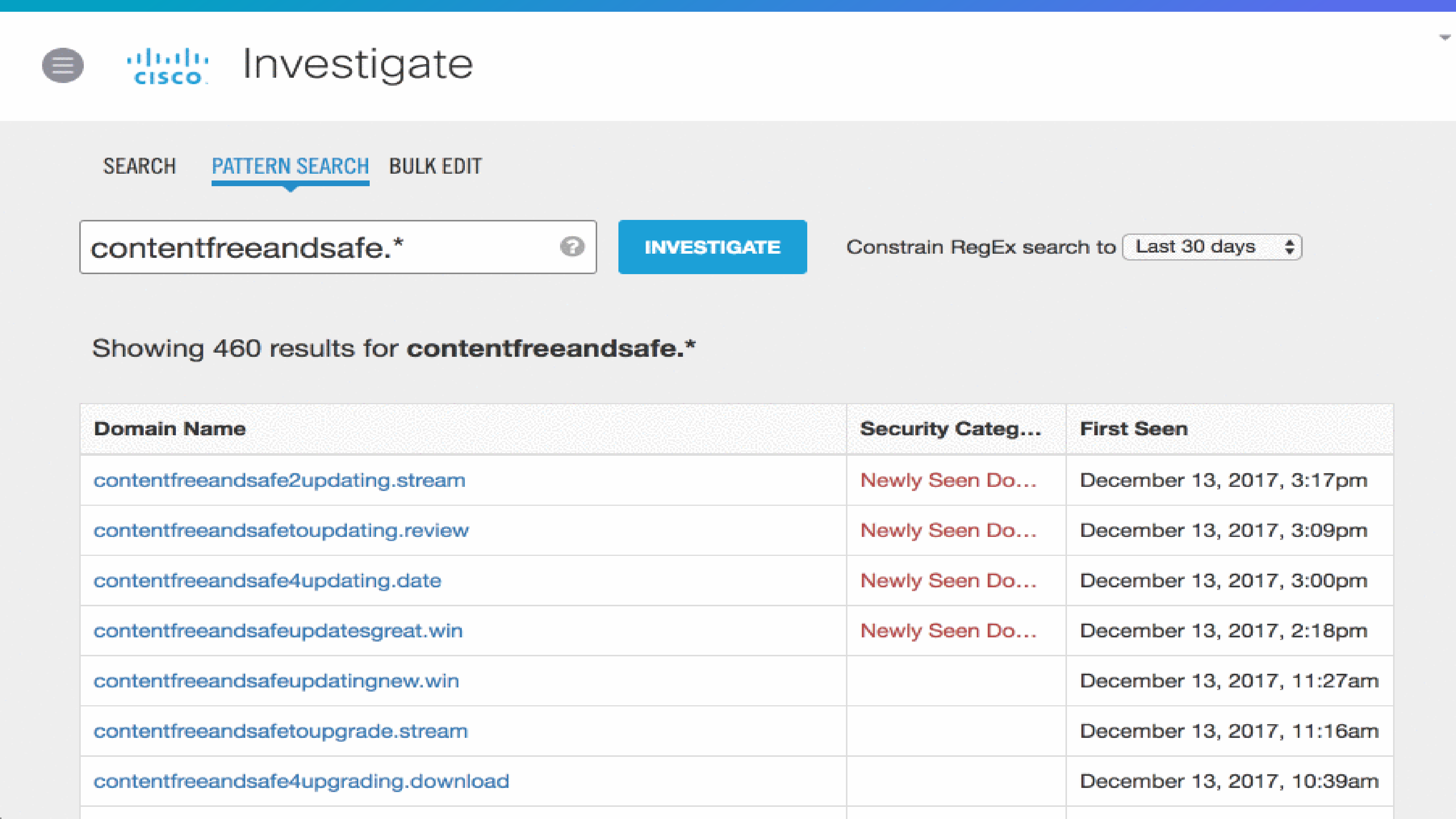
Adobe

Install the latest update

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.

upnow2app.contentfreeandsafe4update.bid © 2017 | All Rights Reserved.



