Triplet Censors: Demystifying Great Firewall's DNS Censorship Behavior

Anonymous, Arian Akhavan Niaki†, Nguyen Phong Hoang‡,

Phillipa Gill†, Amir Houmansadr†

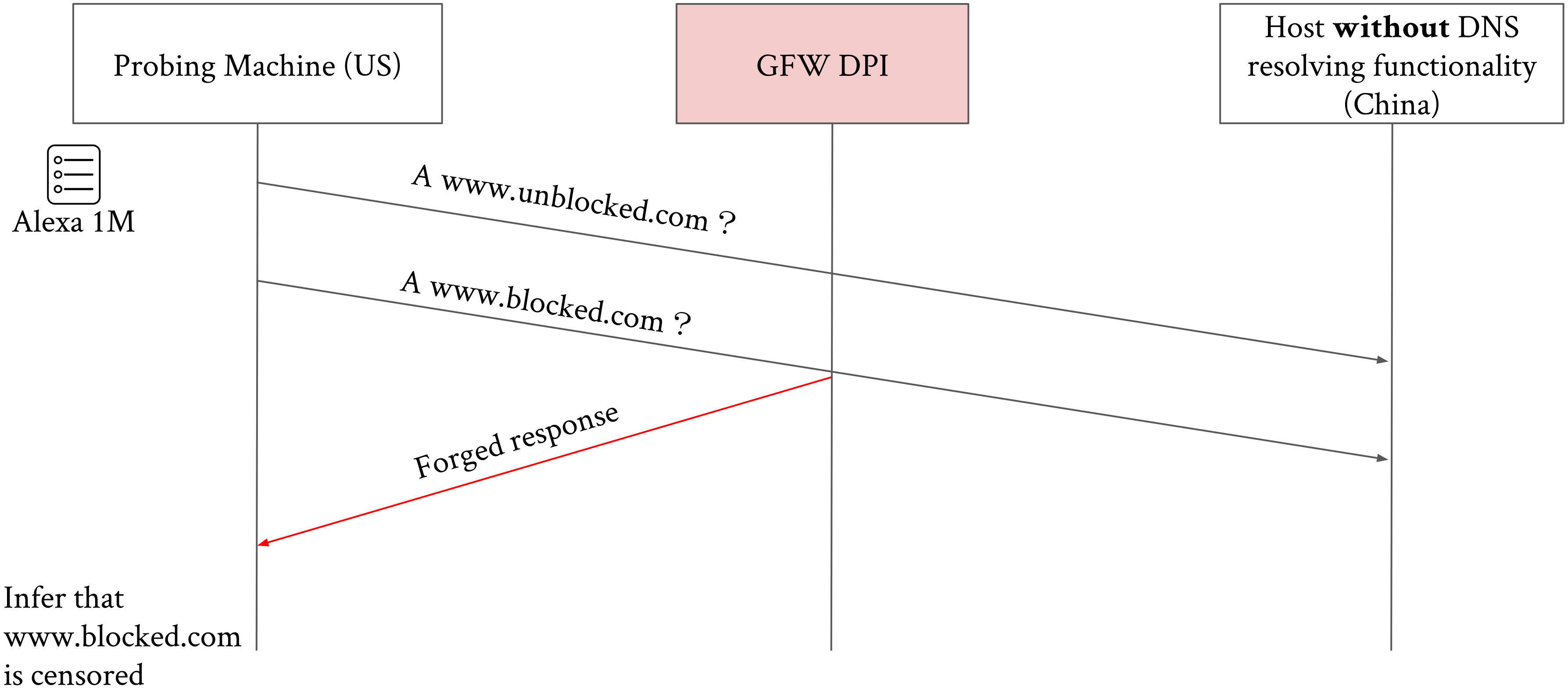
†University of Massachusetts Amherst, ‡Stony Brook University

Overview

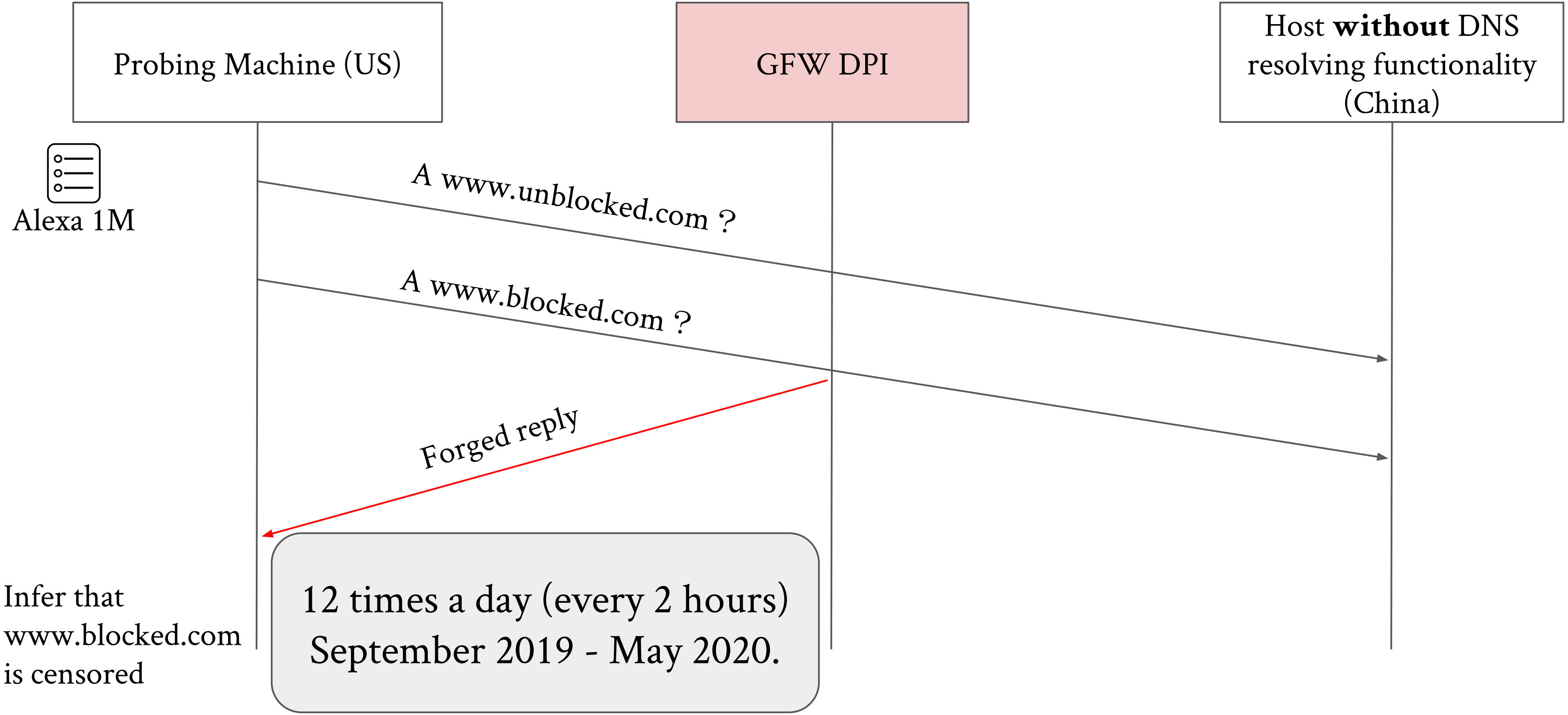
Questions about the DNS filtering of the Great Firewall of China

- What domains are blocked?
- What are the IPs used in the forged DNS responses?
- How are domains being blocked?
- Is the blocking consistent within China?

Methodology


Probing Machine (US)
GFW DPI
Host **without** DNS resolving functionality (China)
Alexa 1M
A www.unblocked.com ?
A www.blocked.com ?
Forged response
Infer that www.blocked.com is censored

Longitudinal Dataset



Probing Machine (US)

GFW DPI

Host **without** DNS resolving functionality (China)

Alexa 1M

A www.unblocked.com ?

