A Wall Behind A Wall: Emerging Regional Censorship in China

Mingshi Wu*
GFW Report
gfw.report@protonmail.com

Ali Zohaib*
University of Massachusetts Amherst
azohaib@umass.edu

Zakir Durumeric
Stanford University
zakir@cs.stanford.edu

Amir Houmansadr
University of Massachusetts Amherst
amir@cs.umass.edu

Eric Wustrow
University of Colorado Boulder
ewust@colorado.edu

***Abstract*—China has long orchestrated its Internet censorship through relatively centralized policies and a unified implementation, known as the Great Firewall of China (GFW). However, since August 2023, anecdotes suggest that the Henan Province has deployed its own regional censorship. In this work, we characterize provincial-level censorship in Henan, and compare it with the national-level GFW. We find that Henan has established TLS SNI-based and HTTP Host-based censorship that inspects and blocks traffic leaving the province. While the Henan Firewall is less sophisticated and less robust against typical network variability, its volatile and aggressive blocking of second-level domains made it block ten times more websites than the GFW at some points in time. Based on the observed parsing flaws and injection behaviors, we introduce simple client-side methods to bypass censorship in the Henan province. Our work documents an alarming sign of regional censorship emerging in China.**

1. Introduction

The People's Republic of China develops and maintains one of the most sophisticated Internet censorship apparatuses, colloquially referred to as the Great Firewall (GFW). Through DNS poisoning [1], [2], [3], [4], [5], HTTP Host header filtering [6], [7], [8], [9], TLS SNI/ESNI filtering [2], [9], [10] [11 §3], IP address blocking [2 §4], active probing [12], [13], [14] [15 §5], and proxy traffic detection [15 §4], China blocks its citizens from accessing large swaths of Internet content and services.

China's censorship apparatus has long been believed to be operated relatively centrally, in terms of both its *policy* and *implementation*. Empirical measurements have revealed China's uniform and coordinated management of censorship policies [3], [4], [9], [15], software updates [16 §4.5] [5 §VII], and infrastructures [14 §3.4] [4 §5]. Censorship devices are positioned at the national border [4], [17], [18], where they inspect and filter traffic entering or exiting the country. As a result, traffic exchanged domestically within China is not inspected or blocked by the GFW.

However, recent anecdotes suggest that this centralized and uniform censorship model may no longer tell the whole story. In August 2023, users in the Henan Province of China—the third-largest province by population and a pivotal labor hub—began reporting an uptick in inaccessible websites that were accessible elsewhere in China [19].

In this work,[1] we first explore a natural question raised by the discovery of regional censorship in Henan (Section 3): have other provinces in China deployed the same or similar regional censorship? We conducted a measurement study in seven provinces and municipalities in China, including Beijing, Shanghai, Guangdong, Zhejiang, Jiangsu, Sichuan, and Henan, to identify potential regional censorship. Likely limited by the vantage points we could access in China, we found no evidence of regional censorship in the six provinces other than Henan.

We then analyze the emerging regional censorship in the Henan Province, comparing its policies and implementations with the national GFW. As illustrated in Figure 1, our investigation reveals that the provincial-level middleboxes in Henan block access to certain HTTP and HTTPS websites through both HTTP Host-based and TLS Server Name Indication (SNI)-based filtering (Section 4.1). Contrasting the GFW that monitors and blocks traffic leaving and entering the country, this regional firewall only censors traffic exiting the province (Section 4.2). It also differs from the GFW in terms of connection tracking and parsing logic (Section 4.3), injection behaviors and fingerprints (Section 4.4), and network location. (Section 4.5).

We conduct a longitudinal study to understand the content blocked by the Henan Firewall and how it differs from the content blocked by the GFW (Section 5). Between November 2023 and March 2025 (with a measurement gap between March and October 2024), we tested Tranco top one million domains on a daily basis, and tested CZDS 227 million domains on a weekly basis. We find that the Henan Firewall employs more aggressive and volatile blocking policies than the GFW. The Henan Firewall blocked a cumulative 4.2 million domains, more than five times

*. Mingshi Wu and Ali Zohaib contributed equally to this work.

1. Project homepage: https://gfw.report/publications/sp25/en.

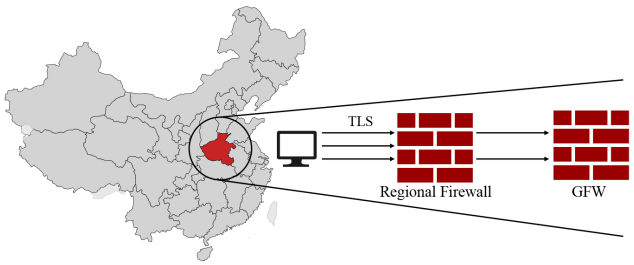


Figure 1: Henan Province has deployed TLS SNI-based and HTTP Host-based censorship middleboxes that inspect and block traffic exiting the province.

the size of the GFW's cumulative blocklist. A key reason for this was its blocking of generic second-level domains (e.g., `*.com.au`). Our testing also revealed periods where it blocked ten times more domains than the GFW.

Based on the observed parsing flaws and injection behaviors, we introduce circumvention techniques to bypass this regional censorship (Section 6), which have been implemented by various popular anti-censorship tools. The regional censorship in Henan marks one of the first formally documented cases of a provincial firewall operating autonomously in China. We hope our study sounds the alarm to the broader censorship research community to identify, investigate, and combat the emergence of regional censorship in China and elsewhere.

2. Background and Related Work

2.1. The Great Firewall of China (GFW)

The Great Firewall of China (GFW) is a set of different censorship mechanisms and devices deployed in China. The GFW utilizes a network of middleboxes distributed across China's border autonomous systems (ASes) to inspect and block Internet traffic [17]. The GFW not only blocks access to specific websites and services, but also tries to identify and block attempts to bypass its censorship.

Website censorship. To block access to specific websites and services, the GFW often employs a combination of techniques, including DNS injection [1], [3], HTTP Host-based filtering [8], TLS SNI/ESNI-based filtering [2], [9], [10], [20], and IP address blocking [2 §4].

To censor DNS traffic, the GFW operates *on-path* to inject forged DNS responses with wrong IP addresses to block access to specific domains [3], [4], [5], [21], [22]. Early reports from 2002 documented that the GFW used a single wrong IP address in its forged responses [23], [24]. Over time, this evolved into a more sophisticated system employing an increasing number of fake addresses and expanding the list of blocked domains [3], [4], [22], [25]. Researchers have uncovered memory disclosure vulnerabilities in the GFW's injection system [5], [16], [26].

To censor HTTP and TLS traffic, the GFW statefully inspects unencrypted text in the connection. Upon detecting a censored domain in a HTTP request's Host field or in a TLS ClientHello's Server Name Indication (SNI) extension, the GFW injects TCP `RST` packets to both sides of the connection to tear it down [6], [7], [9], [27], [28]. Figure 2 shows the GFW's operation on a connection containing a forbidden domain name in the SNI of the TLS Client Hello.

The GFW often operates bidirectionally, meaning both traffic coming into the country and leaving the country can trigger its censorship [4], [9], [29]. The bidirectional operation of censorship middleboxes has enabled researchers to measure censorship from outside the country [3], [30], [31].

Projects such as OONI [32], Censored Planet [33], and ICLab [34] have been measuring censorship globally for years. To monitor website censorship in China, several large-scale projects have been developed, including the GreatFire Analyzer [35], Blocky [36], GFWatch [3], and GFWeb [9]. While longitudinal and large scale studies are excellent at tracking and understanding the blocklist changes in the GFW, sometimes a revisit of the existing censorship mechanisms could still reveal new updates by the censor. For example, Bock et al. [11] discovered secondary TLS censorship middleboxes in China that had operated undetected until an in-depth analysis revealed them.

Proxy censorship. Blocking access to websites is not enough to prevent users from accessing censored content, as users can use circumvention tools to bypass censorship. There has thus been a seemingly endless cat-and-mouse game between the GFW and the Internet users in China [37]. For example, the GFW employs active probing techniques to identify and block circumvention tools, such at Tor [12], [13], [38], [39], [40] and Shadowsocks [14] [15 §5], which have been successfully defended against [41], [42], [43], [44]. The GFW also conducts traffic analysis to identify and block fully encrypted proxies [15].

Other censorship mechanisms. There have also been unique components of China's censorship that appear separate from the GFW's censorship against websites and proxies. Notably, in 2015, researchers discovered the "Great Cannon" of China, which injected Javascript into HTTP traffic in order to co-opt victim browsers into participating in a denial-of-service attack against specific hosts [30].

2.2. Regional Variation in Censorship

Localized or decentralized censorship mechanisms are common in countries with strict censorship policies. In Russia, thousands of privately owned ISPs each implement their own filtering mechanisms, resulting in a varied censorship landscape [45], [46], [47]. Similarly, in India, researchers have shown that ISPs differ significantly in their implementation of government censorship orders, leading to fragmented censorship across the country [48].

