Triplet Censors: Demystifying Great Firewall's DNS Censorship Behavior

Anonymous

Arian Akhavan Niaki
University of Massachusetts Amherst

Nguyen Phong Hoang
Stony Brook University

Phillipa Gill
University of Massachusetts Amherst

Amir Houmansadr
University of Massachusetts Amherst

Abstract

The Great Firewall of China (GFW) has long used DNS packet injection to censor Internet access. In this work, we analyze the DNS injection behavior of the GFW over a period of nine months using the Alexa top 1M domains as a test list. We first focus on understanding the publicly routable IPs used by the GFW and observe groups of IPs used to filter specific sets of domains. We also see a sharp decline in public IPs injected by the GFW in November 2019. We then fingerprint three different injectors that we observe in our measurements. Notably, one of these injectors mirrors the IP TTL value from probe packets in its injected packets which has implications for the use of TTL-limited probes for localizing censors. Finally, we confirm that our observations generally hold across IP prefixes registered in China.

1 Introduction

Many countries are known to use injection of DNS responses to implement censorship [3, 8, 15, 21, 28] with China's use of DNS injection in the Great Firewall (GFW) being a popular topic for study [1, 2, 10, 11, 14, 16–18, 22, 26, 30]. While other countries tend to use NXDOMAIN or reserved IP address space [3,4,8,20], China's use of a range of public IP addresses owned by a variety of organizations is notable. This use of public IP addresses can complicate detection of DNS-based censorship in China [5, 12, 21] and can make evading inadvertent DNS cache poisoning by the GFW challenging [10, 26].

While there have been numerous studies of China's DNS censorship [1, 2, 10, 11, 14, 16–18, 22, 26] (owing in part to the fact that the GFW will inject replies to clients outside of the country), in this study, we take a longitudinal approach focusing on China's use of public IPs for filtering. We measure China's DNS injector for a period of nine months which allows us to observe changes in the set of public IP addresses used by the GFW (§2). We further perform targeted measurements to fingerprint the behavior of the GFW's DNS packet injector and consider the generalizability of our results across 36K prefixes announced by Chinese ASes (§5).

Our study reveals several previously-unknown properties of China's filtering system:

IP groups. First, we observe groups of IP addresses that are used in injected replies to specific sets of domains (§3). These groups may point to groups of domains that are being blocked by a common infrastructure or blocking process. We discuss these groups in the context of blocked domains and IPs used for blocking over time (§3.2)

Three distinct injectors. We also observe that a single DNS query can result in multiple injected DNS replies from the GFW. Using IP ID, IP TTL, DNS TTL and DNS flags, we were able to fingerprint these multiple replies and identify three distinct packet injectors acting on DNS requests (§4.1).

TTL-echoing in injected packets. In the process of fingerprinting the censors, we observe one of the packet injectors will actually echo the TTL of the probe packet which has implications on the popular technique of using TTL-limited probe packets to localize network censors (§4.3).

2 Methodology

We now describe our methodology for monitoring DNS-based censorship in China on a longitudinal basis (§2.1) and how we extend this method to understand regional differences in filtering (§2.2). We also discuss steps taken to address ethical concerns while conducting our experiment (§2.3).

2.1 Baseline Longitudinal Experiment

We use the commonly employed tactic of issuing DNS queries for potentially sensitive domains from a host outside of China towards IP addresses located in China (specifically, those not hosting DNS servers). This allows us to trigger the GFW as our packet crosses the GFW, and the targeting of IP addresses not hosting DNS servers means that any response to our query can be inferred to be injected by the GFW. We issue queries from a Virtual Private Server (VPS) running Ubuntu 18.04 LTS located in a US academic network. We then send DNS queries towards a VPS under our control located in China with the same configuration as our US host. We perform our queries using the standard DNS port (53). We performed an initial test over ports 1-65535 and only observed censorship on DNS queries sent on port 53.

With this source and destination host, we then issue DNS queries for a set of tested domains. In our case, a set of 1 million domains is extracted from the Alexa top million Web sites list (accessed on Feb. 22, 2019). For any domains without the prefix "www." we add this prefix as the GFW does not consistently inject DNS replies in the absence of this prefix [1, 9]. We query these domains every two hours between September 2019 and May 2020. In total, we sent 2.8 billion DNS queries and observed 119.6 million forged responses from the GFW.