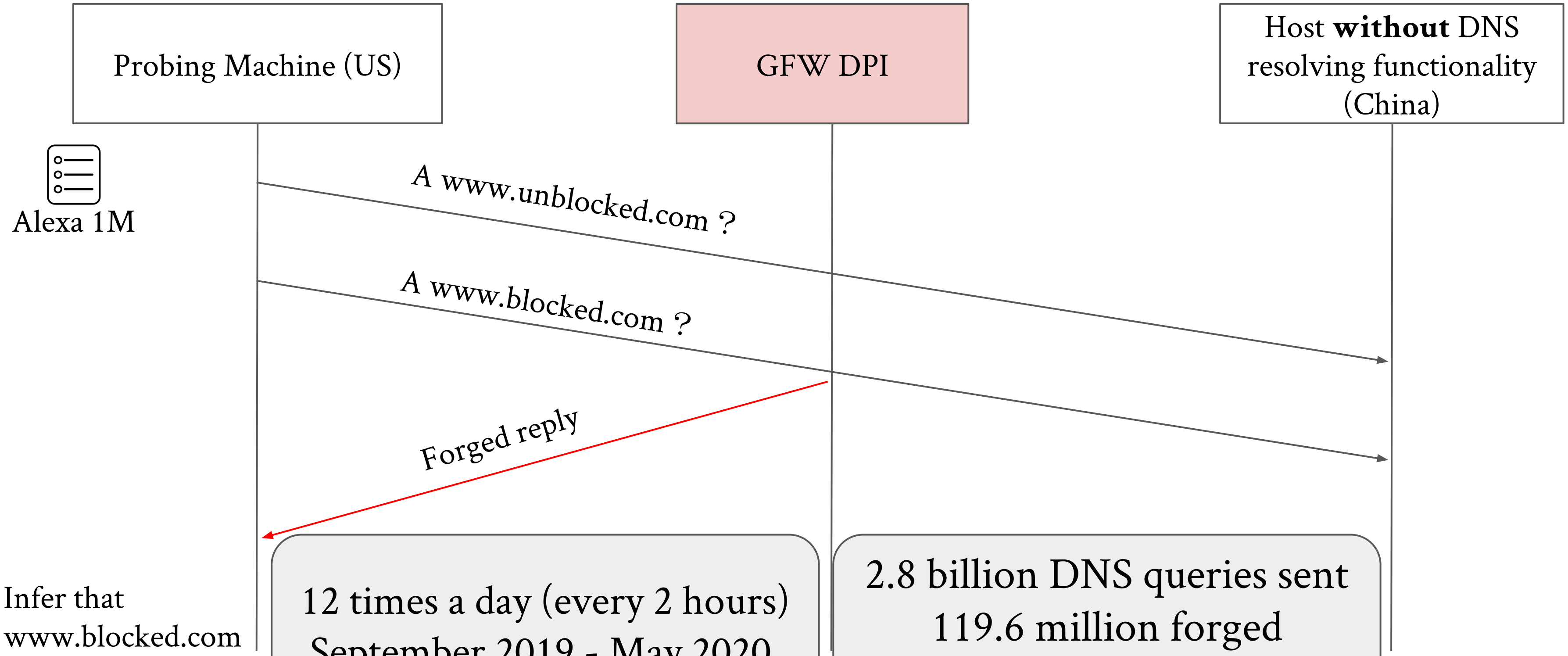
A www.blocked.com ?

Forged reply

Infer that www.blocked.com is censored

12 times a day (every 2 hours) September 2019 - May 2020.

Longitudinal Dataset



Probing Machine (US)

GFW DPI

Host **without** DNS
resolving functionality
(China)

Alexa 1M

A www.unblocked.com ?

A www.blocked.com ?

Forged reply

Infer that
www.blocked.com
is censored

12 times a day (every 2 hours)
September 2019 - May 2020.

2.8 billion DNS queries sent
119.6 million forged
responses

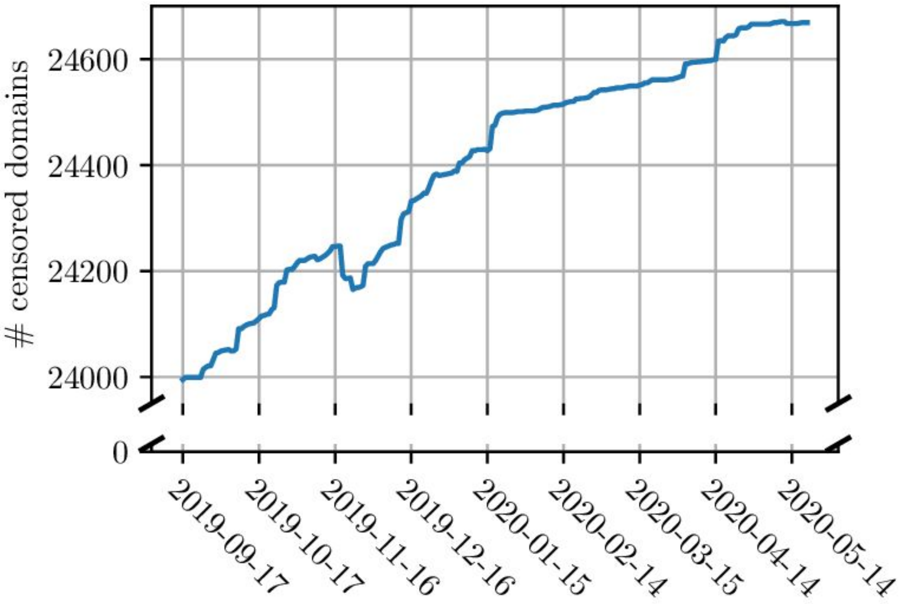
Overview

Questions about the DNS filtering of the Great Firewall of China

- What domains are blocked?
- What are the IPs used in the forged DNS responses?
- How are domains being blocked?
- Is the blocking consistent within China?