However, prior work has suggested that China's censorship systems and policies are largely uniform and centralized across the country. In 2011, Xu et al. [17] measured

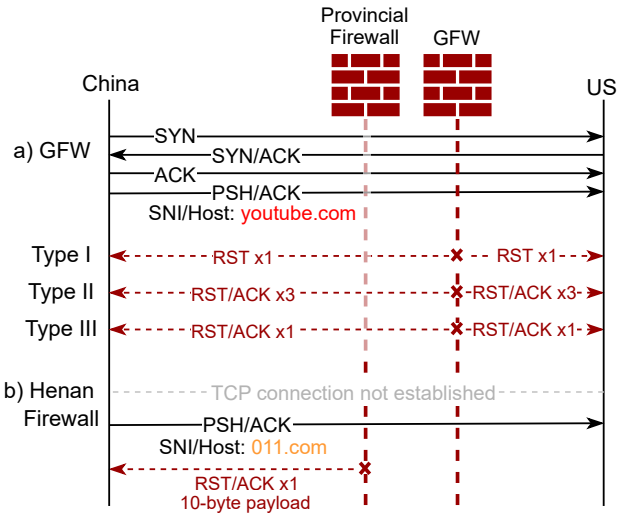


Figure 2: Overview of the Henan Firewall and the three different types of GFW. One can trigger and study each censorship mechanism individually by putting exclusively censored domains in the SNI or HTTP Host field of a probe. For example, as of April 2024, `011.com` was exclusively blocked by the Henan Firewall, and `youtube.com` was exclusively blocked by the GFW.

the location of censorship devices in China. They found that China's keyword-censoring middleboxes were largely at the edges of the network and employed rules in line with nationwide blocking policies of that time. In 2012, Wright [18] performed a small-scale study on DNS censorship in China, finding that DNS responses to queries varied across the country. However, this work did not account for other possible causes of the variation in DNS responses (e.g. geolocation-based load balancing, or changes in DNS configuration). In 2018, Bao et al. [49] measured DNS injection variances in China from residential and cellular IP addresses. Internet-wide and longitudinal measurements have revealed China's uniform and coordinated management of censorship policies [3], [4], [9], [15], software updates [16 §4.5] [5 §VII], and infrastructures [14 §3.4] [4 §5].

3. Detecting Regional Censorship

Anti-censorship researchers outside of China often rely on local user reports to learn about the new censorship shifts and upgrades in China. This is partially because of the difficulty for researchers to obtain a diverse range of vantage points inside China and then constantly monitor various Internet services and protocols. Encouragingly, online discussion forums—such as Net4People BBS [50], NTC Party forum [51], and the GitHub issue pages of popular anti-censorship tools such as Xray [52], V2Ray [53], singbox [54], and Hysteria [55]—enable users to report new censorship behaviors as soon as they encounter them and allow researchers to investigate those reports promptly [37].

This crowdsourced, collaborative approach has also been effective in identifying and combating the provincial censorship in Henan. In particular, our study started with reports from a group of users in Henan who were unable to access certain websites [19], [56], [57], [58], [59]. We then obtained a server in the Henan Province and confirmed the presence of a regional firewall. In particular, as illustrated in Figure 2, we found that the regional Henan Firewall blocked TLS and HTTP connections for some Server Name Indication (SNI) and HTTP Host values, but it operated differently than the GFW. Most distinctively, the regional firewall in Henan blocks a TCP connection by injecting one TCP `RST+ACK` packet containing a fixed 10-byte payload to the client. The unique payload of TCP `RST` packet differentiates the Henan Firewall from all three types of packets injected by the GFW.

The discovery of regional censorship in Henan province led to a natural question: have other provinces in China deployed the same or similar regional censorship? Below, we explore this question with measurement across the country.

3.1. Experiment

Our goal is to quantify the regional variation of TLS censorship across China by comparing the number of domains blocked between each pair of hosts inside or outside of China. As summarized by the second row of Table 1, we obtained two vantage points in each of the seven cities in China, including Shanghai, Beijing, Chongqing, Guangzhou (Guangdong Province), Nanjing (Jiangsu Province), Chengdu (Sichuan Province), and Zhengzhou (in Henan Province). We also setup two VPSes in each of the three locations outside of China: Seattle (U.S.), San Francisco (U.S.), and Singapore. Our selection of vantage points was guided by a set of ethical considerations detailed in Section 7.

For the two VPSes in each location inside or outside of China, we used one as a client and the other as a *sink server*. The sink servers were configured to accept TCP handshakes on all ports between 1 and 65535. They would acknowledge the TCP data sent to them, but would never send any TCP payload back to the client. We configured `iptables` rules on both the client and sink servers to drop any outgoing `RST` packets. This way, any `RST` packets received on either end must be injected by some middleboxes on the network path. We could thus confirm the presence of censorship by checking whether the TCP connection was being reset.

We then sent TLS traffic with various SNI values between each pair of the clients and sink servers on July 10, 2024. In particular, we used the top 10,000 domains from the Tranco list [60] 5YZ7N for testing.[2] To reduce the chances of false negatives due to packet loss, we repeated our test three times on the same day and let the OS to control retransmissions of the packets.

Limitations. Ideally, we would liked to use a diverse set of vantage points to identify potential regional censorship

2. Tranco list ID 5YZ7N, obtained on August 15, 2023: https://tranco-list.eu/list/5YZ7N/1000000.

TABLE 1: Experiment timeline and vantage points. In total, we used 14 VPSes in China VPS Cloud (CVC, AS4837) in Zhengzhou, Henan Province (HN), six VPSes in Akamai Linode (LD, AS63949) in San Francisco (SF), Singapore (SG) and Seattle (SE), 12 VPSes in Tencent Cloud (TC, AS45090) in Beijing (BJ), Shanghai (SH), Chongqing (CQ), Guangzhou, Guangdong Province (GZ), Chengdu, Sichuan Province (CD), Nanjing, Jiangsu Province (NJ), and one bare metal network tap server (TAP) in a U.S. university.

Experiments	Time Span	Duration	China Vantage Points	External Vantage Points	Sections
Identification	7/10/24	1 day	12 (TC), 2 (CVC: HN)	4 (LD: SG,SE)	§3
Characterization	10/2/23 – 11/12/24	13 months	2 (CVC: HN)	1 (LD: SF), 3 (TC: GZ,BJ,SH)	§4
Traffic Analysis	10/31/24	1 hour	-	1 (TAP: US)	§4.3
Locating	10/2/23 – 12/8/23	2 months	1 (CVC: HN), 1 (TC: GZ)	1 (LD: SF)	§4.5
Blocklist	11/5/23 – 3/5/24 & 10/07/24 – 3/31/25	9 months	14 (CVC: HN), 2 (TC: GZ)	2 (LD: SF)	§5

in China. However, due to the difficulty of obtaining VPSes in China, we have only been able to obtain vantage points in a limited number of locations and ASes. While using residential vantage points would have allowed us to observe potential middleboxes from more network locations in China, this could put potential risks on uninformed users and providers of residential proxies [61]. For this reason, we focus on using two large VPS providers in China, China VPS Cloud and Tencent Cloud, to avoid risks or persecution to individuals. We utilized all available locations these two VPS providers offered to maximize our coverage. We acknowledge that our results are limited to measuring TLS censorship, which could potentially miss regional censorship of other protocols. Additionally, due to a configuration error, we did not test using our client in Singapore, potentially missing bidirectional censorship from that perspective.

3.2. Results

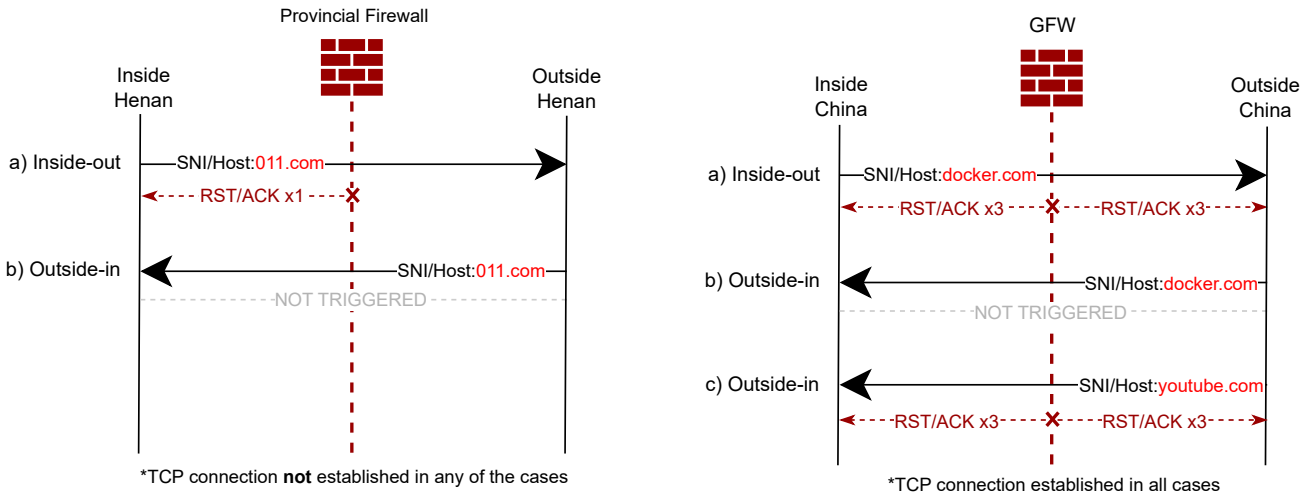
Figure 3 shows the number of blocked domains between different locations. We first observe that, connections originating from China to our sink servers in Singapore and the U.S. were almost equally impacted by the national-level Great Firewall of China (GFW), with around 479 out of the 10,000 domains blocked. The most significant blocking was observed in Zhengzhou, the capital of Henan province, where both provincial (Henan) and national (GFW) censorship mechanisms contributed to the high figure.

Traffic leaving Henan is affected by the regional firewall, regardless of the sink server location, even to other regions within China. On average, 122 domains were blocked by the Henan Firewall. We did not observe any blocking of TLS connections within Henan itself; however, since both of our client and sink servers were in the same data center, we can only cautiously conclude that the Henan Firewall does not affect internal traffic within this data center.