2.2 Multi-path Experiment

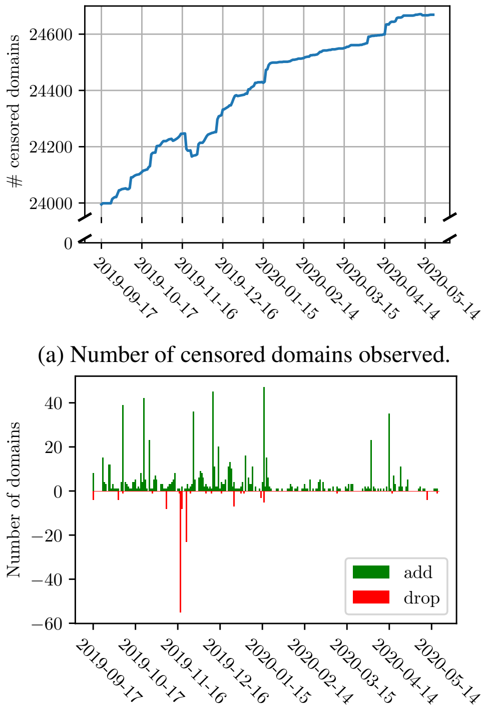
A limitation of our baseline methodology, is that we will only observe filtering on the path between our VPS in the US and our VPS in China. To complement this methodology, we perform an additional experiment where we direct DNS queries towards a broad range of Chinese IP prefixes. We identify Chinese IP prefixes by using CAIDA's AS-to-organization dataset [6] to identify ASNs registered in China. We then use CAIDA's prefix-to-AS mapping tool [7] to collect IP prefixes announced by these ASes, for a total of 36,629 prefixes.

Within each prefix, we select one IP address at random, ensuring that there is not a host at this IP address that will respond to DNS queries. To determine this, we send 10 queries for a non-sensitive domain `www.baidu.com` to the candidate IP address. If there is no reply to any of our DNS queries, we infer that this IP is not hosting a DNS server and proceed with our tests. We exclude an IP prefix from testing if we fail to find a non-responding IP address after 50 attempts. In total, we select 36,146 IP prefix, belonging to 417 Chinese ASes.

For this test, we focus on a single domain `www.google.sm` that we observe triggers censorship by the three packet injectors observed in our baseline experiment (§4) since our goal is to understand the behavior of multiple network paths. We attempt 100 queries for this domain towards each of the Chinese prefixes we identify.

2.3 Ethics

For our baseline experiment, the two hosts that we sent DNS queries to and from are machines under our control. For our multi-path experiment, we first verify that no DNS service was running on the selected IP address. We also note that our experiments are initiated from a host outside of China, thus to the GFW it appears that queries are coming from an external (academic) network, as opposed to any host within China. Finally, our multi-path experiment limits the amount of traffic sent to each IP address to 1 MB.



(a) Number of censored domains observed.

(b) Number of censored domains added and dropped per day.

Figure 1: Censored domain name changes among Alexa top 1 million from September 2019 to May 2020.

3 Characterizing DNS Injection

In this section, we characterize domains filtered over time (§3.1) as well as the IP addresses in the injected replies (§3.2).

3.1 Censored Domains

We see that there exists an increasing trend in the number of domains being censored by the GFW. The number of censored domains increases from 23,995 to 24,636 (2.8% increase) over our nine-month measurement study. Figure 1a presents the number of unique domains censored over time. Interestingly, previous work [1] has also shown a 10% increase in the number of censored domains over time in their 2014 study (also using the Alexa top million as their test domains).

Figure 1b depicts the daily number of domains from the Alexa top 1 million that get added and removed from the set of domains that we observe being blocked. We manually analyzed the dates in which more than 20 domains were removed from blocked set, on November 18 a group of 50 domains that all have the keyword `youtube.com` were removed and on November 22 a group of 22 domains with the keyword `line.me` were removed from the blocked set. This suggests that the GFW still operates on keywords to censor domains as opposed to curating a fixed set of domains.

Category of censored domains. We leveraged the "FortiGuard" URL classification service, operated by FortiNet [13]

Category	Alexa%	Category	Censored%
Business	27.7	Proxy Avoidance	46.0
Information Technology	13.3	Personal Websites	43.0
Shopping	5.9	Explicit Violence	20.5
Education	5.7	Extremist Groups	10.0
Personal Websites	4.4	Other Adult Material	9.4
News and Media	4.1	Content Servers	9.3
Entertainment	3.5	Dynamic DNS	7.3
Pornography	2.8	Pornography	6.2
Health and Wellness	2.7	Distcrimination	5.3
Government and Legal Orgs	2.6	Instant Messaging	4.2

Table 1: FortiGuard Categories. The 10 most common categories for the domains on Alexa 1M test list, and the percentage of censored domains in each category.

