an interesting/influential/important paper from the world of CS every weekday morning, as selected by Adrian Colyer

Who controls the Internet? Analyzing global threats using property traversal graphs

MAY 19, 2017

tags: Security

Who controls the Internet? Analyzing global threats using property traversal graphs (http://trouge.net/gp/papers/gthreats_www2017.pdf) Simeonovski et al., *WWW'17*

Who controls the Internet? How much influence do they have? And what would happen if one of those parties launched an attack or was compromised and used to launch an attack? Previous works have looked at the individual core services, but this paper focuses on their inter-dependencies along attack propagation paths.

> *An increasing number of reports and studies are showing that a limited number of players have an important influence on the overall security of the Internet infrastructure... we have a rather limited capability to assess the impact of attacks against, or performed by, core service providers.*

What kind of attacks are we talking about? Three large-scale security incidents form the initial motivation:

1. The Great Cannon DDoS Attack of March 16th 2015, a massive DDoS attack caused by malicious JavaScript code injected into TCP connections crossing Chinese network borders. The injected code aggressively requested resources from the DDoS targets.
2. The PRISM program (2013), an NSA surveillance program with direct access to Internet communications and stored information. "*While the direct involvement of popular tech providers is still unclear, in this paper we make the assumption that establishing the this type of collaboration is possible and can be voluntary, or coerced by authorities by means of law and court orders.*"
3. The DDoS attack against Dyn.com of October 21st 2016. The attack caused Dyn.com customers including Amazon, Netflix, Twitter, Reddit, and Spotify to experience outages on name resolution, affecting hundreds of millions of Internet users who could not access their services.

Four different attack vectors are analysed: email sniffing, redirection via malicious domain resolution, in-path content injection, and hosting malicious content.

Gathering information

The authors crawl the web starting from the top 100K Alexa domains, expanding to server and network information, and then adding in organisations and countries. This ultimately leads to a labeled graph containing 1.8M nodes, of which 350K are unique IP addresses. The nodes are connected by 4.7M relationships.

The following table shows labels (think node and edge types) in the graph:

Labels	Description
Address	Node for IP address
Domain	Node for a domain name; the source data set, e.g., Alexa or JS, is a node property
DNS Zone	The zone administrated by an authoritative name server
AS	IANA number assigned to the AS; The hosted IPs is a node property
Country	Code Country code, number of IPs
Organization	Service provider name
`ORIG_FROM`	*AS* where an *Address* originates from
`LOC_IN`	*Country* where an element is located
`CTRL_BY`	*Organization* controlling, e.g., a *Domain*
`A`	DNS record mapping *Address* to *Domain*
`MX`	DNS record mapping *Domain* for email delivery
`NS`	DNS record for name servers
`ZONE`	DNS record for authoritative information of a *DNS zone*
`CNAME`	Aliases from *Domain* to *Domain*
`PTR`	PTR DNS record type maps an *Address* to a *Domain*
`INCL_JS_FROM`	*Domain* name or *Address* hosting JS library

Table 1: Labels of nodes and relationships

When considering the impact of an attack nodes can be marked at one of three different compromise levels: comprised, partially compromised, and non-compromised. Taint-style propagation rules can then be written which capture how attacks can spread through the network. For example, if a node n is compromised and there is an edge from n to m labeled as A (name lookup) then m is marked as compromised.

Identifying the most promising attack targets

> *Before assessing attacks, we use our model to select entities that can be either attack victims or the attackers. The selection criteria are based on metrics that reflect the popularity and the influence of entities.*

The most promising attack targets (or viewed from another perspective, the entities with the most power over Internet infrastructure) are identified via six metrics.

Who hosts the most Alexa domains?

The analysis is done by country (giving a lens into the power of nation-state attackers), and also by 'Autonomous Systems' (AS) – a collection of IP networks and routers under the control of a given network operator.

Under this metric, these are the most powerful countries:

Country	Dom.
United States	30,582
Netherlands	4,296
Germany	4,178
China	4,158
Japan	3,053
France	2,526
Great Britain	2,400
Russia	1,678
Canada	1,186
India	1,087
Ireland	986
EU	950
Spain	848
South Korea	755

And these are the most powerful network operators:

ASN	Name	Dom.
13335	CloudFlare	7,170
16509	Amazon-1	2,816
14618	Amazon-2	1,892
20940	Akamai	1,830
16276	Ovh	1,025
37963	Alibaba	779
24940	Hetzner	725
15169	Google	525
36351	Softlayer	518
4134	ChinaNet	468
19551	Incapsula Inc	397
54113	Fastly	361
63949	Linode	358
4808	China Unic.	348

Who has the most JavaScript hosting servers?