When connections were made from Zhengzhou, Henan Province, to locations outside China (Singapore and Seattle), a total of 594 domains were blocked. This indicates the simultaneous operation of two firewalls with independent blocklists, with the Henan Firewall intercepting traffic before it reaches the GFW and thus, increasing the total number of domains that are blocked. We, however, did not observe any blocking of connections from other client locations in China to Henan or other sink server regions within China. This finding suggests that the *Henan Firewall is the first known deployment of a regional firewall in China.*

Clients \ Servers	Beijing	Shanghai	Guangdong	Sichuan	Chongqing	Jiangsu	Henan	Singapore	U.S.
Beijing	0	0	0	0	0	0	0	478	479
Sichuan	0	0	0	0	0	0	0	479	479
Chongqing	0	0	0	0	0	0	0	479	479
Guangdong	0	0	0	0	0	0	0	479	479
Jiangsu	0	0	0	0	0	0	0	479	479
Shanghai	0	0	0	0	0	0	0	479	479
Henan	123	122	122	122	122	124	0	594	594
U.S.	411	411	411	411	411	440	411	0	0

Figure 3: The matrix shows the number of domains blocked between each pair of hosts in various locations. For each host pair, we sent TLS ClientHello messages with SNI values of the top 10,000 domains from the Tranco list [60] 5YZ7N, generated on August 15, 2023. The result suggests that 1) regional censorship in Henan province exists evidenced by the non-zero number of blocked domains when testing from Zhengzhou, Henan to sink servers in other regions of China; 2) the censorship in Henan is not bidirectional, as initiating TLS connections from the outside to Henan did not trigger any blocking; 3) the GFW maintains a blocklist that is only censored when accessed from within China, as evidenced by the differences in the numbers of blocked domains when testing inside-out and outside-in.



(a) Henan Firewall (b) GFW

Figure 4: (a) The Henan Firewall does not censor inbound TLS or HTTP traffic to Henan, contracting the bidirectional censorship employed by the GFW. (b) The GFW's TLS and HTTP censorship machines inspect bidirectional traffic coming in and out of China; however, certain domains are only censored when accessed from within China. In this example, while a TLS ClientHello with SNI value `docker.com` can trigger the three TCP `RST` packets by the GFW when sent from within China, it does not trigger any blocking when sent from outside of China.

Moreover, as presented in the last row of Figure 3, tests from the U.S. to various locations in China consistently identified the same 411 domains blocked by the GFW, with only one exception: tests from the U.S. to Jiangsu Province detected 440 blocked domains. Further analysis indicates that the additional 29 domains blocked in the outside-in direction for Jiangsu is a subset of the 479 domains blocked by the GFW in the inside-out direction. This finding suggests that the additional censorship of these 29 domains likely does not reflect regional censorship specific to Jiangsu. Instead, it indicates that the GFW is configured to block these domains bidirectionally within Jiangsu.

Overall, these results are particularly noteworthy as they show the infeasibility of remote measurements to trigger the regional firewalls and more importantly, the asymmetric behavior of the GFW. In particular, while 479 domains were blocked on average when connections were initiated from within China, only 411 domains were blocked when connections were initiated from outside China. This discrepancy suggests that the GFW enforces a different blocklist for traffic originating from within China. Until recently, it was widely believed that the GFW operated symmetrically, triggering and applying the same blocklist to traffic regardless of direction. However, recent work has suggested this assumption is incorrect [9], and our findings here are consistent with this recent result.

We note that both the GFW and the Henan Firewall exhibit asymmetric interference to varying degrees. As shown in Figure 4(a) , while traffic going out of Henan is subject to the regional firewall (inside-out), inbound traffic (outside-in) to Henan does not trigger the regional firewall at all. This stands in contrast to the GFW, which, although bidirectional, behaves asymmetrically based on the domains queried.

Figure 4(b) provides a clear example of this behavior. When a TLS ClientHello with SNI value `docker.com`, in our case, is sent from within China (inside-out), the GFW triggers blocking via three TCP `RST` packets. However, when the same TLS ClientHello is sent from outside of China (outside-in), the GFW does not trigger any blocking. On the other hand, when a TLS ClientHello packet with the SNI value `youtube.com`, in this example, is sent, the GFW triggers blocking in both scenarios: whether the packet is sent from inside or outside China. This behavior demonstrates an apparent blocklist of domains that are exclusively censored by the GFW when accessed from within China.

In our experiment designed to detect any regional censorship, we inadvertently uncovered a significant aspect of the GFW's operational mechanics that has only recently been documented [9]. The newly observed asymmetric nature of the GFW and regional firewalls highlights the critical need for inside-out measurements to fully capture the extent and nuances of censorship. Relying solely on remote measurements, as is common in many other studies, fails to provide a comprehensive picture of such censorship events.

To further substantiate the asymmetric behavior of the GFW, we provide a list of domains that are exclusively blocked when sending TLS ClientHello messages from within China, as shown in Table 2. In our experiment, 68 out of the 10,000 domains did not trigger any censorship when tested from outside China but only were blocked when probed from inside China. These domains include popular websites such as `google.com`, `nyt.com`, and `docker.com`.

TABLE 2: A sample of domains that are exclusively blocked by the GFW when sending TLS ClientHello messages from within China. These domains did not trigger censorship when sent from outside China to within, as of July 10, 2024. Among the 10,000 Tranco top domains we tested, 68 domains were exclusively blocked inside-out and no domains were exclusively blocked outside-in.

binance.com	godaddy.com	note.com
cdninstagram.com	google.com	nyt.com
docker.com	google.com.hk	tiktokcdn.com
gmail.com	linktr.ee	torproject.org

The list serves as concrete evidence of the selective enforcement of the GFW's blocklist based on the origin of the traffic and the domain in question.

4. Characterizing the Censorship Devices

Since October 2023, we conducted a series of experiments to characterize censorship devices and understand the differences between the Great Firewall (GFW) and Henan regional censorship devices. In this section, we answer several research questions: where are the regional censorship devices located? What packets can trigger the Henan SNI Firewall? Which ports are monitored by the Henan Firewall? Does the TCP `RST` injections have any specific fingerprints? And does the Henan Firewall induce residual censorship?

4.1. Methodology

We developed a methodology tailored to the specific characteristics of two firewalls, i.e., regional and national, as discussed earlier. To precisely assess the impact of each firewall, our approach involves isolating and analyzing these two systems individually. This method, which was devised based on our preliminary observations, serves as the foundation for our comprehensive measurement experiments. Key aspects of our methodology are outlined below.

Obtaining Vantage Points. In total we use 10 vantage points in Zhengzhou, China (in Henan) acquired via China VPS Cloud (AS 4837), two VPSes in Guangzhou, one in Beijing, and one in Shanghai via Tencent Cloud (AS 45090), and two VPSes in San Francisco, U.S. through Akamai's Linode (AS 63949). The VPSes in Guangzhou, Beijing, Shanghai and San Francisco served as sink servers that were programmed to listen on ports 1 to 65535 and accept TCP connections but did not send any other data back to the sender. All our machines ran Ubuntu 22.04, and we verified their advertised locations using the IP2Location [62] database. We have summarized the timeline of our experiments and the vantage points used in each in Table 1.

Dropping Outgoing RSTs on the VPSes. We configured `iptables` rules on both the client and sink servers to drop all outgoing `RST` packets. This configuration ensures that any `RST` packet received by the client side can be reliably attributed to middlebox injections.

Triggering TLS SNI-based Censorship. We trigger censorship by sending a TLS ClientHello with potentially censored domain names in the SNI field. Since the sink servers are configured to not respond with any data packets and not tear down connections before observing a `FIN` or a `RST` packet, we expect any `RST` packets received are indeed injected packets from a firewall. We mark a domain name as censored if a `RST` packet is received for a TLS ClientHello containing the domain name.

Triggering HTTP Host-based Censorship. To trigger HTTP censorship, we sent HTTP GET requests with the forbidden domain name in the Host header of the request:

```
GET / HTTP/1.1\r\nHost: example.com\r\n
```

While later we found that the Henan Firewall does not require a full TCP handshake to trigger blocking, we still complete a TCP handshake before sending the HTTP request, making our testing methods consistent against the Henan Firewall and the GFW. We mark a domain name as censored if a `RST` packet is received for an HTTP GET request containing the domain name.

Isolating the Henan Firewall. To distinguish Henan Firewall responses from the GFW's, we identify several fingerprints unique to each firewall. Prior work has documented the GFW disrupts connections by injecting up to three `RST+ACK` to both sides of the connection whenever a TLS ClientHello message with a forbidden Server Name Indication field is observed [2], [63]. In contrast, the Henan Firewall injects a single `RST+ACK` packet to only the client side of a connection. In addition, the Henan Firewall's `RST+ACK` packet contains a payload, making it easy to distinguish from GFW responses. We expand on this in Section 4.4.

Finally, we send probes from vantage points in Henan to servers in Guangzhou, Beijing, and Shanghai to make sure that our traffic is not routed outside China (where it may encounter the GFW) but is still subject to the regional firewall in Henan.

Limitations. Our measurements in Henan Province are limited to a single Autonomous System (AS), China Unicom (AS 4837), due to the difficulty of obtaining diverse vantage points in China that could be ethically used for censorship measurement. Consequently, our empirical findings are confined to this single Internet Service Provider (ISP), limiting our ability to confirm or characterize censorship practices across other ISPs or ASes in Henan.

While user reports suggest that ISPs in Henan employ region-specific censorship, the censorship implementations are reportedly distinct [64]. For instance, Github user 5e2t reported that China Mobile Henan censored traffic on its cellular network and was capable of reassembling closely spaced TCP packets [64], which differs from the behavior we observed on China Unicom in Section 4.3. Therefore, our results should be interpreted as reflective only of China Unicom Henan's censorship implementation, not necessarily indicative of province-wide practices across all ISPs.

4.2. What Traffic Is Targeted

Does the Henan Firewall sample traffic to monitor and censor? The censor has been observed to only monitor and censor a fraction of traffic, potentially as a way to reduce the computation load on its censorship devices [15 §6.3]. We, however, did not observe any traffic sampling or probabilistic blocking behaviors from the Henan Firewall. We observed that the Henan Firewall consistently blocked domains listed on its blocklist. We sent 1,000 consecutive ClientHello messages containing a forbidden domain name, each request made over a unique port pair with small delays. We received the TCP `RST` packets for every connection we made, indicating a 100% triggering rate of censorship for censored domains of the Henan Firewall.