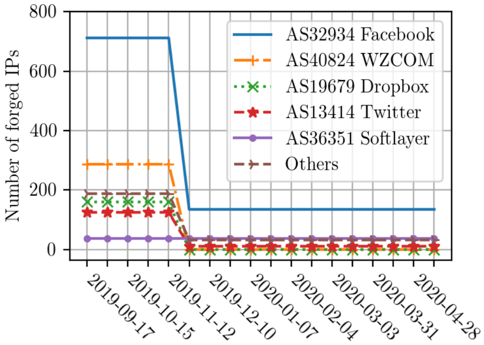


Figure 2: Top ASNs and the number of injected IP addresses used by the GFW belonging to each of them.

to categorize the Alexa top domains. The top categories within the Alexa list are listed in the left column of Table 1. We further analyze the percentage of censored domains in each category of the Alexa top million list. The top 10 categories with the highest percentage of domains censored are shown in the right column of Table 1. We can see that 46% of the domains in the "Proxy Avoidance" category are censored by the GFW. The high number (42.9% of domains censored) for the "Personal Websites" category is because 42.7% of the censored domains within the "Personal Websites" category are domains containing the keywords .blogspot.com, or .tumblr.com which appear to be filtered by the GFW. We further analyzed and found that this is in fact a keyword based block list, i.e any domain that ends in .blogspot.com or .tumblr.com will be censored by the GFW.

3.2 Injected IPs

Longitudinal trends. We observe a set of 1,510 distinct IP addresses returned in type A DNS records injected by the GFW. While the majority of responses we observe are type A DNS records, we observe injected CNAME records for a single domain (www.sunporno.com). We focus on the type A records in this paper and plan to dig into the use of CNAME records by the GFW in future work.

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%

Table 2: Overview of the relationship between the sensitive domain, forged IP groups and injectors after the decrease in the number of injected IP addresses.

Figure 2 shows the top ASes associated with the IPs injected by the GFW. We observe a total of 41 ASes associated with the injected IP addresses. Most of these ASes correspond to organizations in the US, particularly Facebook, WZCOM, Dropbox and Twitter. We note a striking decrease in the number of distinct IPs injected by the GFW on November 23, 2019 from 1,510 IPs (associated with 41 ASes) to only 216 IPs (associated with 21 ASes). We investigate this drop in injected IPs further in Section 4.

Groups of injected IPs. One property of the injected IPs that we note, is that certain subsets of blocked domains resolve to a fixed set of public IPs. That is, a group of public IPs is used to filter a given group of censored domains. Table 2 depicts the six distinct groups of domains we identified. We further categorized the domains in each group. The top category of domains in group 1, 2, and 3 belong to the "Proxy Avoidance" category, while 97% of the domains from group 4 and 5, include the word google, belonging to the "Search Engines" category. Group 6 consists of the remaining websites that are censored on the Alexa 1M that are mostly blogspot and tumblr related websites. We analyzed the IPs that were dropped from the IP pool on November 23 and found that 99% of the domains that received those IPs currently receive 197 injected IPs (Group 6), the majority (99%) of these domains have the keyword tumblr.com in them.

Reachability of the injected IP addresses. Given that China is using publicly routable IP addresses, a natural question is whether these IPs are hosting content or are otherwise reachable on the broader Internet. We test the reachability of the injected IPs from our VPS in China and the United States by initiating TCP handshakes on port 80 and port 443. We

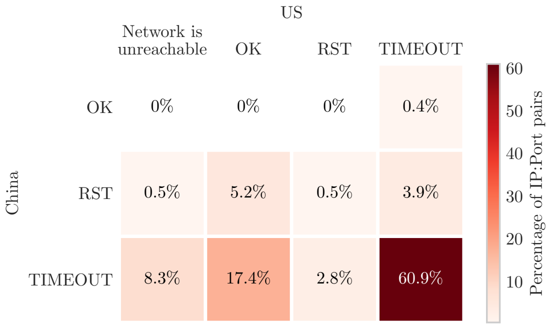


Figure 3: Reachability of the ports 80 and 443 of the injected IPs from China and from the US. The numbers are averaged over seven days.