By country:

Country	JS
United States	47,910
Germany	7,830
China	7,273
Netherlands	6,963
Great Britain	4,455
Japan	4,205
France	4,048
Russia	2,865
Ireland	1,919
EU	1,581
Canada	1,347
Italy	1,159
Spain	952
Poland	943

And by network operator:

ASN	Name	JS
16509	Amazon-1	10,085
13335	CloudFlare	5,489
20940	Akamai	3,004
14618	Amazon-2	2,207
16276	Ovh	1,970
24940	Hetzner	1,508
15133	EdgeCast	1,360
37963	Alibaba	940
36351	Softlayer	910
4134	ChinaNet	814
15169	Google	814
4837	China169	728
54994	Quantil	606
35415	Webzilla	551

Who hosts the most email servers?

By country:

Country	MX
United States	41,434
Germany	12,047
Great Britain	6,811
France	6,261
Netherlands	6,091
Japan	4,314
Russia	3,923
Italy	3,293
Canada	3,042
Ireland	2,897
Spain	2,703
Turkey	2,094
Iran	1,946
India	1,892

And by network operator:

ASN	Name	MX
8075	Microsoft	8,503
16276	Ovh	2,669
24940	Hetzner	2,497
46606	Unified L.	1,353
36351	Softlayer	865
26496	GoDaddy	799
16509	Amazon-1	643
60781	Leaseweb	579
15169	Google	568
39572	Advancedh.	522
12876	AS12876	452
63949	Linode	438
14618	Amazon-2	329
32475	SingleHop	298

Who hosts the most name servers?

By country:

Country	NS
United States	34,235
Germany	6,697
France	3,865
Great Britain	3,139
Netherlands	3,116
Canada	2,244
Russia	2,167
Turkey	2,143
Japan	2,126
Spain	1,662
China	1,617
Iran	1,552
Brazil	1,070
India	954

And by network operator:

ASN	Name	NS
16276	Ovh	2,415
24940	Hetzner	2,131
16509	Amazon	1,907
46606	Unified L.	1,524
36351	Softlayer	1,345
32475	SingleHop	1,155
13335	CloudFlare	699
32244	Liquid Web	674
16552	Tiggee	611
26496	GoDaddy	535
60781	Leaseweb	398
33517	DynDNS	364
12876	AS12876	354
4808	China Unic.	351

Who has the most power over JavaScript providers?

This metric measures the number of JS hosting servers whose authoritative name server is hosted in a given country or by a given network operator.

By country:

Country	JS/NS
United States	41,231
Germany	3,101
Netherlands	3,045
China	3,009
Russia	2,254
France	2,084
Japan	2,000
EU	1,636
Great Britain	1,364
Spain	1,219
Canada	801
Singapore	787
Poland	540
Iran	474

And by network operator:

ASN	Name	JS/NS
16509	Amazon-1	15,429
13335	CloudFlare	4,933
33517	DynDNS	3,570
4837	China169	2,008
26496	GoDaddy	1,938
4812	China Tlc.	1,875
16552	Tiggee	1,467
16276	Ovh	1,307
15169	Google	1,012
24940	Hetzner	873
15395	London Off.	822
36351	Softlayer	753
4808	China Unic.	494
20940	Akamai	414

Who controls the most email server name servers?

The number of domains of email servers hosted by a given country or network operator.

By country:

Country	MX/NS
United States	28,800
Netherlands	14,213
Ireland	11,440
Germany	5,380
Great Britain	3,116
France	2,996
Russia	2,588
Japan	1,663
Spain	1,421
Iran	1,123
Canada	933
China	842
Italy	798
Turkey	797

And by network operator:

ASN	Name	MX/NS
8075	Microsoft	11,596
13335	CloudFlare	6,790
16509	Amazon-1	2,018
16276	Ohn	1,969
26496	GoDaddy	1,750
24940	Hetzner	1,708
33517	DynDNS	1,523
36351	Softlayer	575
39572	Advancedh.	560
60781	Leaseweb	478
16552	Tiggee	475
49505	Selectel	433
63949	Linode	428
4837	China169	352

Evaluating the impact of potential attacks

Now we're in a position to evaluate the potential impact of three different attacks: distribution of malicious JavaScript content, email sniffing, and a DDoS attack against a core service provider. In each case a target can be selected by consulting the tables above.

Distributing malicious JavaScript content

The authors consider three ways to do this: – directly compromising (or colluding with) web servers hosting JS code; injecting malicious JavaScript when JS libraries are accessed over unprotected connections (HTTP instead of HTTPS); and redirecting requests for JS content via compromised name resolution.