What ports does the Henan Firewall monitor? Previous works have shown that the GFW TLS ESNI censorship middleboxes monitor all ports i.e. 1-65535 [10]. To measure the Henan Firewall, we sent TLS ClientHello messages, with a known blocked SNI to all ports of our sink server in Guangzhou, China. We found that the Henan Firewall, similar to the GFW, monitors TLS traffic going to any TCP port number, ranging between 1 and 65535.

Is the Henan Firewall bidirectional? Owing to the inherent limitations of obtaining vantage points in a censored region, researchers typically opt for performing measurements from the outside in rather than inside out. Particularly in China, works that study the GFW [3], [4], [9] used vantage points outside China because of its bidirectional nature. However, as mentioned in Section 3, sending probes from outside China does not trigger the Henan Firewall as it only censors traffic going out of Henan. As shown in Figure 3, we tested this by sending TLS ClientHello messages with different SNI values in the Tranco list [60] 5YZ7N, between nodes in Henan and nodes in other regions of China. We found that only traffic going out of Henan was blocked by the regional firewall. Similar asymmetric blocking behaviors were also observed in the GFW by prior work [9], [10].

4.3. How the Henan Firewall Parses Connections

In this section, we look at the parsing logic of the Henan Firewall and the GFW. We perform experiments to check the TCP handshake requirements for triggering the Henan Firewall and the GFW. We also use DPYProxy [65] to test for TCP and TLS reassembly capabilities, as well as the presence of residual censorship in the two firewalls. We summarize our findings in Table 3.

TCP handshake completeness requirements. The middleboxes designers often need to make a trade-off between the complexity of the parsing logic and the efficiency of the traffic analysis operations. For example, due to asymmetric routing nature of the Internet, and the fact that Henan Firewall and the GFW are not always immediate neighbors of the client or the server (as shown in Table 5), the middleboxes may only be able to observe flows in one direction. This nature often makes the middleboxes' designers not require to observe a complete TCP three-way handshake to track TCP connection and conduct censorship. On October 10, 2024, we tested the requirements of the TCP handshake completeness for the Henan Firewall and the GFW from our vantage point in Henan. We sent a single TCP packet whose payload is a TLS Clienthello message contained a forbidden domain name `011.com` as the SNI, preceded by 1) a SYN packet from client, or 2) a SYN packet from the client and a SYN+ACK packet from the server, or 3) no packet at all.

TABLE 3: Parsing logic of the GFW and the Henan Firewall. The Henan Firewall appears to be stateless and less robust against typical network variability than the GFW.

	GFW	Henan Firewall
Require SYN	✓	✗
Require SYN+ACK	✗	✗
TCP Reassembly	✓	✗
TLS Reassembly	✗	✗
TCP Header Length	Arbitrary	20 bytes Only

As summarized in Table 3, while the GFW requires to observe a SYN packet from the client (but not a SYN+ACK packet from the server) to trigger the censorship [9], the Henan Firewall does not require to observe any TCP handshake packet to be triggered.

TCP segmentation. TCP segmentation enables the splitting of larger TCP payloads into smaller ones. In the context of circumvention, splitting a TLS ClientHello message into multiple TCP segments has been used to confuse stateless censors that do not reassemble packets. However, we confirm that the GFW performs TCP reassembly and thus, is stateful. On the other hand, we found that the Henan Firewall does not perform TCP reassembly and thus, it is possible to bypass it by splitting the TCP payload of the ClientHello into multiple TCP segments, with the SNI distributed between the segments. We tested this by initiating a TLS connection from our vantage point in Henan to our VPS in Guangzhou with a forbidden SNI and splitting the ClientHello into two segments, with the second segment containing the forbidden domain name. We observed that while a complete ClientHello message was blocked by the Henan Firewall, not putting a complete SNI extension in the first segment would bypass the Henan Firewall.

TLS fragmentation. While TCP segmentation has been long known to be used to bypass stateless censors, the use of TLS fragmentation was only recently analyzed by Niere et al. [65] and implemented in their DPYProxy tool. Before a TLS message is encapsulated within a TCP segment, it is first enclosed in what is known as a TLS record. Given that the maximum size of a TLS message exceeds the maximum allowable size for a TLS record, the TLS standard permits the division of TLS messages across several TLS records. Niere et al. [65] found that the GFW did not perform TLS reassembly and it is thus possible to bypass it by

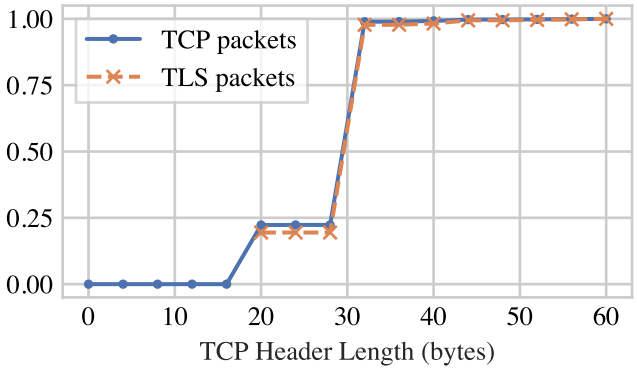


Figure 5: The distribution of the TCP header length fields of all TCP and TLS packets during a one-hour captured on a university network on October 31, 2024. In total, we captured approximately 23.1 billion TCP packets and 5.0 billion TLS packets. Only 22% of any TCP packets have a header length of 20 bytes, while only 19% of any TLS packets have a header length of 20 bytes. This evaluation result suggests that the Henan Firewall has only being able to censor around 20% of the targeted connections.

fragmenting TLS ClientHello messages over multiple TLS records, wherein the SNI is split into multiple TLS segments within the same TCP payload. We confirm that, as of April 4, 2024, both the Henan Firewall and the GFW do not perform TLS reassembly and thus, it is possible to bypass them via TLS ClientHello fragmentation.

TCP header length has to be 20 bytes. The four most significant bits of the 13th byte of the TCP header represent the TCP Data Offset, which specifies the length of the TCP header in 32-bit words. The minimum value of the TCP Data Offset field is 5 words (20 bytes) when no TCP options are present, and the maximum value is 15 words (60 bytes).

We found that the Henan Firewall required the TCP header length to be exactly 20 bytes to correctly parse and block the TLS ClientHello or HTTP request messages. We tested this by sending forbidden messages (e.g. TLS ClientHello messages with a forbidden SNI `011.com`) from our vantage point in Henan to our sink server in Guangzhou on October 17, 2024, with different TCP options set in their TCP headers. While varying the TCP header length, we made sure that the TCP options are always a multiple of four bytes to comply with the 32-bit word alignment requirement of TCP header. The TCP options we tested include common TCP ones like Maximum Segment Size (MSS), Window Scale, Timestamps, Selective Acknowledgment Permitted (SAckOk), No Option (NOP), End of Option List (EOL), as well as self-defined TCP options that are not commonly used. We found that as long as any TCP option was set, the Henan Firewall did not block the connection.

An intuitive hypothesis to explain this strange behavior is that the Henan Firewall does not parse the TCP header length field in the TCP header, and falsely assumes that the TCP header length is always 20 bytes. This way, when a TCP header has more than 20 bytes due to TCP options, it will treat the TCP options as part of the TCP payload and will thus fail to recognize a complete TLS ClientHello or an HTTP request message. However, we falsified this hypothesis and confirmed the Henan Firewall did parse the TCP header length field. In particular, we sent TLS ClientHello messages with a forbidden SNI `011.com` with no TCP option set in its TCP header, confirming that this message was blocked by the Henan Firewall. If the Henan Firewall does not parse the TCP header length field in the TCP header, then regardless of the TCP header length value we put in the TCP header, this message should be blocked. We changed the 4-bit TCP header length field in the TCP header to be all 2^4 possible values from 0 to 15, and recomputed the correct TCP checksum for each TCP packet, and found that the Henan Firewall only blocked a connection when its TCP header length value was 5 words (20 bytes). This experiment indicates that *the Henan Firewall did parse the TCP header length field in the TCP header, but had a condition to only block a connection when its TCP header length is 20 bytes.*

Although we were unable to determine the rationale behind this condition—possibly an oversight by the censor—it raises an important question about how much real-world traffic evaded detection due to this condition. We conducted a test on a university network in the United States. Specifically, we used Retina [66] to capture the TCP header length fields for all traffic on the campus network over a one-hour period from 3:56:14 PM to 4:56:14 PM (UTC–7) on October 31, 2024. In total, we collected 23.1 billion TCP packets and 5.0 billion TLS packets. As shown in Figure 5, only 22% of the TCP packets had a header length of 20 bytes, and only 19% of the TLS packets had a header length of 20 bytes. This result suggests that the Henan Firewall may only be able to censor around 20% of the targeted connections.

4.4. How the Henan Firewall Blocks Traffic

Does the Henan Firewall employ residual censorship? Residual censorship is a mechanism used by censors in which after a censorship event is detected between two hosts, the censor continues to block all subsequent connections between the two hosts (SrcIP, DstIP, DstPort - three tuple) for a certain duration typically 90 s or 180 s. The phenomenon has been documented by multiple previous works, studying the GFW [6], [8], [67]. We found that the Henan Firewall does not perform any residual censorship. We were able to make connections with the same three-tuple subsequent to any reset injections from the Henan Firewall.

Fingerprinting the injection behaviors. Continuing the efforts of fingerprinting the GFW's evolving injection behaviors [68] [69] [65 §3.1] [70 §7.1.6] [7 §2.1], we fingerprint the TCP `RST` packets injected by the GFW and the Henan Firewall. Using the `RST` packets collected in Section 4.5, we analyze their packet features such as IP ID, IP TTL, TCP Flags, TCP Payload, and Payload Length.