Injector	Description	IPs	Domains	IP Group
1	DNS: TTL=60; AA=1 IP: DF=0 incrementing IP TTL	4	88	4, 5, 6
2	DNS: AA=0 IP: DF=1 randomized IP TTL	1,506	24,729	1, 2, 3 5, 6
3	DNS: AA=0 IP: DF=0; ID=0 fixed IP TTL	958	22,948	1, 2, 3, 5

Table 3: Summary of the three DNS injectors. "DNS AA" refers to the DNS Authoritative Answer flag. "IP DF" refers to the IP "do not fragment" flag.

perform this experiment daily for 7 days and present the averaged result in Figure 3. We note that each days results looked similar. In the majority of cases (60.9%), the TCP handshake attempt results in a TIMEOUT both for source hosts in the US and China, indicating there is likely no content being served from these IPs at the time of our measurements. It is possible these IPs were observed serving content at some point in the past which resulted in their addition to the set of injected IPs.

4 Understanding the GFW Injectors

We now characterize cases where multiple injected DNS replies are observed. We are able to fingerprint these replies and identify three distinct injection processes (§4.1). We characterize longitudinal trends of the injectors (§4.2). Finally, We also localize these injectors and observe peculiar mirroring of the probe-TTL value by one injector (§4.3).

4.1 Fingerprinting the Injectors

In our measurements, we observed cases where a single DNS query may result in multiple injected DNS replies. Upon closer inspection, we were able to identify three distinct fingerprints within these multiple injected replies based on IP Do-not-Fragment (DF), IP TTL , DNS Authoritative Answer (AA), and DNS TTL fields. Table 3 summarizes the fingerprints of the three injectors and Figure 4 plots the IPID and TTL values for these three injectors when queries are sent in rapid succession[1]. We also find that the three injectors also behave slightly differently in how they format their DNS responses. Specifically, Injector 1 uses the domain from the query as-is in the DNS response, whereas Injectors 2 and 3 use a "compression pointer" [19] to reduce repetition of the query domain in the response, perhaps a sign of these injectors using a different code base in their operation.

Similar to prior work [1], we observe Injector 1 with an incrementing IP TTL value between subsequent packets. However, we see this injector is considerably less active in terms of the number of domains it filters. Figure 5 shows the number of domains that observed an injected reply from each injector. We can see that Injector 1, which most closely resembles the injector seen in 2014 [1], only filtering a total of 88 domains.

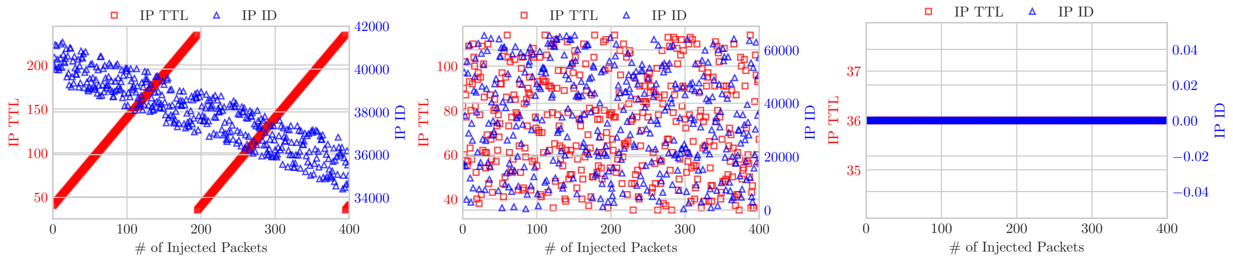
Interestingly, we do not observe any domains that *only* trigger Injector 3, with it acting on a subset of Injector 2's domains. When we consider the relationship between the Injectors and the IP/domain groups (Table 3), we see that Injector 1 is the only injector filtering IPs in the fourth IP/Domain group with 33 domains that are mostly in the "Search Engines" category (cf. Table 2).

While Figure 5 gives a sense of the number of domains filtered by each injector, it doesn't necessarily reflect how often the injector would be triggered. For this, we consider the popularity of domains that each injector acts on. Figure 6 shows the cumulative percentage of domains filtered by each injector relative to their Alexa ranking. Here we see that domains filtered by Injector 1 tend to be more popular than those filtered by the other injectors. Most of the domains (97%) censored by Injector 1 are domains that contain the keyword `google`, and 90% of them are in the top 350K domains in the Alexa top 1M list. While, the majority (80%) of domains censored by Injectors 2 and 3 are `*.blogspot` and `.*tumblr` domains which are in the long tail of the Alexa 1M list [25].