contentfreeandsafe.*

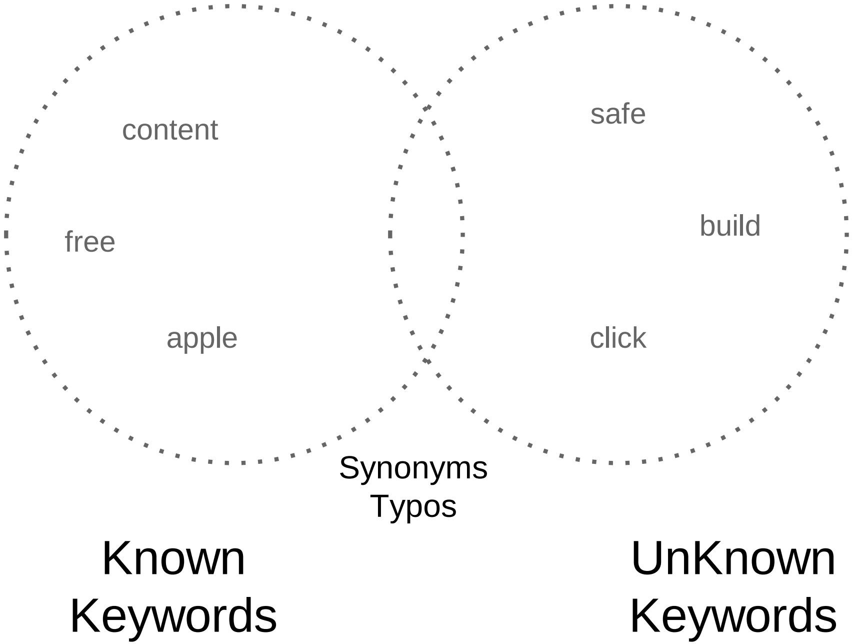
Constrain RegEx search to Last 30 days

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
contentfreeandsafeupdatingnew.win		December 13, 2017, 11:27am
contentfreeandsafetoupgrade.stream		December 13, 2017, 11:16am
contentfreeandsafe4upgrading.download		December 13, 2017, 10:39am