TABLE 4: A comparison of the injection behaviors and packet fingerprints of the Henan Firewall and the three types of GFW TCP RST injectors. All injections were triggered by TLS SNI-based censorship. The IP TTLs shown are the observed values; their initial values should be higher. The 'C' and 'S' refer to the client and server.

	GFW (I)	GFW (II)	GFW (III)	Henan Firewall
Observed IP TTL	55–118	39–238	248	58
IP ID	00 00	00 A3 – FE 5F	99 16 – 99 33	00 01
IP Flag (DF)	0	1	0	0
TCP Payload Len	0 byte	0 byte	0 byte	10 bytes
TCP Payload	-	-	-	01 02 03 04 05 06 07 08 09 00
TCP Flags	RST	RST+ACK	RST+ACK	RST+ACK
Packet Counts	x1	x3	x1	x1
Targeted Hosts	C&S	C&S	C&S	C
Residual Duration	180 s	180 s	180 s	-

Table 4 compares the reset packet-injection behaviors of the Henan Firewall against three types from the GFW (I, II, III). While the GFW injection mechanisms target both client (C) and server (S), the Henan Firewall exclusively injects reset packets to the client side.

Examining IP and TCP flags of the RST packets from firewalls, we observed that the Henan Firewall sent a single TCP+RST packet with the IP DF (Do Not Fragment) flag unset. Among the GFW injectors, type I sends a single RST packet without ACK with the IP DF flag unset, type II sends three duplicate and identical RST+ACK with IP DF set, and type III sends a single RST+ACK packet with IP DF unset.

The observed IP TTL values of the TCP RST packets by the three GFW injectors exhibited a range of values: 55–118 for type I, 39–238 for type II, and a fixed value of 248 for type III. We observed a fixed IP TTL value of 58 for the RST+ACK packets from the Henan Firewall. We note that these values are the IP TTL values observed by the client; the initial TTL values set by the censorship devices would have been higher, subsequently reduced by the number of network hops from the censorship devices to the client.

Regarding IP ID values, we observed that the Type I GFW inject one RST packet with a fixed IP ID of 0x0000, the Type II GFW injects three RST+ACK packets with a range of IP ID values from 0x00A3 to 0xFE5F (163–65119), and the Type III GFW injects RST+ACK packets with a range of IP ID values from 0x9916 to 0x9933 (39190–39219). The Henan Firewall, on the other hand, had a fixed IP ID value of 0x0001 for its TCP RST packets.

The most distinctive fingerprint of the Henan Firewall's RST packets is their 10-byte TCP payload pattern 01 02 03 04 05 06 07 08 09 00, a characteristic not found in any of the GFW injectors. While RFC 9293 states that "TCP implementations SHOULD allow a received RST segment to include data (SHLD-2)" [71 §3.5.3], it is still very rare to see a RST packet with a payload in real world. In Section 6, we introduce a circumvention technique that leverages this distinct fingerprint to bypass the Henan Firewall.

4.5. Where Are the Censorship Devices Deployed

To find where in the network the Henan regional firewall devices are located, we used a variant on our methodology to measure the network time and TTL-hop distance of the censorship devices from our Henan client.

First, we sent ClientHello packets from our vantage point in Zhengzhou, Henan Province to our sink servers in Guangzhou and San Francisco independently, and measured the time difference between when we sent a ClientHello and when we received a RST for any of the connections. We utilized the top one million domains from the Tranco list performed the experiment four times in a day and recorded any RST packets that we received.

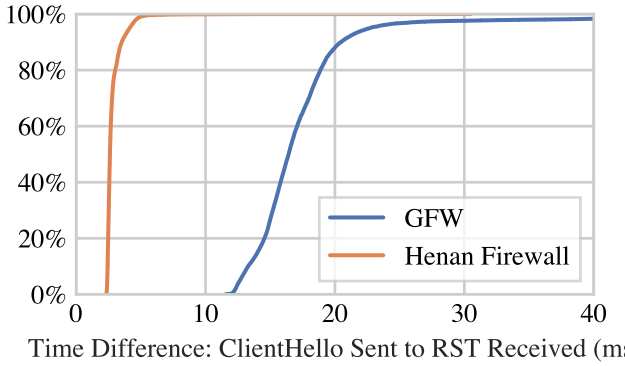


Figure 6: Cumulative distribution of the time difference between sending a TLS ClientHello packet containing a forbidden domain name and receiving the first forged TCP RST packet from the censorship devices.

Figure 6 shows the cumulative distribution of the time difference between sending a ClientHello message and receiving the first TCP RST packet by the Henan censorship devices and the GFW. The analysis is based on 36,480 RST packets received from Henan and 16,649 RST packets collected from the GFW between October 2 and December 8, 2023. Although the GFW can inject more than three RST packets for a blocked connection, we account only for the first RST packet received since it is the one that initiates the connection tear down. The graph clearly shows the difference in latencies: the delta timing differences indicate that Henan censorship devices were located closer to the client, whereas the GFW was situated at the national gateway. Specifically, the delta times for the GFW ranged from 11.52 ms to 445.38 ms (with a mean of 17.98 ms), while those for the Henan devices ranged from 2.30 ms to 30.49 ms (with a mean of 2.82 ms). This evidence strongly suggests that the regional censorship in Henan is independently deployed and in closer proximity to our vantage points, implying that these censorship devices are located within the Henan province.

Second, to identify the exact network hop where censorship occurs, we used a TTL-limited probing method based on traceroute. Specifically, we sent TLS ClientHello packets containing a known censored domain, gradually increasing

the IP TTL value of the probes until an injected RST packet was observed. The TTL of the probe that triggered the RST reflects the hop count to the censoring device. This approach is similar to that used in prior work such as CenTrace [72].

TABLE 5: Results from our TTL-limited probing experiment, showing that the Henan middleboxes are two hops closer to our client compared to the GFW. We sent TLS ClientHello probes from Zhengzhou, Henan to a sink server in San Francisco, US, triggering two distinct middleboxes at different hops.

	Hops Away	ASN	ISP
Henan	5	4837	China Unicom Henan Province Network
GFW	7	4837	Backbone - China Unicom

Table 5 shows results from our measurements conducted in Zhengzhou, targeting a sink server in the US. We used `011.com` to trigger regional censorship (Henan) and `youtube.com` for national-level censorship (GFW). Our findings indicate that the Henan middlebox is located at hop 5 within China Unicom's provincial network, while the GFW appears at hop 7, deeper in the national backbone network. These results confirm that both censoring entities operate as on-path middleboxes, with the Henan device positioned closer to the client.

5. Understanding the Blocklists

We monitored and analyzed the websites blocked by the Henan Firewall and the GFW across time. We also inferred the underlying blocking rules employed.

5.1. Analyzing the Blocked Domains

Experiment setup. Due to the challenge of obtaining high-bandwidth machines in Henan, we divide our measurements into two parts. First, we perform daily tests on the top one million websites from the Tranco list 5YZ7N. Second, carried out weekly, we test 227 million domains sourced from the zone files of more than 1,000 Top-Level Domains (TLDs), obtained from the Centralized Zone Data Service (CZDS) of the Internet Corporation for Assigned Names and Numbers (ICANN) [73].

For our daily test of the Tranco top one million domains, we tested both TLS SNI and HTTP Host-based blocking by sending respective requests to servers we controlled in China. For each domain, for TLS SNI-based censorship, we sent four requests per day; for HTTP Host-based censorship, we sent two per day. For a given day, we mark a domain as blocked for that protocol if it receives a TCP RST in response to any of our requests.

Due to the bandwidth constraints, for the 227 million tested weekly, we send a single TLS request per domain each week to our server, and mark the domain as blocked if our request receives a TCP RST.

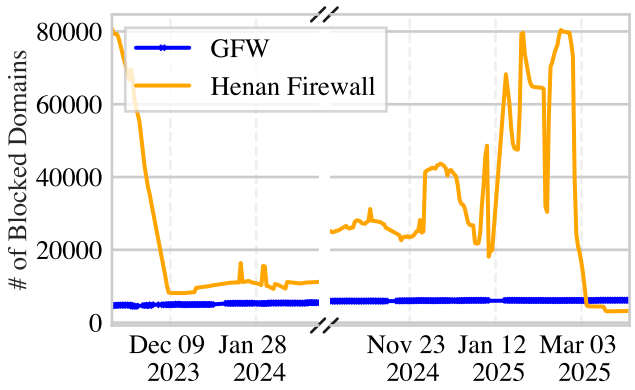


Figure 7: The numbers of domains blocked by the Henan Firewall and the GFW over time. We tested with a Tranco top one million domain list ID 5YZ7N, between November 5, 2023 and March 31, 2025, with a measurement gap between March 5 and October 7, 2024.

Experiment timeline. Table 1 summarizes the specific experiment timeline and vantage point usage. In particular, we failed to run the longitudinal experiments between March 5 and October 7, 2024. There were also minor data gaps, also reflected in Figure 7, due to unexpected disruptions of our VPSes in Guangzhou. Since we used the same machines to measure both the Henan Firewall and the GFW, the disruptions experienced by our sink servers in Guangzhou impacted our measurements of both firewalls. We thus removed these monior measurement gaps, counting towards an additional 25 days, from our analysis.

The Henan Firewall uses the same blocklist for HTTP Host-based and TLS SNI-based censorship. Prior work has shown that the GFW maintains different domain-based blocklists to censor different protocols [2 §4.1] [9 §5.2]. In contrast, we find the Henan Firewall uses the same blocklist for both HTTP Host-based and TLS SNI-based censorship. In particular, we compare the lists of domains that were blocked by Henan's HTTP Host-based and TLS SNI-based censorship on the same day (November 14, 2024). A similar number of domains is blocked in each protocol: 24,795 domains blocked by HTTP Host-based censorship, and 24,974 domains blocked by TLS SNI-based censorship. The small 1% difference between these two lists is explained by measurement noise: we repeated the same detection for the divergent domains twice to reduce false negatives, and found the difference between lists disappeared.