4.2 Longitudinal trends

Halting interval of injectors. Figure 7 shows the total number of injected packets on a daily basis. Due to the frequency of our measurements, we are not able to discover any gaps less than two hours. When analyzing the data on a bi-hourly basis, we discover that while Injector 2 has been working consecutively, Injector 1 and Injector 3 occasionally stopped working for a few hours. Specifically, the three halting intervals of Injector 1 are between 13:00 and 15:22 on September 18, 2019; between 9:26 and 13:00 September 19, 2019; and between 17:06 to 10:22 on September 19, 2019. The only halting intervals of Injector 3 are between 2:36 and 8:00 on

[1] In this test, we injected packets as fast as we could using a multi-threaded Python program while using tcpdump to capture the response packets.



(a) Injector 1 (b) Injector 2 (c) Injector 3

Figure 4: IPID and TTL values observed for the three DNS Injector behaviors observed in our measurements. Injector 1 is similar to what has previously been observed in [1]. We observe that the third injector reflects the IP TTL value, leading to a fixed value when the initial IP TTL values of our queries are not varied.

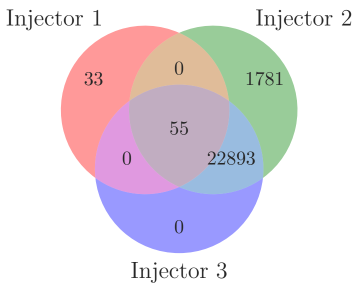


Figure 5: Venn diagram showing the number of domains receiving different combinations of injected responses by the three observed DNS injectors.

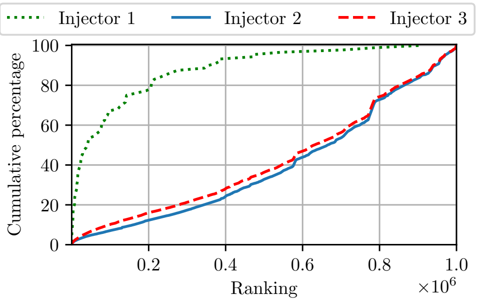


Figure 6: CDF of the popularity ranking of censored domains by each injector.

May 1 (in Beijing Time). We note the actual halts are likely to be a sub-interval of what we have discovered. All of these occasionally happened halts lasted less than 6 hours and most of them happened during work hours in China.

Relationship between injectors and the IP drop seen in Figure 2. We analyzed the IPs used by the injectors over time, specifically before and after the decrease in the number of distinct IPs injected on November, 2019. The decrease has no effect on Injector 1 as it always uses the same four distinct IPs. However, Injector 2 and Injector 3 initially use a pool of 958 and 1,506 IPs to send injected DNS replies, respectively. After the drop, both Injector 2 and 3 use the same IP pool (with 212 IPs) for their injected DNS replies.

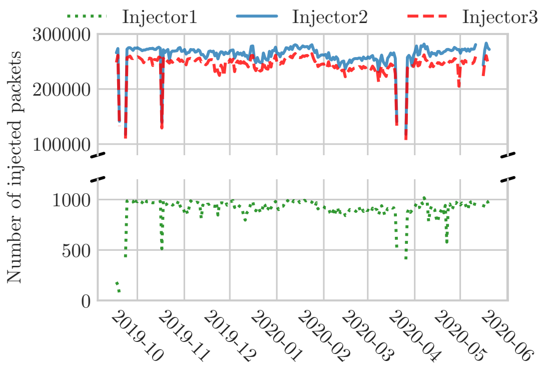
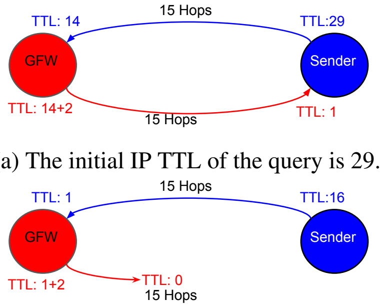


Figure 7: Total number of injected packets per injector received each day across time. The gaps are all due to disruptions of the measurements.

4.3 Localizing the Injectors

We next attempt to localize the three injectors identified in §4.1. We use the commonly employed method of sending packets with incrementing IP TTL values until we receive an injected DNS reply to identify where on our path the packet injector lies [1, 18, 23, 26, 29]. For this test, we focus on a single domain that we observed to trigger all three injectors: `www.google.sm`. We then send DNS queries for this domain from our VPS in the US to the VPS in China.