More or Less Traveled Roads

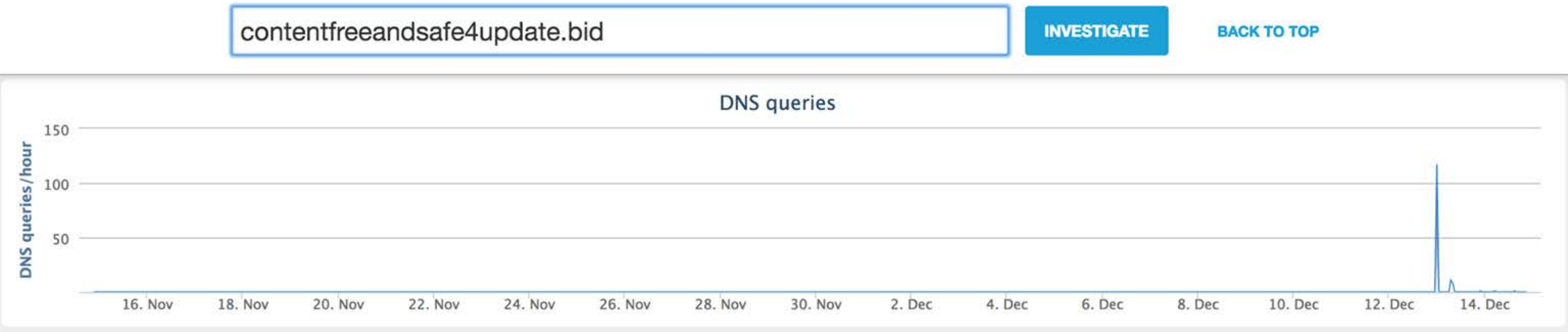
Consider the almighty RegeX Keywords



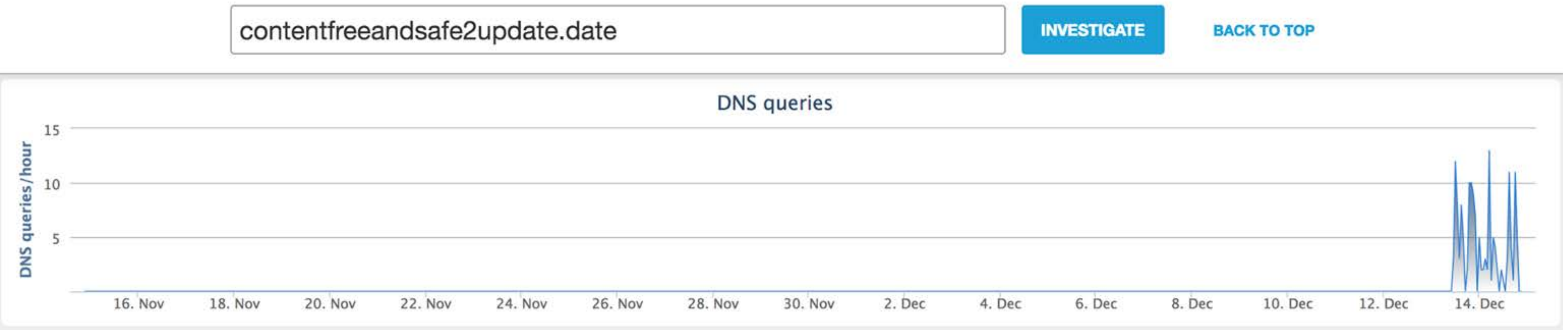
Consider the almighty RegeX

~~grep “*.fake.*”~~

Traffic Pattern of Fake Update Sites



contentfreeandsafe4update.bid



contentfreeandsafe2update.date

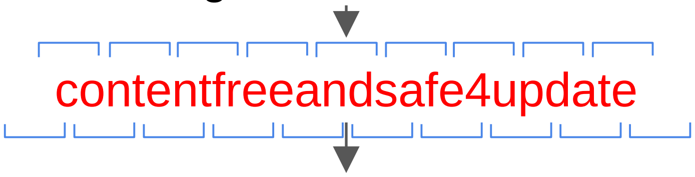
Cisco Umbrella

Traffic Pattern of Fake Update Sites

Look for burst in traffic

For one word, many

Shingling Fake Flash and Java Update