Comparing the sizes of the blocklists over time. We monitored the changes to the blocklists of the Henan Firewall and the GFW over time. Figure 7 shows the total number of domains blocked by the Henan Firewall and the GFW. The Henan Firewall has a blocklist that remains much larger than the GFW blocklist until March 4, 2025.

The Henan Firewall frequently added and removed generic second-level domain blocking rules (e.g. *.com.au, *.net.br, *.gov.co), causing dramatic changes in the number of blocked domains. For example, Figure 7 shows a large

TABLE 6: Top ten TLDs censored by the GFW and the Henan Firewall over a period of three months. The Henan Firewall blocked more country code top-level domain (ccTLDs) than the GFW.

GFW		Henan	
TLD	**Blocklist %**	**TLD**	**Blocklist %**
.com	45.8%	.com	37.4%
.org	6.1%	.au	11.4%
.net	5.6%	.za	4.6%
.jp	2.4%	.net	4.5%
.cc	2.1%	.uk	4.1%
.de	1.7%	.org	4.0%
.xyz	1.7%	.in	2.9%
.in	1.7%	.jp	2.4%
.tw	1.5%	.tw	1.1%
.io	1.3%	.de	1.0%

consistent drop in the number of domains being blocked by the Henan Firewall, between November 10 and December 8, 2023. This drop was mostly due to the removal of at least 112 generic second-level domain blocking rules. In particular, the removal of the blocking rule *.com.au itself contributed to the unblocking of more than five thousands domains on November 22, 2023.

We observe that the blocklist used by the Henan Firewall also targets websites that are related to state or city governance from other countries. For instance, a majority of state government websites from the United States such as *texas.gov, seattle.gov, alabama.gov, nc.gov* are all blocked in Henan but not by the GFW. Compared to the 83 **.gov** domains that are seen in the GFW blocklist, we found 1002 **.gov** domains blocked by the Henan Firewall, showing an inclination to block anything that exhibits governance data or news from around the world. In fact, we noticed a trend in the Henan Firewall to target country code top-level domains (ccTLDs) more than the GFW as can be seen in Table 6. Some of these blocks were widespread: In 2024, Henan blocked all 5,334 **.com.au* domains we tested on Jan 19 and Feb 1–2, all 2,075 **.co.za* domains Feb 15–Mar 4, and all 1,547 **.org.uk* domains Feb 8–Mar 4. These may be instances of overblocking, where the firewall contains an overly broad rule. It is unclear to us why the Henan Firewall would repetitively block and unblock these country code second level domains.

Henan Firewall's blocklist is more volatile than the GFW's. As shown in Figure 8, the Henan Firewall has more volatile blocking policy than the GFW's blocklist. While 75% of blocked domains were censored for fewer than 51 days by the Henan Firewall, more than 50% of the domains ever censored by the GFW were blocked during the entire measurement period (256 days). Domains blocked by the GFW had longer censorship durations (mean: 173.8 days; median: 256 days) compared to those blocked by the Henan Firewall (mean: 35.7 days; median: 21 days).

As mentioned above, this volatile blocking policy of the Henan Firewall is also mostly due to the frequent addition and removal of generic second-level domain blocking rules. For example, Figure 7 shows two spikes in the number of domains blocked by the Henan Firewall between January 11 and January 12, 2024, as well as between February 1 and February 3, 2024. They are mostly due to the addition and removal of the blocking rule *.com.au. It worth noting that even when the rule *.com.au was removed, for example on January 12 and February 3, 2024, the Henan Firewall still blocked 44 and 26 domains ended with .com.au, respectively. This observation suggests that the blocking rule can be finer grained than the second-level domain.

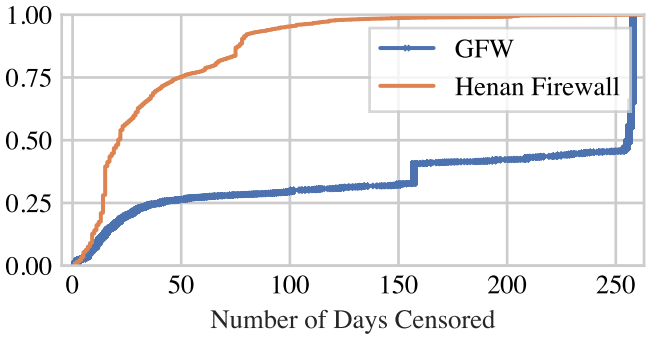


Figure 8: The censorship duration of all domains (ever) blocked by the GFW and the Henan Firewall between November 5, 2023 and March 31, 2025, with a measurement gap between March 5 and October 7, 2024. Compared to the GFW, the Henan Firewall has a more volatile blocking policy, with a larger proportion of domains being blocked for a shorter duration.

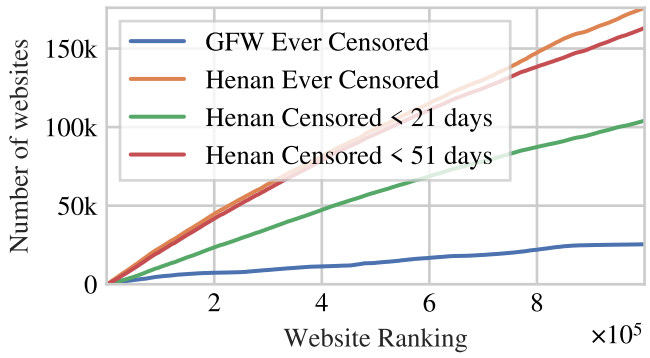


Figure 9: Cumulative distribution of the domains blocked by the GFW and the Henan Firewall in the Tranco top one million list 5YZ7N. The data is collected between November 5, 2023 to March 31, 2025, with a measurement gap between March 5 and October 7, 2024.

Do the two firewalls target similar websites? Figure 9 shows the cumulative distribution of the domains blocked by the GFW and by the Henan regional censorship devices among the top one million Tranco domains over our measurement period of nine months.

For the GFW, we classify a domain as blocked if it was blocked at least once during our measurement period. Due to the volatility of the Henan Firewall's blocklist, we categorize domains into three classes: those that were ever blocked, those blocked for less than 21 days, and those blocked for less than 51 days. We selected these thresholds based on our observation of average blocking durations for both firewalls, as shown in Figure 8.

During our measurement period, we cumulatively observed 25,441 domains censored by the GFW, while 175,925 domains were blocked at least once by the Henan Firewall. Of the domains censored by the Henan Firewall, our analysis identified 104,100 domains with blocking periods under 21 days, while 163,083 domains experienced blocking durations shorter than 51 days.

Looking at the cumulative distribution and the ranking of the domains, we found that the most popular domains were more likely to be blocked by both the GFW and the Henan Firewall. The Henan Firewall is more homogeneous in blocking domains in terms of their popularity whereas the GFW's blocklist exhibits a more heterogeneous distribution. While the GFW firewall targets the more popular websites, as can be seen from the graph, the Henan Firewall targets the websites more uniformly. However, the sizes of the two blocklists provide a stark contrast between the two firewalls.

Overlap between the two blocklists. To understand the sizes and overlap of the two blocklists, we perform a long-running experiment to test 227 million domains on a weekly basis, between December 26, 2023 and March 31, 2025. Figure 10 shows the accumulated blocklists of the GFW and the Henan Firewall. During the experiment, the Henan Firewall blocked 4,196,532 domains—more than five times the 741,542 domains ever blocked by the GFW. There are 479,247 domains blocked by both firewalls. The Jaccard index between the two blocklists is approximately 0.0885, indicating they share under 9% similarity and are therefore largely independent yet complementary in their coverage.

Categorizing the blocked domains. we used the *whoisxmlapi.com* [74] website categorization service to classify the blocklists obtained for each firewall between November 21, 2023, and January 15, 2024. We acknowledge that not all domains could be categorized, as some were inactive or did not host content. Table 7 shows the top ten categories of censored domains for each firewall.

An interesting point that we note here is that the Henan Firewall targets Business, Economy, Computer and Internet Information domains more than the GFW. More than 35% of the total domains appearing on the blocklist of the Henan Firewall were from these two categories. To find the reason behind the focus on these categories, we hypothesize that the province of Henan has been a center of a lot of financial controversies, with the most prominent being the mass protests in 2022 that were a result of a financial scandal involving local lenders [75]. Given the financial scandals targeting state-controlled financial institutions, it is very probable that the state wants to limit access to information that is relevant to the economy of the area. On the other side, it could be a part of the national policy to censor critics of the country's business and economic policies.

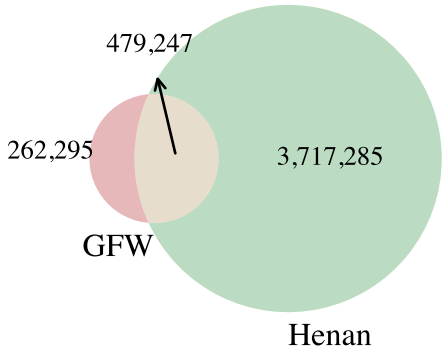


Figure 10: Venn diagram of the cumulative domains ever blocked by the GFW and the Henan Firewall. We conducted weekly testing of 227 million domains between December 26, 2023 and March 31, 2025 (with a measurment gap between March 5 and October 7, 2024). The Henan blocklist is more than five times the size of the GFW blocklist.

TABLE 7: The top categories of domains blocked by the Henan Firewall and the GFW among the top one million Tranco domains. Categories not in the top ten of each firewall are marked as "-".