Based on these TTL limited probes, we were able to observe that Injectors 1 and 2 are located 15 hops away from our US VPS. For comparison, our Chinese VPS is 25 hops from our US VPS. However, we observed an unusual behavior with Injector 3, where we did not see an injected DNS reply from Injector 3 until the initial TTL on our probe packet is set to 29. Given that the destination IP of our probe packet was only 25 hops away, this behavior seemed unusual. However, upon closer inspection, we determined that this behavior stemmed from Injector 3 echoing the incremented TTL of the probe packet in its injected reply.



(a) The initial IP TTL of the query is 29.

(b) The initial IP TTL of the query is 16.

Figure 8: Illustration of how Injector 3 mirroring the IP TTL of the DNS query impacts the results of TTL-limited probing. Figure 8a shows that when the IP TTL of the DNS query is 29 the corresponding injected packet has a high enough TTL to reach the sender. Figure 8b shows that when the IP TTL of the DNS query is below 29, the initial IP TTL of the forged response is too small to reach the sender.

Figure 8 illustrates this phenomenon. We find that when the probe packet has a TTL of 29, the injected reply has an IP TTL of 1 when it reaches our US host. Similarly, when the probe packet has a TTL of 30 the TTL of the injected reply is 2, and so on. The precise probe TTL needed to observe this behavior is $2n-1$ where n is the number of hops between the probing host and the packet injector. We note, that this discussion implicitly assumes symmetric paths between the injector and the probing host. This behavior could potentially be used to identify asymmetric routing on paths (when a domain that will trigger multiple injectors is used), but we leave more in depth analysis of this to future work.

We also compare the time between sending our DNS query and when we receive the injected reply to get a sense of where the injectors are located. Specifically, we compare the delays of the three injectors and find that more than 90% of the time the delays are within 0.2 ms of each other. This would support the theory that these three devices are installed in the same physical location.

We repeat these experiments from seven hosts outside of China (our VPS in the US and cloud-hosted VMs in the Netherlands, Singapore, UK, France, Canada and India) with consistent results.

5 Multi-path Results

In Section 2, we describe our method to send our DNS queries to 36K Chinese prefixes. Our goal here is to confirm that our results are robust, to the location of the host that we focus on for our longitudinal experiment. Figure 9 shows the result: each bar corresponds to the combination of injectors that were observed and the height of the bar corresponds to the percent of prefixes where this combination of injectors was observed.

Of the 36K prefixes we direct our query towards, we find that 62% of them observe all three DNS injectors. We observe 12% of cases where two of the three injectors are observed, and 13% of cases where only one of the three injectors are observed. For each IP address, we send 100 queries which suggests that these cases are not just caused by transient packet loss. We also observe some injectors that are not seen in our longitudinal data in this broader study (denoted by Injector X in Fig. 9). In total, there are around 4% of the prefixes where we observe injectors not matching our fingerprints.

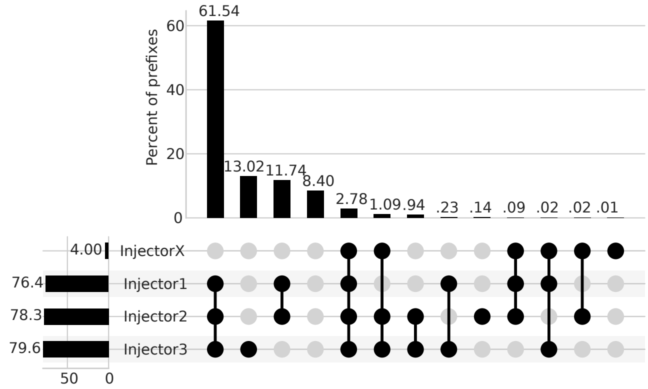


Figure 9: Number of unique IP prefixes responding with different types of responses. InjectorX refers to the injectors that have fingerprints other than the summarized ones.

Interestingly, we see 8% of the prefixes, registered to 134 ASes, where no DNS injector is triggered. Using the RIPE NCC AS visibility tool [24], we find 22% of these prefixes have less than 15% visibility, suggesting our queries may never reach these prefixes. For the remaining prefixes, we use the RIR-based IP-to-ASN mapping provided by Team Cymru [27] and find that half of these prefixes are registered outside of China (e.g., a Chinese-based company registering an IP address with ARIN). In these cases, the prefixes may be located outside of China and not subject to censorship. It worth noting that there are still 1,027 IP prefixes that seem to be within China's territory, but with no injected packet observed. These IP prefixes correspond to 120 ASes. Upon closer inspection we find that these ASes tend to be related to technology companies or government agencies.