Here we see the number of Alexa domains that can be reached via the first two of these:

Country/AS	Host. coll.	In-path Inj.
United States	15,658	12,267
Netherlands	3,292	2,639
Russia	1,701	1,409
Germany	1,622	1,317
Japan	1,311	1,151
China	1,141	1,079
Great Britain	1,094	895
Ireland	1,048	828
EU	905	824
France	713	603
Canada	399	246
Poland	176	151
Italy	105	97
Spain	83	69
15169 Google	9,469	5,553
13335 CloudFlare	4,310	3,165
15133 EdgeCast	3,404	2,306
16509 Amazon-1	3,216	2,264
20940 Akamai	2,279	1,800
14618 Amazon-2	572	351
35415 Webzilla	515	479
24940 Hetzner	379	330
16276 Ovh	342	287
36351 Softlayer	334	286
4837 China169	227	226
4134 ChinaNet	198	185
37963 Alibaba	148	146
54994 Quantil	71	71

The attack results show that countries can be very powerful attackers. For example, the United States hosts 47K JS hosting providers, which could distributed malicious code to about 16% of the top 100K Alexa domains. However, ASes are also very powerful and affect a fraction of websites that is even larger than than of individual countries, and even groups of countries. For example, the AS of Google can affect about 9% of Alexa domains.

When we look at JS inclusion over unprotected connections, 1,079 of them cross the Chinese network borders, but the United States, the Netherlands, Russia, Germany, and Japan all have even greater influence.

In *malicious name resolution redirection* the authoritative name server of a domain hosting JS redirects users to a malicious server. The attack result is the number of websites including a resource hosted on a server whose name server is colluding or compromised.

Country/AS	DNS redir.
United States	12,375
Russia	1,362
Netherlands	1,225
China	1,032
Japan	880
EU	743
Germany	621
France	454
Singapore	317
Great Britain	225
Spain	173
Iran	124
Canada	117
Poland	65
15169 Google	7,859
33517 DynDNS	4,311
16509 Amazon-1	3,685
13335 CloudFlare	3,012
4812 China Tlc	595
4837 China169	555
4808 China Unic	401
16552 Tiggee	361
26496 GoDaddy	316
24940 Hetzner	227
16276 Ovh	199
36351 Softlayer	196
20940 Akamai	88
15395 London Off	88

The United States, Google, and DynDNS stand out here.

Email sniffing

> *To acquire a large number of emails, an attacker can rely on various techniques. In this paper we consider two. The first one is by acquiring them directly from the email server. The second one is by redirecting an email client toward a malicious mail server, which will accept the email, keep a copy, and forward it to the intended recipient. This attack can be performed by a provider or by a country. Tables 3(c)and 3(d) show the attack results. All values are the number of Alexa domains that will be affected by this attack grouped by technique and attacker.*

Email sniffing by a malicious email provider:

Country/AS	MX coll.
United States	24,459
Germany	2,301
Great Britain	1,838
Russia	1,602
France	1,382
Japan	1,317
Netherlands	1,279
Ireland	809
Canada	614
India	496
Spain	410
Iran	392
Italy	384
Turkey	319
15169 Google	11,127
8075 Microsoft	2,465
26496 GoDaddy	1,267
16276 Ovh	565
24940 Hetzner	347
16509 Amazon-1	332
36351 Softlayer	237
60781 Leaseweb	170
12876 AS12876	134
46606 Unified L	113
63949 Linode	108
14618 Amazon-2	104
39572 Advancedh	96
32475 SingleHop	93

The United States alone can acquire emails for 25% of the most popular websites!

Malicious name resolution for email sniffing:

Country/AS	MX+NS
United States	13,077
Netherlands	3,933
Ireland	3,006
China	2,300
Germany	1,405
Great Britain	1,735
Russia	1,466
France	910
Japan	902
Iran	344
Spain	338
Canada	265
Italy	242
Turkey	213
8075 Microsoft	3,003
13335 CloudFlare	2,280
4837 China169	1,784
26496 GoDaddy	1,447
16509 Amazon-1	1,256
33517 DynDNS	1,178
16276 Ovh	555
24940 Hetzner	335
16552 Tiggee	227
36351 Softlayer	179
39572 Advancedh.	96
60781 Leaseweb	75
49505 Selectel	65
63949 Linode	57

Note how Google has much less power in this attack vector – most websites that use Google's email servers do so via name servers which are not hosted by Google.

DDoS against a core service provider

What happens if a service provider is the victim of an attack and is made unavailable? The data we already have can be used to figure this out. For example, consider the Dyn.com DoS attack from October 2016. DynDNS does not host a relevant number of mail servers and JS hosting providers, but it does host 364 domain servers.

> *These name servers are authoritative for 3,570 domains hosting JS that provide JS to 5,559 top 100K Alexa domains (not shown in Table 3), of which 4,331 are unprotected JS inclusion. Furthermore, the name servers hosted by DynDNS are authoritative for 1,523 domains running mail servers which are used by 1,178 top Alexa domains. If the Dyn.com DNS infrastructure is attacked, then a fraction that ranges from 1 to 5% of the top 100K Alexa domains would be affected.*

So who controls the Internet?

> *Our results show that already just a few players may have an extensive power: 14 countries and 14 autonomous systems can, directly or indirectly, affect the security of about 23% of websites… In addition, our results show that little has been learned from past attacks. For example, 70% of JavaScript (JS) inclusion is still done over unprotected connections, i.e., via HTTP URLs, which can be used to mount the Great Cannon attack.*

from → Uncategorized

No comments yet