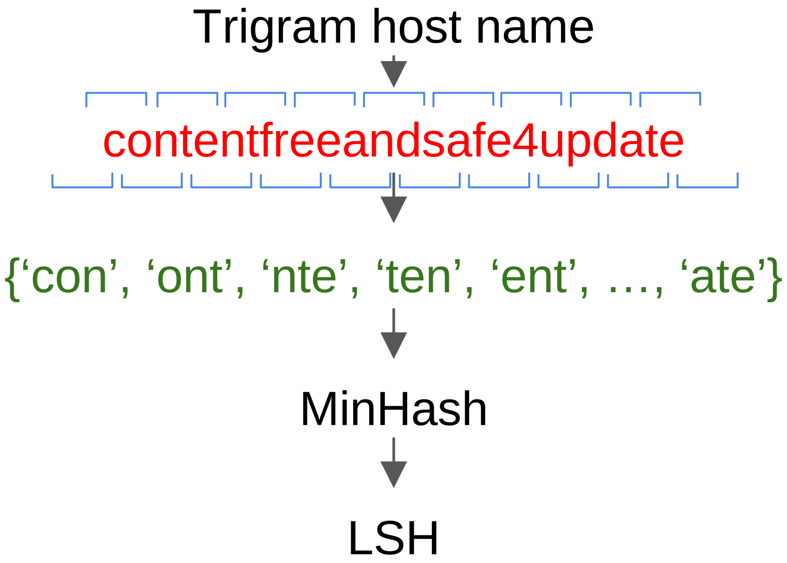


Trigram host name

contentfreeandsafe4update

{‘con’, ‘ont’, ‘nte’, ‘ten’, ‘ent’, …, ‘ate’}

Shingling Fake Flash and Java Update



Trigram host name

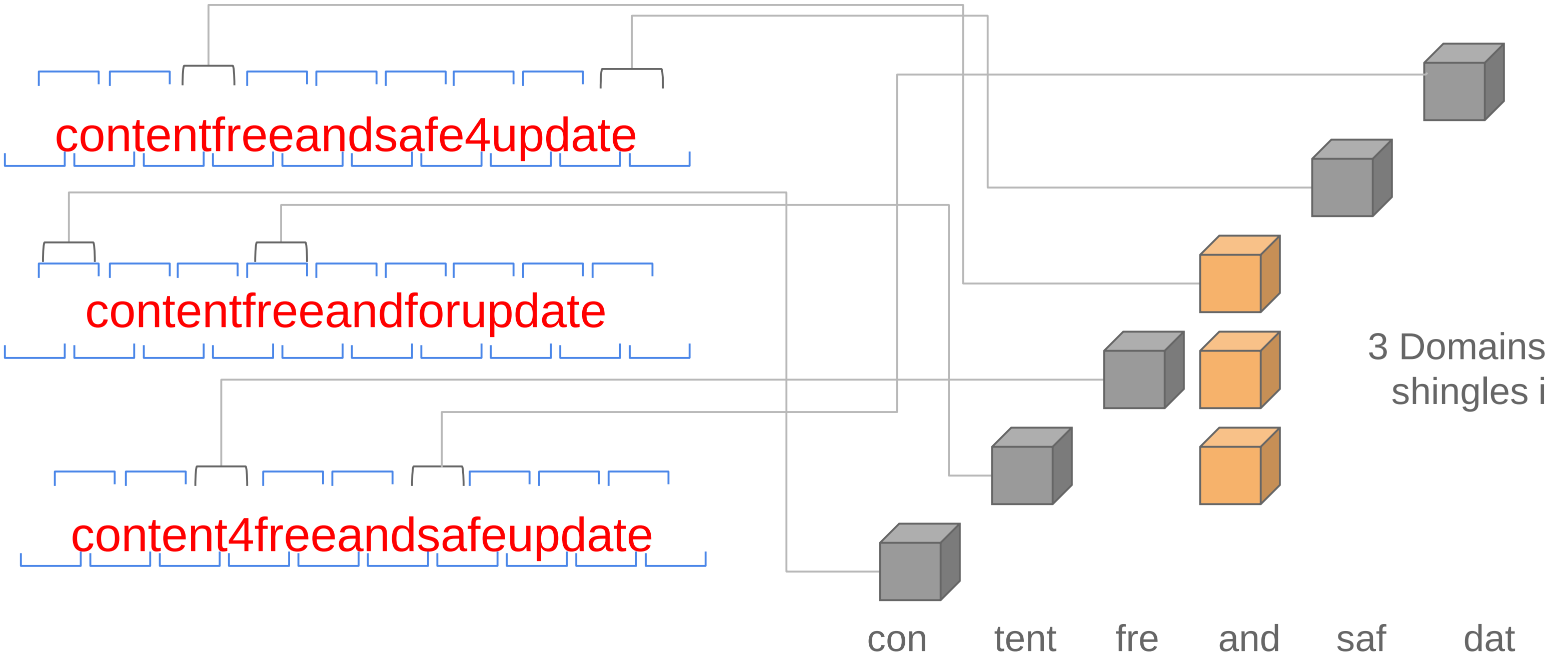
contentfreeandsafe4update

{‘con’, ‘ont’, ‘nte’, ‘ten’, ‘ent’, …, ‘ate’}

MinHash

LSH

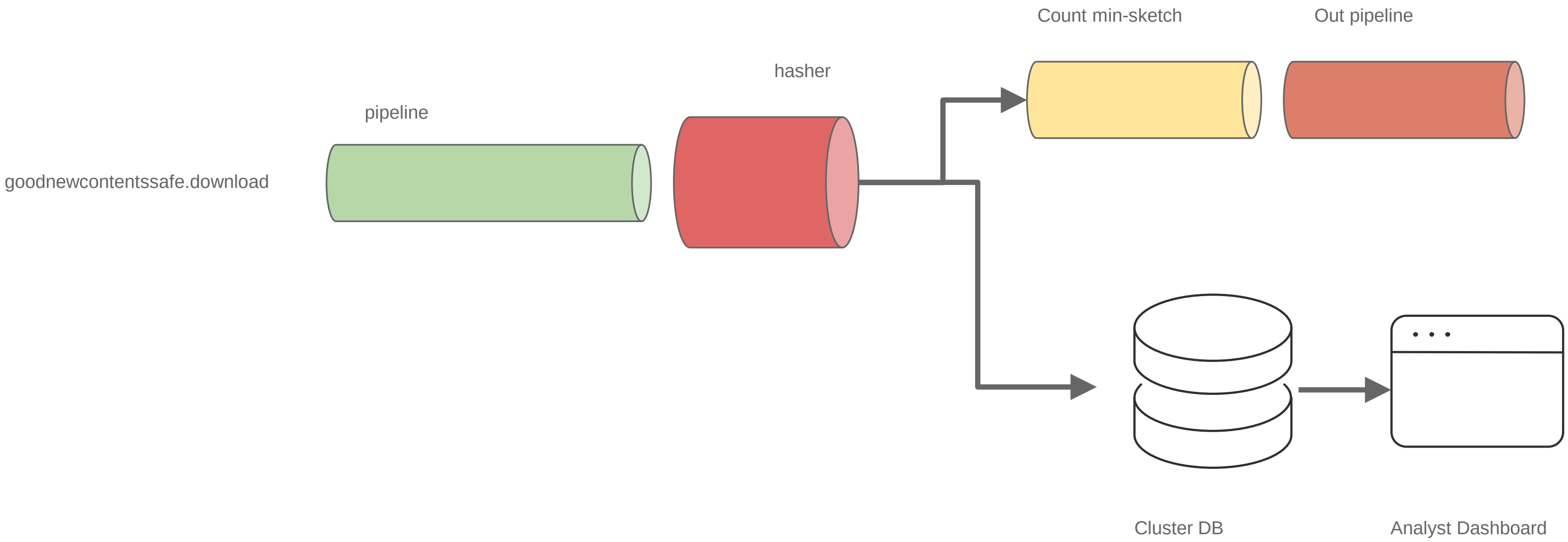
Locality Sensitive Hashing Fake Flash



3 Domains with a lot of shingles in common

On to production

Clustering Pipeline Realtime/Batch