Category	Henan		GFW	
	Count	Portion (%)	Count	Portion (%)
Business	4861	26.9	1183	15.3
Computer	2517	13.9	642	8.3
Pornography	2394	13.2	2207	28.6
Gambling	1276	7.1	–	–
Society	1265	7.0	459	5.9
Shopping	1261	7.0	288	3.7
Travel	1230	6.8	–	–
Entertainment	1134	6.3	548	7.1
Education	1104	6.1	–	–
Uncategorized	1057	5.8	395	5.1
News	–	–	1378	17.9
Personal Sites	–	–	313	4.1
Streaming Media	–	–	305	4.0

The GFW on the other hand, targets more of the news and media, as well as adult content domains. This is in line with the long-standing understanding of the GFW that it aims to limit more of the news, morally sensitive and politically sensitive content.

5.2. Identifying the Blocking Rules

Another way to view how each of the firewalls configures filter rules is to infer likely regular expressions used for blocklist matching. As noted by Anonymous et al. [22 §6] and Hoang et al. [3 §4.1] in their study of the GFW's DNS censorship, the GFW blocks domains using rules that may target second-level domains, top-level domains, and/or subdomains. They developed a methodology to encompass the blocking rules applied by the GFW. We used a similar

TABLE 8: Permutations used to test the blocking rules of the Henan Firewall and the GFW. The placeholder `{str}` represents strings that, alone or combined with others, should not trigger censorship. In this work, we used the string ZZZZ.

Test	Pattern	Test	Pattern
Test 0	{str}domain{str}	Test 5	{str}domain
Test 1	domain	Test 6	{str}.domain.{str}
Test 2	domain.{str}	Test 7	{str}.domain{str}
Test 3	domain{str}	Test 8	{str}domain.{str}
Test 4	{str}.domain		

TABLE 9: We infer the regex equivalents of blocking rules employed by the GFW and the Henan Firewall. In total, the GFW and the Henan Firewall employ 24 and 5 unique regex patterns, respectively. The table only shows the regex patterns that have more than ten occurrences for the GFW.

Inferred Regex	Tests Hit	Rule Count (Portion)			
		GFW		Henan	
`^(.*\.)?keyword$`	1&4	163,355	85%	248,770	64%
`^keyword$`	1	17,764	9.3%	3	0.0%
`^(.*\.)?keyword`	1–4&6&7	7,272	3.8%	–	–
`keyword$`	1&4&5	2,483	1.3%	139,575	36%
`keyword`	0–8	647	0.3%	–	–
`\.keyword$`	4	429	0.2%	4	0.0%
`^keyword`	1&2&3	36	0.0%	–	–

methodology based on the permutations listed in Table 8 to infer blocking rules for both the Henan Firewall and GFW. We note that our inferred regular expressions may not fully reflect the rules employed by the censor, as our permutations can miss regular expressions based on second-level domains or more complex regular expressions such as `*.gov*` that we observe the Henan Firewall blocking. Nonetheless, our inferred rules allow us to identify structural differences in the blocklists of the Henan Firewall compared to the GFW.

As shown in Table 8, we generated nine permutations for each censored domain identified in our daily measurement experiment (Section 5), by prepending and/or appending a fixed string to the domain name. This methodology was used by Anonymous et al. [22 §6] in 2014 and Hoang et al. [3 §4.1] in 2021. We chose the pattern string ZZZZ to construct each permutation in this work. We then sent ClientHellos with SNI containing each permutation independently to our sink servers and recorded the results for each testing. This experiment was conducted four times daily during our measurement period.

As shown in Table 9, the most popular blocking regex pattern used by both the Henan Firewall and the GFW was `^(.*\.)?keyword$`. This pattern meant to be used to block a domain and its subdomains. The second most popular blocking regex pattern used by the GFW was `^keyword$`, which was used to only block the domain name itself, not its subdomains. The third most popular blocking regex pattern used by the GFW was `^(.*\.)?keyword`, which was likely to be a mistake of not including the end anchor in the regex pattern. Interestingly, unlike the GFW, which sometimes employs regex patterns without end anchors, the Henan Firewall always includes end anchors in its regex patterns. This result could be because of a more carefully and consistently maintained blocklist, or perhaps the censorship implementation itself enforces the use of end-anchored regex patterns to prevent potential mistakes made by human.

6. Circumvention Strategies

Based on the parsing logic flaws we identified in Section 4.3, as well as the injection behaviors and fingerprints we observed in Section 4.4, we introduce simple but effective strategies to bypass the Henan Firewall. All strategies require only changes from the client-side, without cooperation from the server side, making them easy to employ and adopt. These strategies have already been implemented in various popular circumvention tools, including but not limited to Xray [76], GoodbyeDPI [77], and Shadowrocket [78].

Enable any TCP option field. As detailed in Section 4.3, the Henan Firewall can only parse and block TCP packets with a 20-byte header. Enabling any TCP option on an operating system will result in a TCP header longer than 20 bytes. While this circumvention solution relies on the unusual implementations of the Henan Firewall, it is nonetheless a feature that users or circumvention tools could easily employ to evade censorship. For instance, enabling TCP Timestamps (disabled by default on some version of the Windows) would prevent the Henan Firewall from blocking connections [56], [59].

Discard TCP `RST` packets with specific payload. As shown in Section 4.4, the Henan Firewall injects a TCP `RST` packet with an unusual 10-byte payload `01 02 03 04 05 06 07 08 09 00`. Its uniquesness allows the client drop only the `RST` packets injected by the Henan Firewall, while keeping the `RST` packets sent by the server. Typically, dropping TCP `RST` packets sent to the client is not enough to evade TCP `RST` censorship by the GFW, as the GFW also injects `RST` packets to the server. However, as explained in Section 4.4, the Henan Firewall only injects `RST` packets to the client, and thus dropping the `RST` packets sent to the client is sufficient to evade censorship. This circumvention strategy can be easily applied via `iptables` rules, similar to the ones introduced by Clayton et al. [6 §5].

Segment or fragment TLS ClientHello into multiple packets. As expalined in Section 4.3, the Henan Firewall does not perform TCP reassembly, and neither the Henan Firewall nor the GFW performs TLS reassembly [65]. Thus, clients can segment TCP packets or fragment TLS ClientHello messages over multiple TLS records to evade the Henan firewall [65]. As long as the TCP packets carrying the beginning part of ClientHello messages does not contain a complete SNI extension, one can bypass the Henan Firewall. Performing this fragmentation may require TLS libraries such as uTLS [79] that provide fine-grained control over the messages sent, or purposely built circumvention tools like DPYProxy [77] that can fragment records made by

a browser. Popular circumvention tools such as Xray [76] and Shadowrocket [78] have also implemented this TCP segmentation strategy [80].

7. Ethics

Censorship measurement studies, especially in authoritarian regimes, require careful ethical considerations and continuous evaluation of the potential risks involved throughout the entire research process. In this work, we conducted all of our censorship measurements from machines we controlled, with network traffic generated automatically by our programs. This approach is a common practice in censorship measurement studies to mitigate the risk of overwhelming other hosts on the Internet and imposing any risks on users [3], [4], [9], [14], [15]. When analyzing the real-world traffic on the university network tap, we only collected the TCP header length fields of the packets without capturing any human identifiable or sensitive information. Since IRB approval is thus not applicable for this study (as it does not involve human subjects), we followed the ethical guidelines outlined in the Menlo Report [81]. Our research team also consulted experts with a deep understanding of Chinese censorship and its legal concerns. Below, we discuss the potential risks we identified and the steps we took to mitigate them [81 §C.3.2].

Traffic analysis. To evaluate the effectiveness of the Henan Firewall, we measured the TCP header length fields of all TCP packets on the university network tap. The use of this network tap was approved by the university's privacy and security office. We also worked closely with the campus networking and security teams who has experience managing similar projects. This approval and collaboration ensured that we followed standard security procedures, complied with network use policy, respected user privacy, and minimized the network's attack surface. Additionally, we designed the tap to only receive a copy of traffic, ensuring no impact on network users in case of system failure.

We designed our experiment to avoid collecting any potentially sensitive information, such as IP addresses, which could be linked to individuals. Specifically, we only collected the 4-bit Data Offset fields from all TCP headers in an aggregated manner. We never inspected or logged any raw traffic data. We practiced the principle of least privilege by restricting access to the network tap to a limited, authorized subset of our team.

Vantage points. Obtaining vantage points within censored areas has become increasingly challenging. However, two key research questions we aim to answer require diverse vantage point coverage in China: 1) Is the Henan Firewall also deployed in other provinces in China? 2) Do other provinces deploy their regional censorship apparatus as well? We took extra care to find the right balance between finding as diverse vantage points as possible and the potential risks it carries [81 §C.3.2]. For example, while using residential vantage points would have allowed us to observe censorship in China from more network locations, we decided not to use them due to the potential risks to the uninformed users [61].

We also explored the possibility of measuring the Henan Firewall remotely from outside of the province or China, which will further reduce the risks of initiating connections from within the region; however, as introduce in Section 4.2, the Henan Firewall could not be triggered that way.

We thus, following the rationale and common practices outlined in prior work [14], [15], strategically selected vantage points provided by large commercial cloud providers to mitigate potential legal risks for individuals. We registered our VPS accounts with the accurate identity and contact information of one of our researchers who is neither a citizen nor a resident of China. Throughout our research, we received no complaints from the providers. To avoid the possibility of getting other cloud users' resources blocked by the censor, we assigned dedicated IP addresses to each of our machines.

Probing rate and design. To avoid overwhelming our vantage points and the network paths our probes traversed, we restricted the transmission speed directed to our sink servers. For experiments in Section 3 and Section 4, we limited the probing rate to no more than one connection per second; for experiments in Section 5, we set a hard limit on each client to send no more than 1 Mbps of traffic. While the risks and potential harms of our probes being logged by the censor is minimal, we also designed our experiments with plausible deniability in mind. That is, since our sink servers never replied with any ServerHello messages or HTTP responses, and no full TLS or HTTP connections were ever established, our measurement behaviors do not resemble users accessing censored websites.