6 Conclusion

In this work, we analyze the DNS poisoning behavior of the GFW across nine months. We observe groups of IPs used to censor specific groups of domains and identify three distinct DNS packet injectors. We localize and characterize the behavior of these injectors and identify one injector mirroring the TTL of the probe packets which has implications for studies that use TTL-limited packets to localize DNS censors.

We have released our code and dataset to maintain reproducibility and to stimulate future work, obtainable at https://gfw.report/publications/foci20_dns/en/.

Acknowledgments

We would like to thank our shepherd Anita Nikolich for her thorough feedback and guidance. We also thank Nicholas Weaver for helpful discussions about the injectors' behavior. This work funded by a Google Faculty Research Award, NSF CAREER grant CNS1553301, NSF grants CNS1740895 and CNS1719386, and by the Open Technology Fund under an Information Controls Fellowship.

References

[1] Anonymous. Towards a comprehensive picture of the Great Firewall's DNS censorship. In *Free and Open Communications on the Internet*. USENIX, 2014. https://www.usenix.org/system/files/conference/foci14/foci14-anonymous.pdf.

[2] Anonymous. GFW Archaeology: gfw-looking-glass.sh, March 2020. https://gfw.report/blog/gfw_looking_glass/en/.

[3] Simurgh Aryan, Homa Aryan, and J. Alex Halderman. Internet censorship in Iran: A first look. In *Free and Open Communications on the Internet*. USENIX, 2013. https://censorbib.nymity.ch/pdf/Aryan2013a.pdf.

[4] S. Bortzmeyer and S. Huque. NXDOMAIN: There Really Is Nothing Underneath. RFC 8020, IETF, November 2016. https://tools.ietf.org/html/rfc8020.

[5] Censored Planet: Satellite and Iris. Available at https://censoredplanet.org/projects/satellite.

[6] Center for Applied Internet Data Analysis. Inferred AS to Organization Mapping Dataset. Web page, Accessed 2020. https://www.caida.org/data/as-organizations/.

[7] Center for Applied Internet Data Analysis. Routeviews Prefix to AS mappings Dataset for IPv4 and IPv6 . Web page, Accessed 2020. http://www.caida.org/data/routing/routeviews-prefix2as.xml.

[8] Abdelberi Chaabane, Terence Chen, Mathieu Cunche, Emiliano De Cristofaro, Arik Friedman, and Mohamed Ali Kaafar. Censorship in the wild: Analyzing Internet filtering in Syria. In *Internet Measurement Conference*. ACM, 2014. http://conferences2.sigcomm.org/imc/2014/papers/p285.pdf.

[9] Zimo Chai, Amirhossein Ghafari, and Amir Houmansadr. On the importance of encrypted-SNI (ESNI) to censorship circumvention. In *Free and Open Communications on the Internet*. USENIX, 2019. https://www.usenix.org/system/files/foci19-paper_chai_update.pdf.

[10] Haixin Duan, Nicholas Weaver, Zongxu Zhao, Meng Hu, Jinjin Liang, Jian Jiang, Kang Li, and Vern Paxson. Hold-On: Protecting against on-path DNS poisoning. In *Securing and Trusting Internet Names*. National Physical Laboratory, 2012. http://conferences.npl.co.uk/satin/papers/satin2012-Duan.pdf.

[11] Oliver Farnan, Alexander Darer, and Joss Wright. Poisoning the well – exploring the Great Firewall's poisoned DNS responses. In *Workshop on Privacy in the Electronic Society*. ACM, 2016. https://dl.acm.org/authorize?N25517.

[12] Arturo Filastò and Jacob Appelbaum. OONI: Open observatory of network interference. In *Free and Open Communications on the Internet*. USENIX, 2012. https://www.usenix.org/system/files/conference/foci12/foci12-final12.pdf.

[13] FortiGuard Labs Web Filter, Accessed 2018. https://fortiguard.com/webfilter.

[14] gfwrev. 深入理解GFW：DNS污染, November 2009. https://gfwrev.blogspot.com/2009/11/gfwdns.html.