Payday

Fake Flash and Java Update Lexical Clustering

cluster_1:
goodnewcontentssafe.download
goodnewfreecontentsload.date
goodnewfreecontentall.trade
...

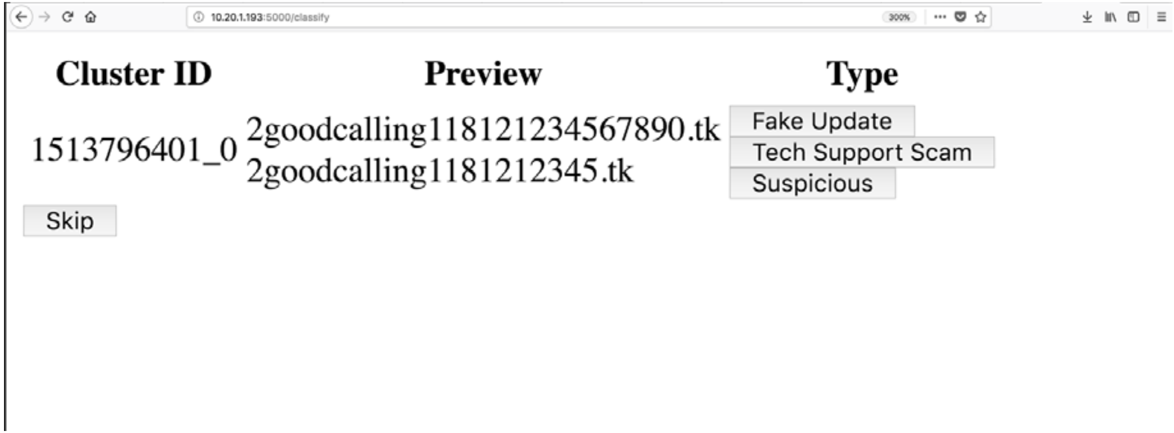
cluster_2:
call-mlcrosoftnw-err81711102.win
call-mlcrosoftnw-err99817109.win
call-mlcrosoftnw-err81711101.win
...

cluster_3:
artificialintelligencesweden.se
artificialintelligencechip.com
artificialintelligence.net.cm
...

cluster_4:
mkto-sj220048.com
mkto-sj220146.com
mkto-sj220162.com
...