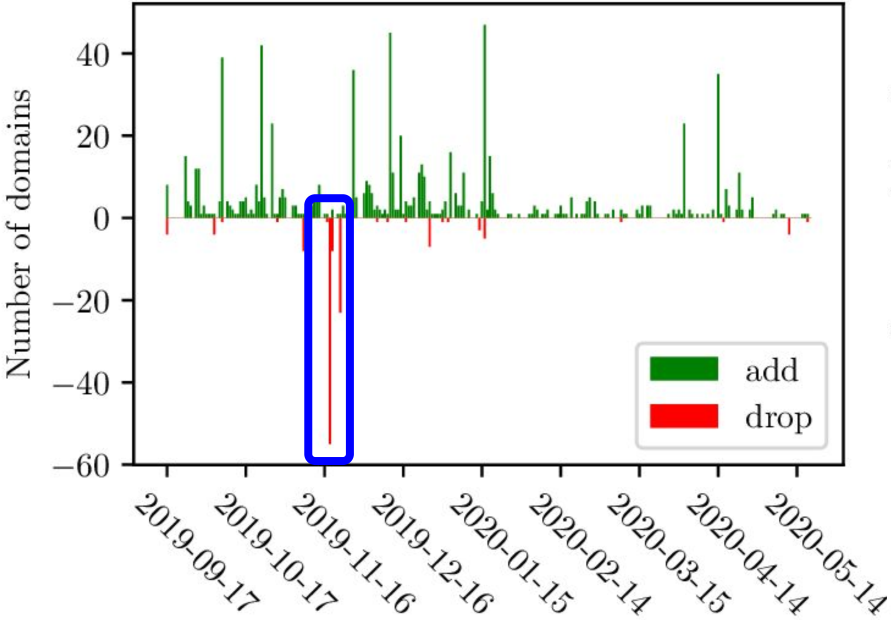
What domains are blocked

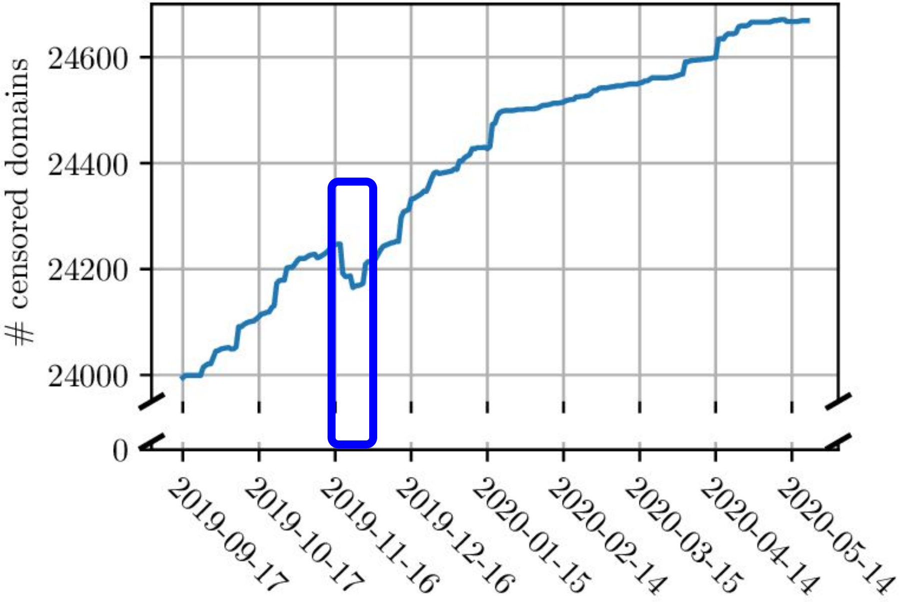
- Number of censored websites increases from 23,995 to 24,636



What domains are blocked

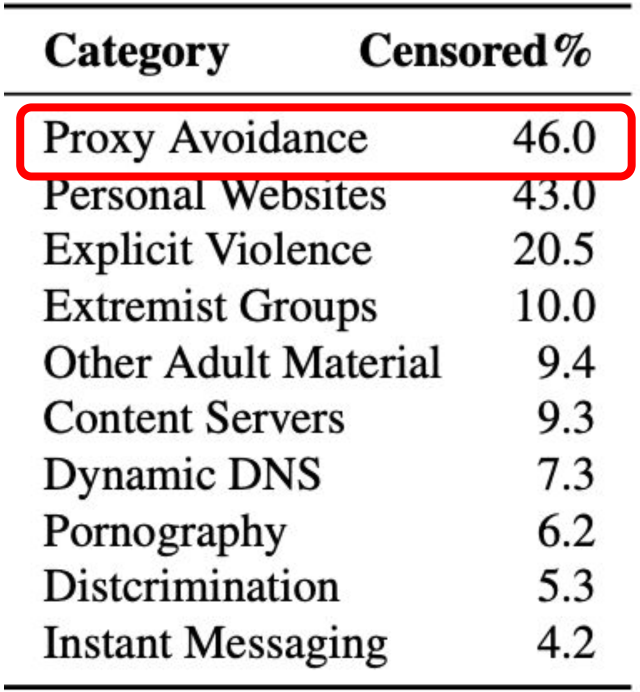
- Number of censored websites increases from 23,995 to 24,636
- A major drop partly due to a rule change: “*youtube.com” -> “*.youtube.com”





What domains are blocked - Categories

- What types of domains are mostly censored?

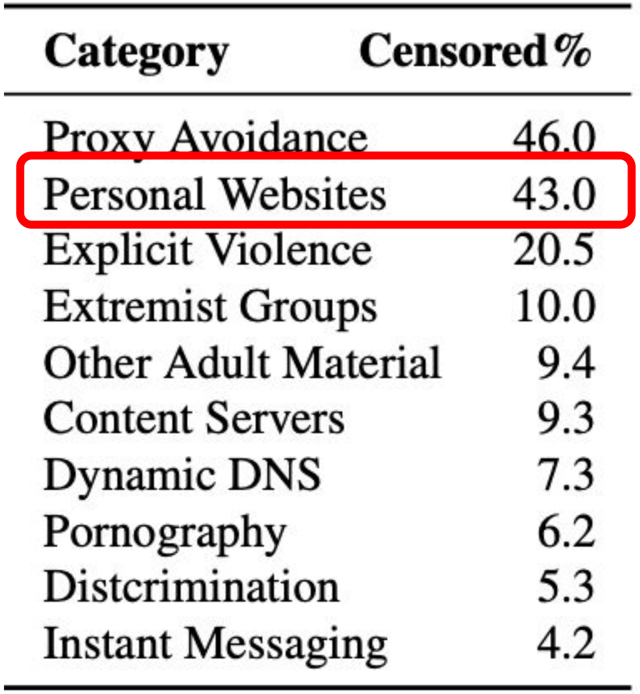


Category	Censored %
Proxy Avoidance	46.0
Personal Websites	43.0
Explicit Violence	20.5
Extremist Groups	10.0
Other Adult Material	9.4
Content Servers	9.3
Dynamic DNS	7.3
Pornography	6.2
Distcrimination	5.3
Instant Messaging	4.2

www.purevpn.com
www.hideipvpn.com
www.hideip.co
www.anonymizer.com

What domains are blocked - Categories

- What types of domains are mostly censored?



Category	Censored %
Proxy Avoidance	46.0
Personal Websites	43.0
Explicit Violence	20.5
Extremist Groups	10.0
Other Adult Material	9.4
Content Servers	9.3
Dynamic DNS	7.3
Pornography	6.2
Distcrimination	5.3
Instant Messaging	4.2

*.blogspot.com

*.tumblr.com

Overview

Questions about the DNS filtering of the Great Firewall of China

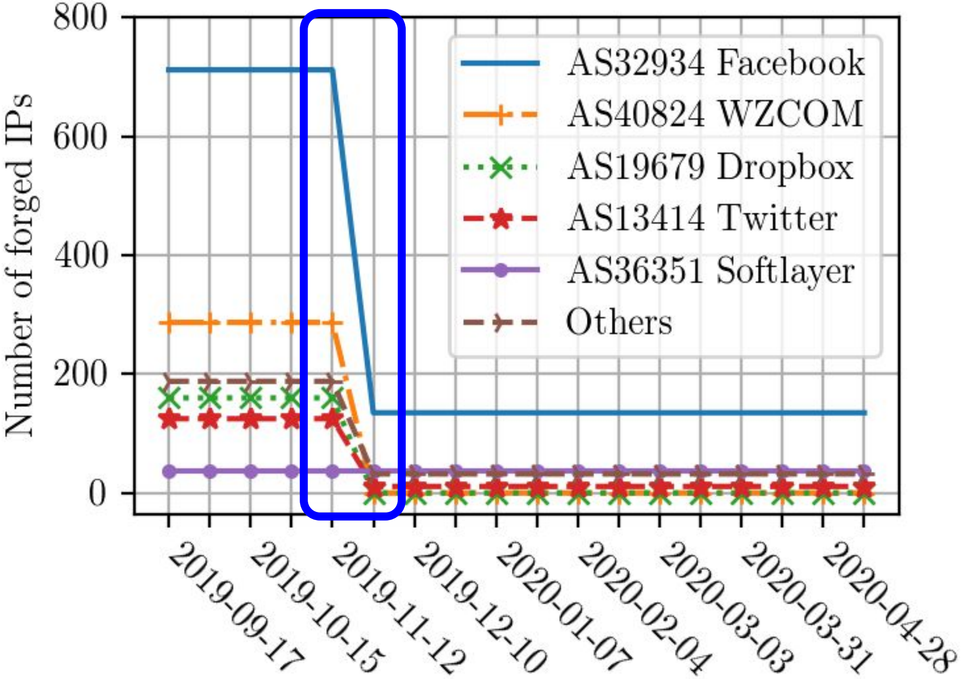
- What domains are blocked?
- What are the IPs used in the forged DNS responses?
- How are domains being blocked?
- Is the blocking consistent within China?

IPs used in forged DNS responses

- How do these IPs change?
- Where do these IPs belong to?