[15] Phillipa Gill, Masashi Crete-Nishihata, Jakub Dalek, Sharon Goldberg, Adam Senft, and Greg Wiseman. Characterizing web censorship worldwide: Another look at the OpenNet Initiative data. *Transactions on the Web*, 9(1), 2015. https://censorbib.nymity.ch/pdf/Gill2015a.pdf.

[16] Nguyen Phong Hoang, Sadie Doreen, and Michalis Polychronakis. Measuring I2P censorship at a global scale. In *Free and Open Communications on the Internet*. USENIX, 2019. https://www.usenix.org/system/files/foci19-paper_hoang.pdf.

[17] Nguyen Phong Hoang, Arian Akhavan Niaki, Nikita Borisov, Phillipa Gill, and Michalis Polychronakis. Assessing the Privacy Benefits of Domain Name Encryption. In *ACM ASIACCS 2020*. https://arxiv.org/pdf/1911.00563.pdf.

[18] Graham Lowe, Patrick Winters, and Michael L. Marcus. The great DNS wall of China. Technical report, New York University, 2007. https://censorbib.nymity.ch/pdf/Lowe2007a.pdf.

[19] P. Mockapetris. DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION. RFC 1035, IETF, November 1987. https://tools.ietf.org/html/rfc1035.

[20] Zubair Nabi. The anatomy of web censorship in Pakistan. In *Free and Open Communications on the Internet*. USENIX, 2013. https://censorbib.nymity.ch/pdf/Nabi2013a.pdf.

[21] Arian Akhavan Niaki, Shinyoung Cho, Zachary Weinberg, Nguyen Phong Hoang, Abbas Razaghpanah, Nicolas Christin, and Phillipa Gill. ICLab: A global, longitudinal internet censorship measurement platform. In *Symposium on Security & Privacy*. IEEE, 2020. https://people.cs.umass.edu/~phillipa/papers/oakland2020.pdf.

[22] Paul Pearce, Ben Jones, Frank Li, Roya Ensafi, Nick Feamster, Nick Weaver, and Vern Paxson. Global measurement of DNS manipulation. In *USENIX Security Symposium*. USENIX, 2017. https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-pearce.pdf.

[23] Thomas H Ptacek and Timothy N Newsham. Insertion, evasion, and denial of service: Eluding network intrusion detection. Technical report, Secure Networks inc Calgary Alberta, 1998. https://users.ece.cmu.edu/~adrian/731-sp04/readings/Ptacek-Newsham-ids98.pdf.

[24] RIPE NCC AS Visibility Tool, Accessed 2020. https://stat.ripe.net.

[25] Quirin Scheitle, Oliver Hohlfeld, Julien Gamba, Jonas Jelten, Torsten Zimmermann, Stephen D Strowes, and Narseo Vallina-Rodriguez. A long way to the top: Significance, structure, and stability of internet top lists. In *Proceedings of the Internet Measurement Conference 2018*, pages 478–493, 2018. https://dl.acm.org/doi/pdf/10.1145/3278532.3278574.

[26] Sparks, Neo, Tank, Smith, and Dozer. The collateral damage of Internet censorship by DNS injection. *SIGCOMM Computer Communication Review*, 42(3):21–27, 2012. http://conferences.sigcomm.org/sigcomm/2012/paper/ccr-paper266.pdf.

[27] Team Cymru IP to ASN Mapping Service, Accessed 2020. https://team-cymru.com/community-services/ip-asn-mapping/.

[28] John-Paul Verkamp and Minaxi Gupta. Inferring mechanics of web censorship around the world. In *Free and Open Communications on the Internet*. USENIX, 2012. https://www.usenix.org/system/files/conference/foci12/foci12-final1.pdf.

[29] Xueyang Xu, Z. Morley Mao, and J. Alex Halderman. Internet censorship in China: Where does the filtering occur? In *Passive and Active Measurement Conference*, pages 133–142. Springer, 2011. https://web.eecs.umich.edu/~zmao/Papers/china-censorship-pam11.pdf.

[30] Boru Yan, Binxing Fang, Bin Li, and Yao Wang. DNS欺骗攻击的检测和防范. 计算机工程, 32(21):130–132, 2006. https://web.archive.org/web/20200726140258/https://tomcat.one/files/papers/DNS%E6%AC%BA%E9%AA%97%E6%94%BB%E5%87%BB%E7%9A%84%E6%A3%80%E6%B5%8B%E5%92%8C%E9%98%B2%E8%8C%83_%E9%97%AB%E4%BC%AF%E5%84%92.pdf.