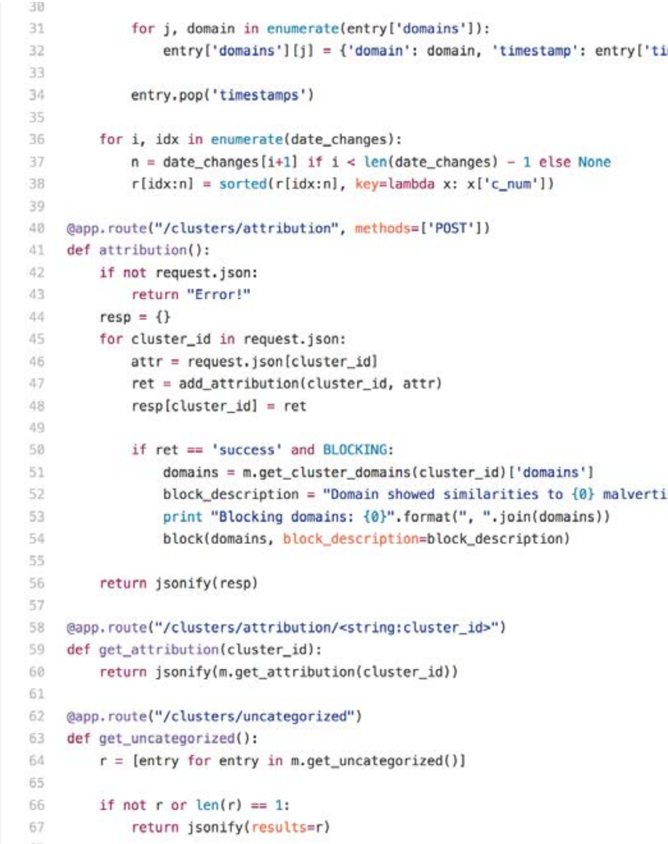
We need help

Simple Flask App Dashboard



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

Skip


```
        for j, domain in enumerate(entry['domains']):
            entry['domains'][j] = {'domain': domain, 'timestamp': entry['ti

        entry.pop('timestamps')

    for i, idx in enumerate(date_changes):
        n = date_changes[i+1] if i < len(date_changes) - 1 else None
        r[idx:n] = sorted(r[idx:n], key=lambda x: x['c_num'])

@app.route("/clusters/attribution", methods=['POST'])
def attribution():
    if not request.json:
        return "Error!"
    resp = {}
    for cluster_id in request.json:
        attr = request.json[cluster_id]
        ret = add_attribution(cluster_id, attr)
        resp[cluster_id] = ret

        if ret == 'success' and BLOCKING:
            domains = m.get_cluster_domains(cluster_id)['domains']
            block_description = "Domain showed similarities to {0} malverti
            print "Blocking domains: {0}".format(", ".join(domains))
            block(domains, block_description=block_description)

    return jsonify(resp)

@app.route("/clusters/attribution/<string:cluster_id>")
def get_attribution(cluster_id):
    return jsonify(m.get_attribution(cluster_id))

@app.route("/clusters/uncategorized")
def get_uncategorized():
    r = [entry for entry in m.get_uncategorized()]

    if not r or len(r) == 1:
        return jsonify(results=r)
```