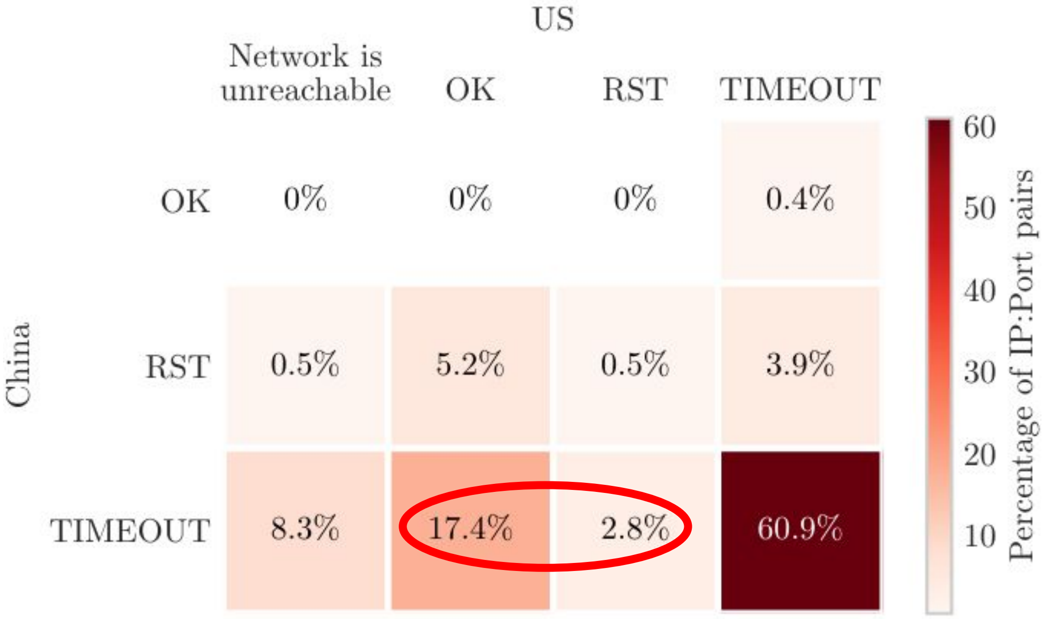
IPs used in forged DNS responses

- How do these IPs change?
- Where do these IPs belong to?
- Drop on November 23, 2019
  - Before 1,510 IPs (41 ASes)
  - After 216 IPs (21 ASes)



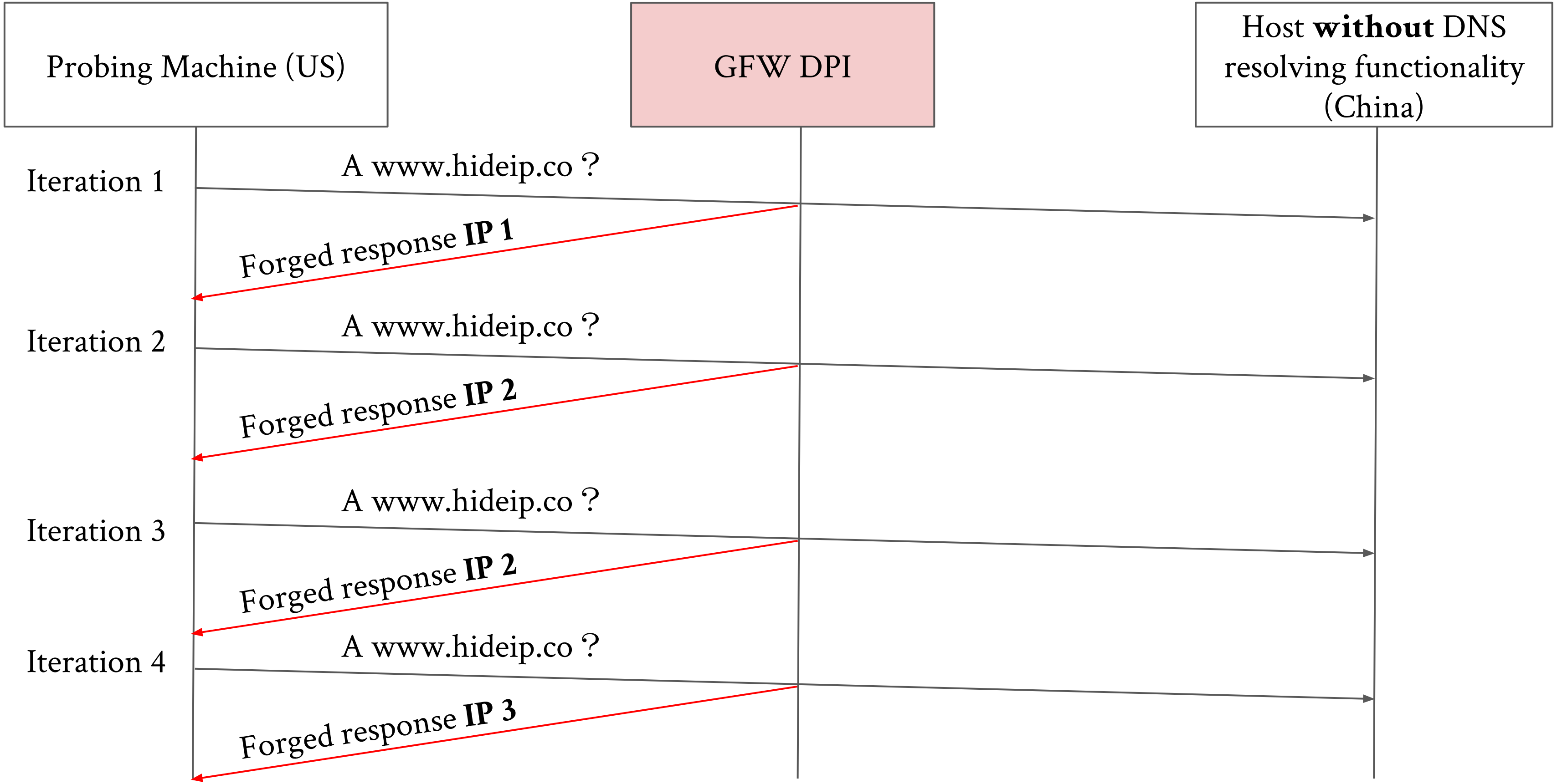
IPs used in forged DNS responses

- Reachability of the 216 currently injected IPs over a week
- Connection scans for each IP
  - Port 80 and 443



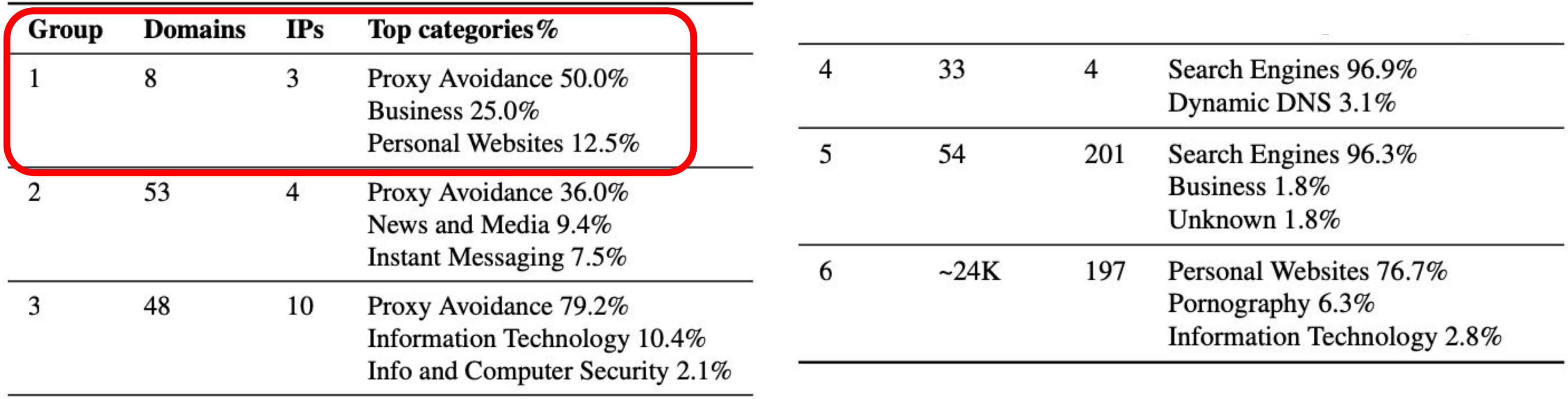
China \ US	Network is unreachable	OK	RST	TIMEOUT
OK	0%	0%	0%	0.4%
RST	0.5%	5.2%	0.5%	3.9%
TIMEOUT	8.3%	17.4%	2.8%	60.9%

IPs used in forged DNS responses



IPs used in forged DNS responses

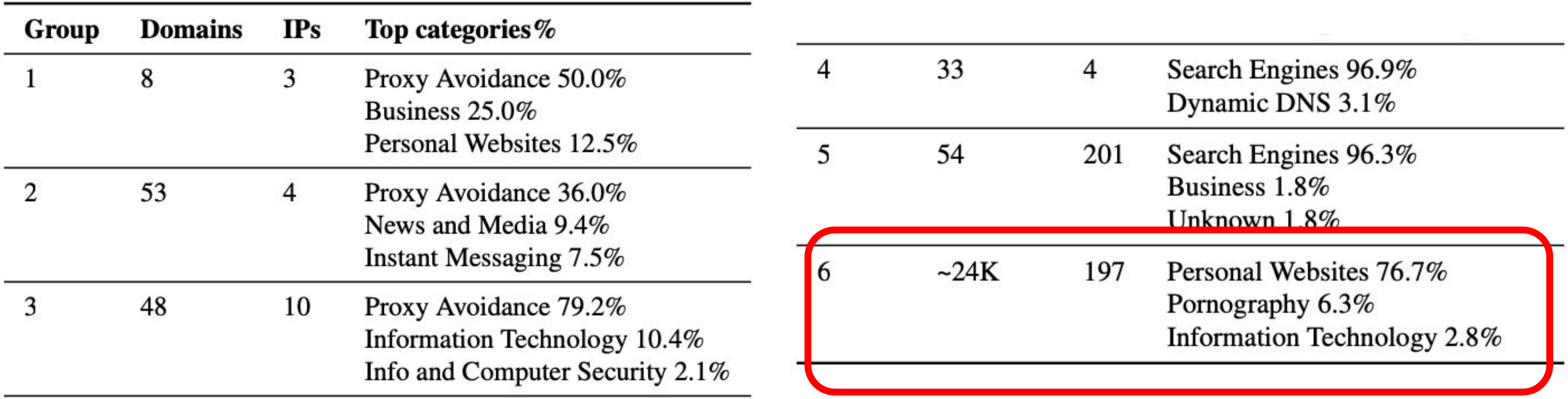
- GFW injects different set of IPs to censor different set of domains



Group	Domains	IPs	Top categories%
1	8	3	Proxy Avoidance 50.0% Business 25.0% Personal Websites 12.5%
2	53	4	Proxy Avoidance 36.0% News and Media 9.4% Instant Messaging 7.5%
3	48	10	Proxy Avoidance 79.2% Information Technology 10.4% Info and Computer Security 2.1%
4	33	4	Search Engines 96.9% Dynamic DNS 3.1%
5	54	201	Search Engines 96.3% Business 1.8% Unknown 1.8%
6	~24K	197	Personal Websites 76.7% Pornography 6.3% Information Technology 2.8%

Characterizing GFW's DNS Injection

- GFW injects different set of IPs to censor different set of domains



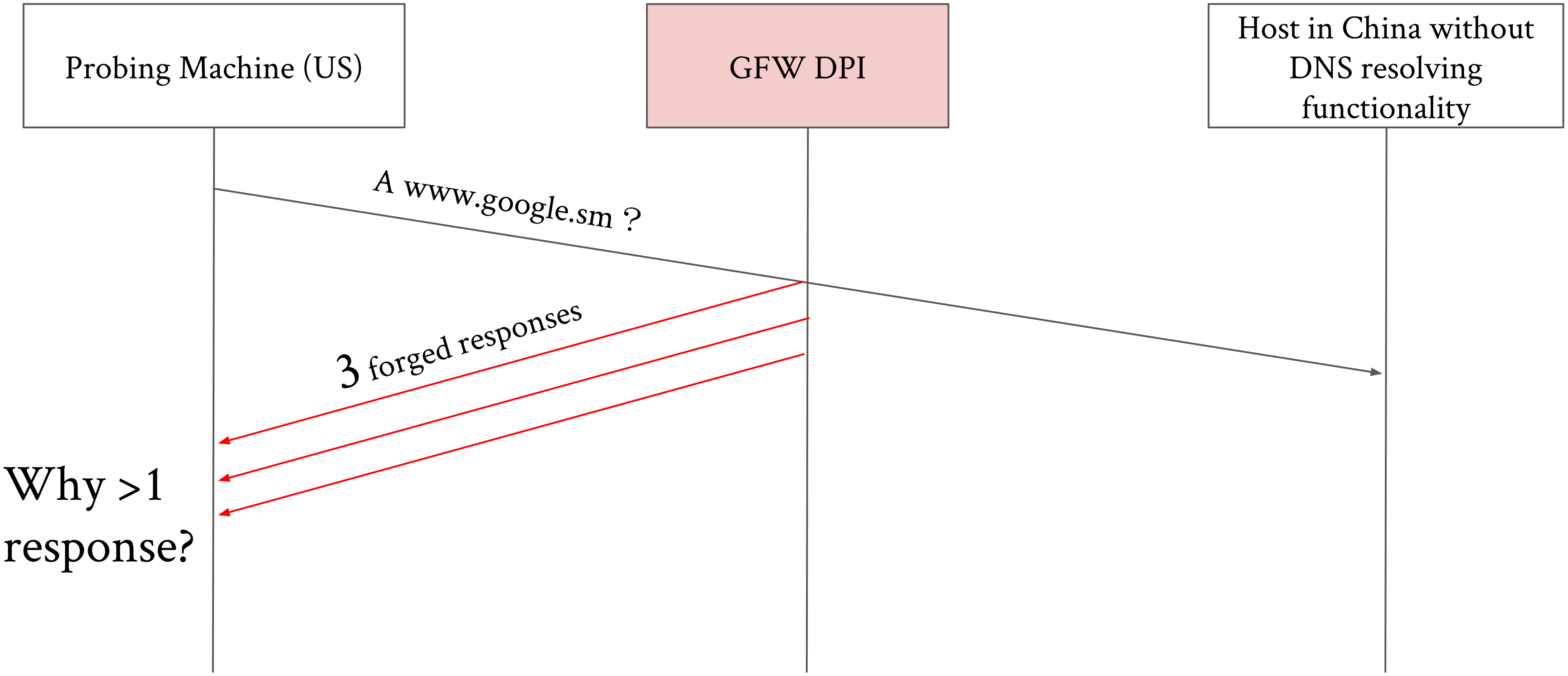
Group	Domains	IPs	Top categories%
1	8	3	Proxy Avoidance 50.0% Business 25.0% Personal Websites 12.5%
2	53	4	Proxy Avoidance 36.0% News and Media 9.4% Instant Messaging 7.5%
3	48	10	Proxy Avoidance 79.2% Information Technology 10.4% Info and Computer Security 2.1%
4	33	4	Search Engines 96.9% Dynamic DNS 3.1%
5	54	201	Search Engines 96.3% Business 1.8% Unknown 1.8%
6	~24K	197	Personal Websites 76.7% Pornography 6.3% Information Technology 2.8%

Overview

Questions about the DNS filtering of the Great Firewall of China

- What domains are blocked?
- What are the IPs used in the forged DNS responses?
- How are domains being blocked?
- Is the blocking consistent within China?