Cisco Umbrella

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 months

- 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

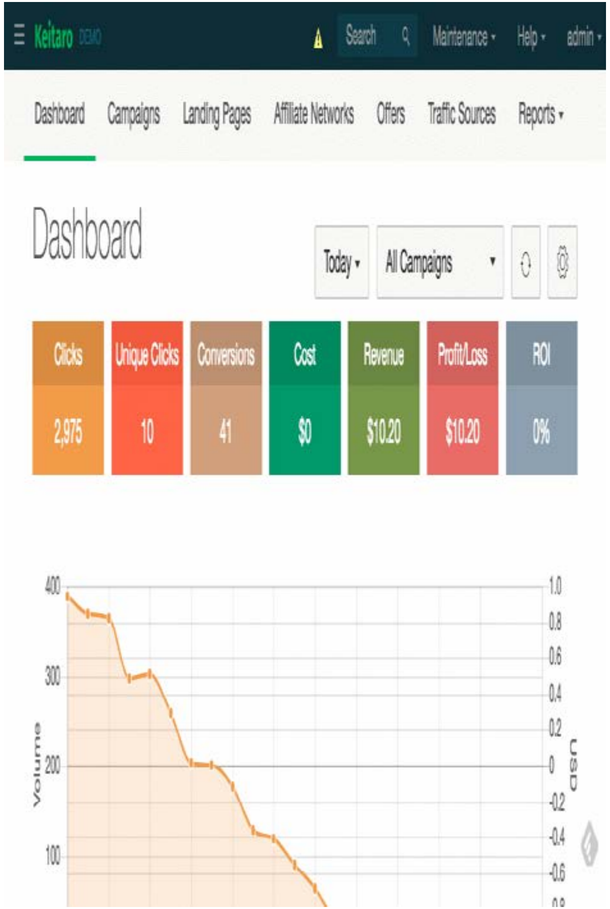
  20000+ user IPs querying 2000+ malvertising domains

- Some top talker clusters emerge

  Security companies owned ranges querying hundreds of domains

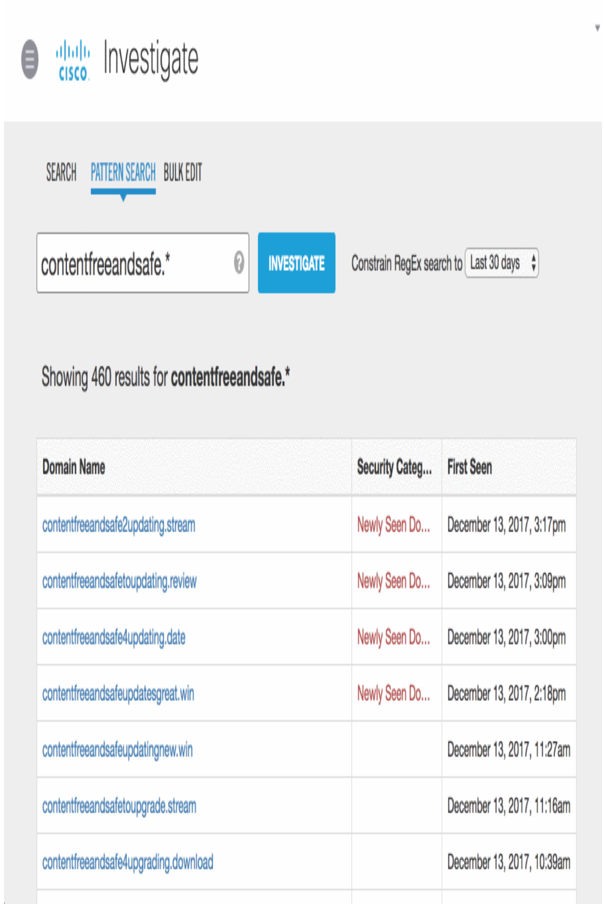
  Some rogue networks querying hundreds of domains

Summary



Dashboard

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



contentfreeandsafe.*

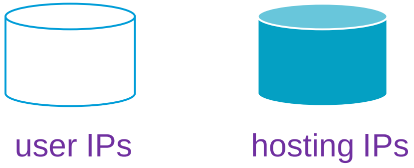
Constrain RegEx search to Last 30 days

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
contentfreeandsafeupdatingnew.win		December 13, 2017, 11:27am
contentfreeandsafetoupgrade.stream		December 13, 2017, 11:16am
contentfreeandsafe4upgrading.download		December 13, 2017, 10:39am

~~grep “*.fake.*”~~

~~Look for burst in traffic~~



Cisco Umbrella

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

David Rodriguez, davrodr3@cisco.com

Dhia Mahjoub, dmahjoub@cisco.com