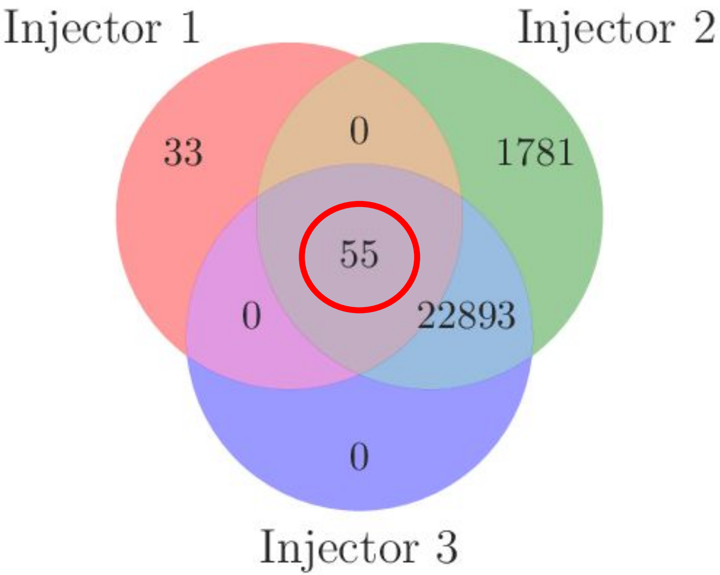
How are domains being blocked



Why >1 response?

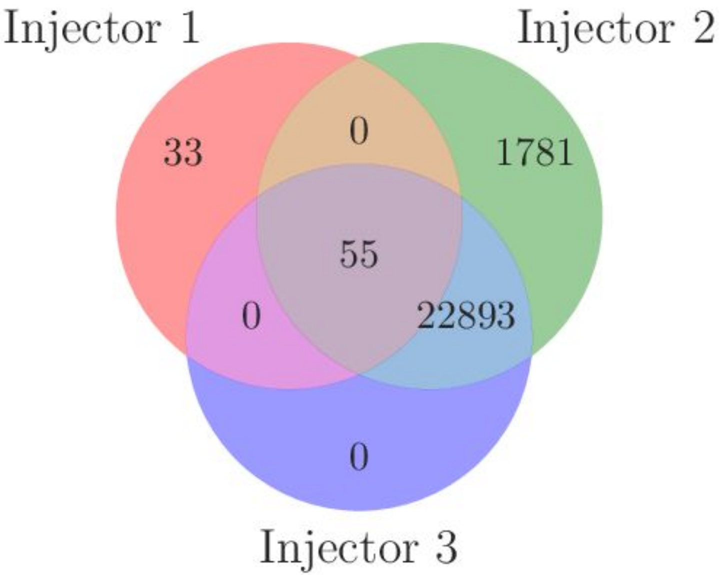
How are domains being blocked

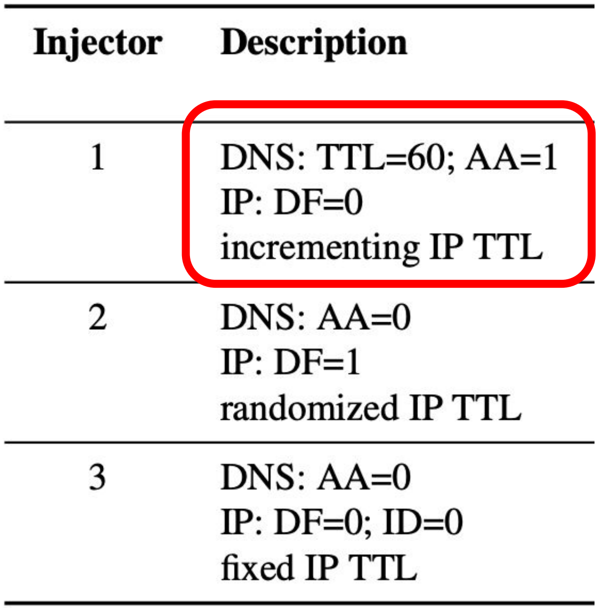
- Each injector maintains a different blacklist



How are domains being blocked

- Each injector maintains a different blacklist
- Each injector has a unique fingerprint

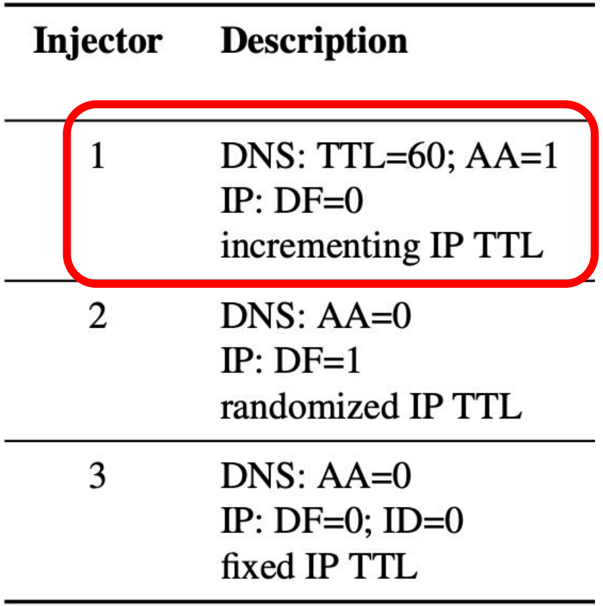




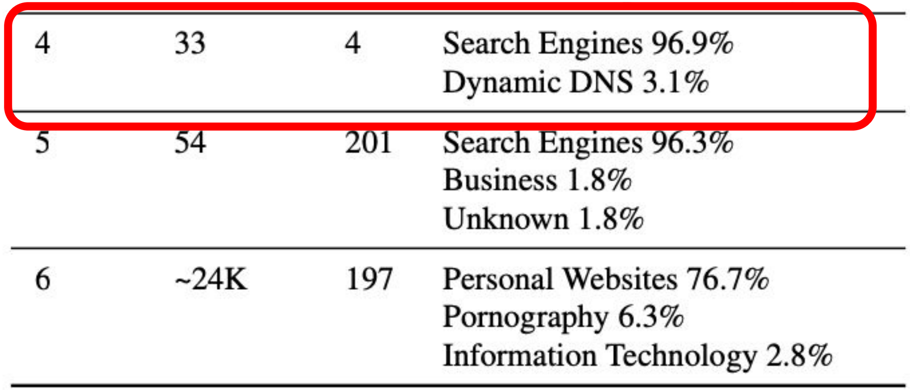
Injector	Description
1	DNS: TTL=60; AA=1 IP: DF=0 incrementing IP TTL
2	DNS: AA=0 IP: DF=1 randomized IP TTL
3	DNS: AA=0 IP: DF=0; ID=0 fixed IP TTL

How are domains being blocked

- Relation between IP/Domain groups and the injectors



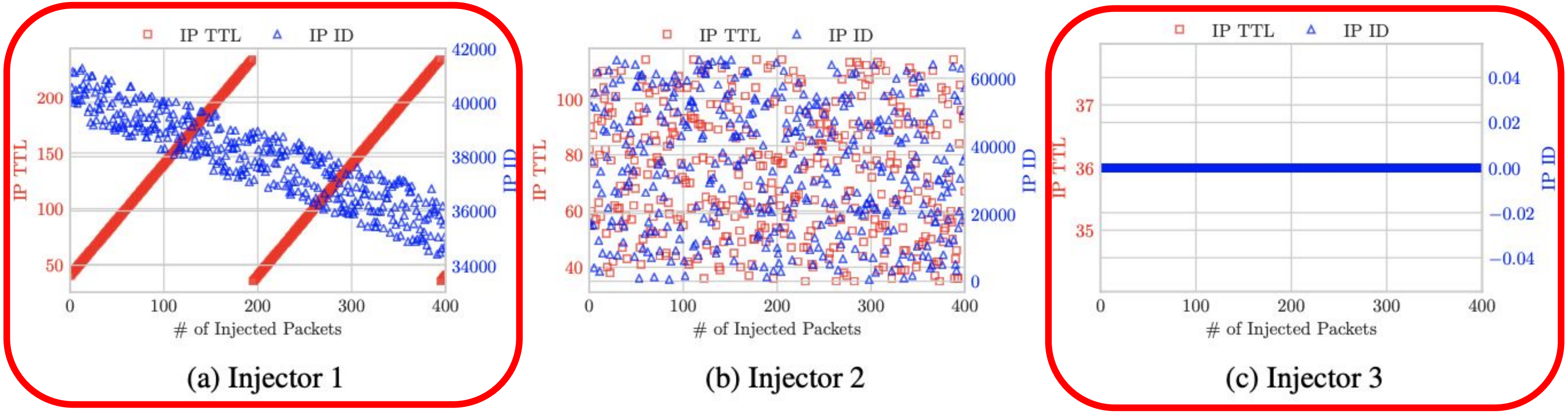
Injector	Description
1	DNS: TTL=60; AA=1 IP: DF=0 incrementing IP TTL
2	DNS: AA=0 IP: DF=1 randomized IP TTL
3	DNS: AA=0 IP: DF=0; ID=0 fixed IP TTL



4	33	4	Search Engines 96.9% Dynamic DNS 3.1%
5	54	201	Search Engines 96.3% Business 1.8% Unknown 1.8%
6	~24K	197	Personal Websites 76.7% Pornography 6.3% Information Technology 2.8%

Fingerprinting the GFW Injectors

- IPID and IP TTL patterns under when sending queries rapidly

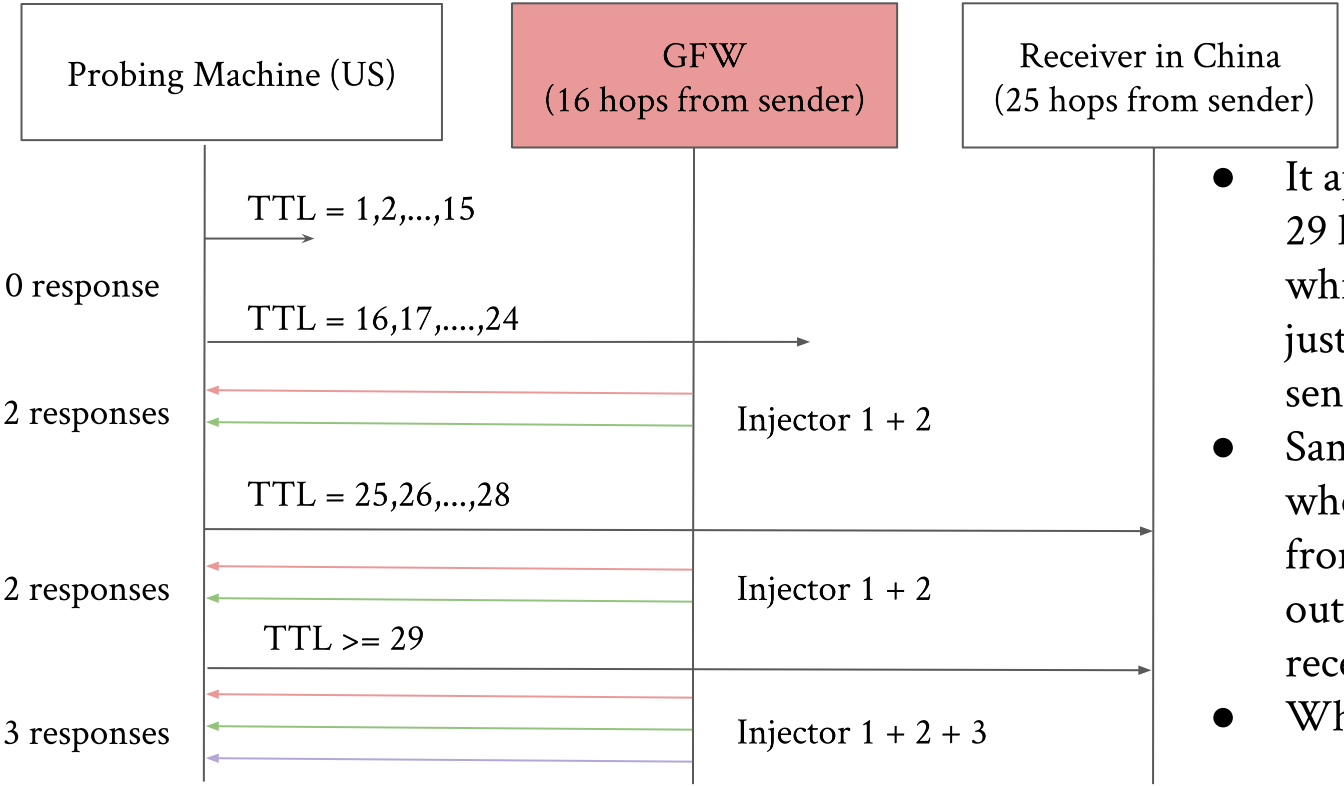


(a) Injector 1

(b) Injector 2

(c) Injector 3

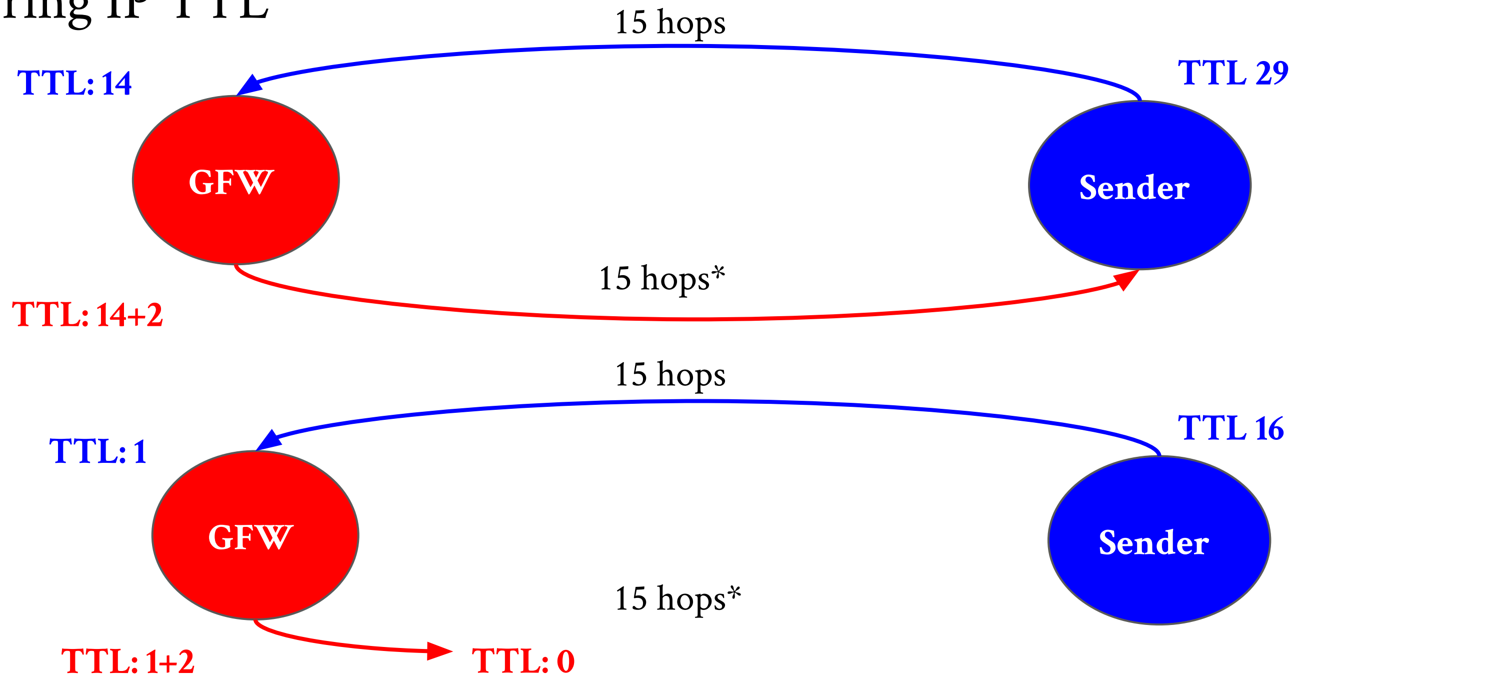
Localizing the GFW Injectors



- It appear one of the injectors is 29 hops away from the sender, while the receiver is actually just 25 hops away from the sender
- Same strange results remain when repeating the experiment from 7 different locations outside of China to the same receiver.
- Why?

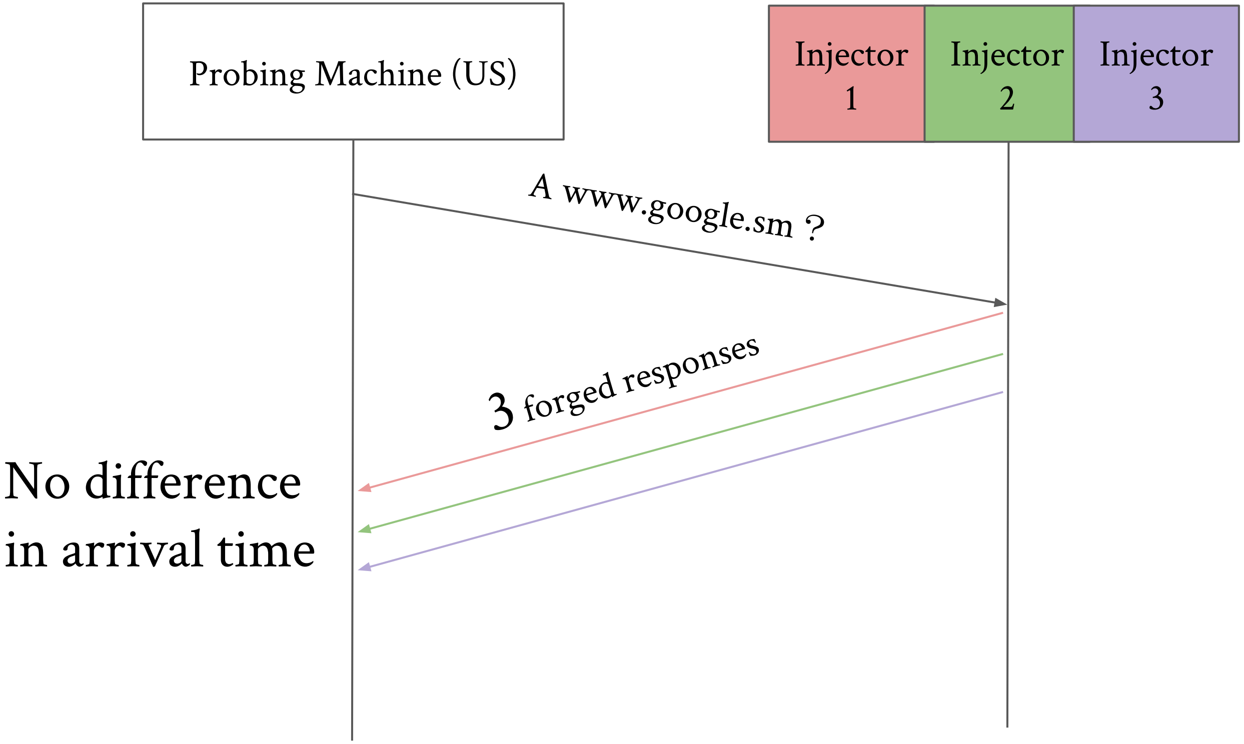
Localizing the GFW Injectors

- Mirroring IP TTL



*Assuming that the routing paths are symmetric

Localizing the GFW Injectors



Overview

Questions about the DNS filtering of the Great Firewall of China

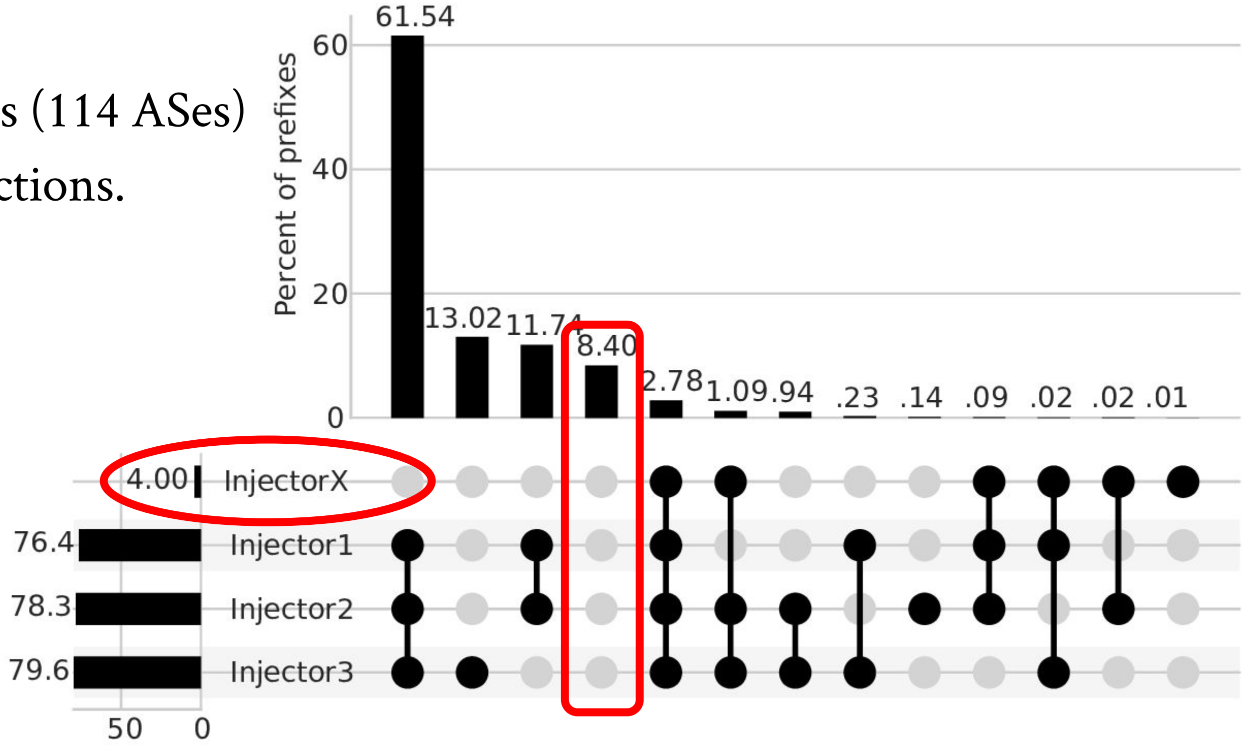
- What domains are blocked?
- What are the IPs used in the forged DNS responses?
- How are domains being blocked?
- Is the blocking consistent within China?

Is the blocking consistent within China

1. Select 36,629 IP prefixes belonging to Chinese organizations from CAIDA
2. Select one non-responding IP for each prefix at random
    a. In total, we get 36,146 non-responding Chinese IPs (417 ASes)
3. Issue 100 sensitive queries for www.google.sm to all selected IPs from one single point outside of China.

Is the blocking consistent within China

- Only 8.4% of prefixes (114 ASes) receive no DNS injections.



Summary

- The GFW injects different sets of IPs to censor different groups of domains
- We have fingerprinted 3 GFW injectors
    - All of them appear to share the same injection point
    - Injector 1's IPID and IP TTL are associated with injection sequence
    - Injector 3's IP-TTL echoing behavior has implications on using TTL-limited probe packets to localize GFW injectors
- Observed DNS injections on 91.6% of the 36K Chinese IP prefixes

We have released all our code and datasets at
https://gfw.report/publications/foci20_dns/en/

gfw.report@protonmail.com (B0C6 EB19 DA7C EAA3)