8. Conclusion

In this paper, we expose and document an alarming sign in China's Internet censorship strategy: our measurements from seven different cities and provinces in China reveal a new regional firewall in Henan province. This Henan Firewall conducts HTTP Host-based and TLS SNI-based censorship for traffic going out of the province. It exhibits distinct characteristics compared to the GFW, including unique packet injection behaviors and fingerprints, different logic in tracking, parsing, and blocking connections, a once ten-times larger and more dynamic blocklist, and closer network location to the client. This localized censorship suggests a departure from China's centralized censorship apparatus, enabling local authorities to exert a greater degree of control within their respective regions. We propose simple but effective circumvention techniques to get around this emerging system in Henan, which have been implemented in vairous popular circumvention tools. We hope our study sounds the alarm to the broader censorship research community to be aware of and further study emerging regional censorship in China, and elsewhere.

Availability

To encourage future research and promote transparency and reproducibility, we have made the code, anonymized data, and constantly updated blocklists available. For improved accessibility, this paper is also available in HTML format in both English and Chinese. The project homepage is at: https://gfw.report/publications/sp25/en.

Acknowledgments

We thank our shepherd and other anonymous reviewers for their valuable comments and feedback. We also thank the brave users in China for immediately reporting and actively studying the blocking incidents, including, but not limited to, 5e2t, Hsukqi, ThEWiZaRd0fBsoD, louiesun, and lemon99ee. We are grateful to ValdikSS, radioactiveAHM, RPRX, Fangliding, GFW-knocker, sambali9, rrouzbeh, nekohasekai, znlihk, the V2Ray developers, the Hysteria developers, the Shadowrocket developers, and many other developers for their helpful discussions and/or for integrating TCP segmentation and/or TLS fragmentation features into their respective circumvention tools. We also thank Jackson Sippe, Jade Sheffey, Paul Flammarion, the Stanford Empirical Security Research Group, the Stanford University security and networking teams, and many others who prefer to remain anonymous for their helpful discussions and support. We thank Net4People BBS, NTC Party forum, Xray community, V2Ray community, and sing-box community for providing online space for censorship discussions. We are grateful to David Fifield for providing feedback, support, and guidance throughout the entire project.

The work was supported in part by the National Science Foundation (NSF) under grant numbers CNS-2145783, CNS-2319080, and CNS-2333965, by a Sloan Research Fellowship, and by the Young Faculty Award program of the Defense Advanced Research Projects Agency (DARPA) under the grant DARPA-RA-21-03-09-YFA9-FP-003. The views, opinions, and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

References

[1] H. Duan, N. Weaver, Z. Zhao, M. Hu, J. Liang, J. Jiang, K. Li, and V. Paxson, "Hold-On: Protecting against on-path DNS poisoning," in *Securing and Trusting Internet Names*. National Physical Laboratory, 2012. [Online]. Available: https://www.icir.org/vern/papers/hold-on.satin12.pdf

[2] Z. Chai, A. Ghafari, and A. Houmansadr, "On the importance of encrypted-SNI (ESNI) to censorship circumvention," in *Free and Open Communications on the Internet*. USENIX, 2019. [Online]. Available: https://www.usenix.org/system/files/foci19-paper_chai_update.pdf

[3] N. P. Hoang, A. A. Niaki, J. Dalek, J. Knockel, P. Lin, B. Marczak, M. Crete-Nishihata, P. Gill, and M. Polychronakis, "How great is the Great Firewall? Measuring China's DNS censorship," in *USENIX Security Symposium*. USENIX, 2021. [Online]. Available: https://www.usenix.org/system/files/sec21-hoang.pdf

[4] Anonymous, A. A. Niaki, N. P. Hoang, P. Gill, and A. Houmansadr, "Triplet censors: Demystifying Great Firewall's DNS censorship behavior," in *Free and Open Communications on the Internet*. USENIX, 2020. [Online]. Available: https://www.usenix.org/system/files/foci20-paper-anonymous_0.pdf

[5] S. Fan, J. Sippe, S. San, J. Sheffey, D. Fifield, A. Houmansadr, E. Wedwards, and E. Wustrow, "Wallbleed: A memory disclosure vulnerability in the Great Firewall of China," in *Network and Distributed System Security*. The Internet Society, 2025. [Online]. Available: https://gfw.report/publications/ndss25/data/paper/wallbleed.pdf

[6] R. Clayton, S. J. Murdoch, and R. N. M. Watson, "Ignoring the Great Firewall of China," in *Privacy Enhancing Technologies*. Springer, 2006, pp. 20–35. [Online]. Available: https://www.cl.cam.ac.uk/~rnc1/ignoring.pdf

[7] Z. Wang, Y. Cao, Z. Qian, C. Song, and S. V. Krishnamurthy, "Your state is not mine: A closer look at evading stateful Internet censorship," in *Internet Measurement Conference*. ACM, 2017. [Online]. Available: https://www.cs.ucr.edu/~krish/imc17.pdf

[8] R. Rambert, Z. Weinberg, D. Barradas, and N. Christin, "Chinese wall or Swiss cheese? keyword filtering in the Great Firewall of China," in *WWW*. ACM, 2021. [Online]. Available: https://censorbib.nymity.ch/pdf/Rambert2021a.pdf

[9] N. P. Hoang, J. Dalek, M. Crete-Nishihata, N. Christin, V. Yegneswaran, M. Polychronakis, and N. Feamster, "GFWeb: Measuring the Great Firewall's Web censorship at scale," in *USENIX Security Symposium*. USENIX, 2024. [Online]. Available: https://www.usenix.org/system/files/sec24fall-prepub-310-hoang.pdf

[10] K. Bock, iyouport, Anonymous, L.-H. Merino, D. Fifield, A. Houmansadr, and D. Levin. (2020, Aug.) Exposing and circumventing China's censorship of ESNI. [Online]. Available: https://github.com/net4people/bbs/issues/43#issuecomment-673322409

[11] K. Bock, G. Naval, K. Reese, and D. Levin, "Even censors have a backup: Examining China's double HTTPS censorship middleboxes," in *Free and Open Communications on the Internet*. ACM, 2021. [Online]. Available: https://doi.org/10.1145/3473604.3474559

[12] R. Ensafi, D. Fifield, P. Winter, N. Feamster, N. Weaver, and V. Paxson, "Examining how the Great Firewall discovers hidden circumvention servers," in *Internet Measurement Conference*. ACM, 2015. [Online]. Available: https://conferences2.sigcomm.org/imc/2015/papers/p445.pdf

[13] A. Dunna, C. O'Brien, and P. Gill, "Analyzing China's blocking of unpublished Tor bridges," in *Free and Open Communications on the Internet*. USENIX, 2018. [Online]. Available: https://www.usenix.org/system/files/conference/foci18/foci18-paper-dunna.pdf

[14] Alice, Bob, Carol, J. Beznazwy, and A. Houmansadr, "How China detects and blocks Shadowsocks," in *Internet Measurement Conference*. ACM, 2020. [Online]. Available: https://censorbib.nymity.ch/pdf/Alice2020a.pdf

[15] M. Wu, J. Sippe, D. Sivakumar, J. Burg, P. Anderson, X. Wang, K. Bock, A. Houmansadr, D. Levin, and E. Wustrow, "How the Great Firewall of China detects and blocks fully encrypted traffic," in *USENIX Security Symposium*. USENIX, 2023. [Online]. Available: https://www.usenix.org/system/files/sec23fall-prepub-234-wu-mingshi.pdf

[16] Sakamoto and E. Wedwards, "Bleeding wall: A hematologic examination on the Great Firewall," in *Free and Open Communications on the Internet*, 2024. [Online]. Available: https://www.petsymposium.org/foci/2024/foci-2024-0002.pdf

[17] X. Xu, Z. M. Mao, and J. A. Halderman, "Internet censorship in China: Where does the filtering occur?" in *Passive and Active Measurement Conference*. Springer, 2011, pp. 133–142. [Online]. Available: https://web.eecs.umich.edu/~zmao/Papers/china-censorship-pam11.pdf

[18] J. Wright, "Regional variation in Chinese Internet filtering," University of Oxford, Tech. Rep., 2012. [Online]. Available: https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2265775_code1448244.pdf?abstractid=2265775&mirid=3

[19] Anonymous, "Issue 2426 | XTLS/Xray-core," https://github.com/XTLS/Xray-core/issues/2426, 2023, (Accessed on 02/06/2024).

[20] G. Report. (2022, Oct.) Large scale blocking of TLS-based censorship circumvention tools in China. [Online]. Available: https://github.com/net4people/bbs/issues/129

[21] O. Farnan, A. Darer, and J. Wright, "Poisoning the well – exploring the Great Firewall's poisoned DNS responses," in *Workshop on Privacy in the Electronic Society*. ACM, 2016. [Online]. Available: https://dl.acm.org/authorize?N25517

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

[23] B. Dong, "A report about national DNS spoofing in China on Sept. 28th," Dynamic Internet Technology, Inc., Tech. Rep., Oct. 2002. [Online]. Available: https://web.archive.org/web/20021015121616/http://www.dit-inc.us/hj-09-02.html

[24] J. Zittrain and B. G. Edelman, "Internet filtering in China," *IEEE Internet Computing*, vol. 7, no. 2, pp. 70–77, Mar. 2003. [Online]. Available: https://nrs.harvard.edu/urn-3:HUL.InstRepos:9696319

[25] G. Lowe, P. Winters, and M. L. Marcus, "The great DNS wall of China," New York University, Tech. Rep., 2007. [Online]. Available: https://censorbib.nymity.ch/pdf/Lowe2007a.pdf

[26] Anonymous. (2020, Mar.) GFW archaeology: gfw-looking-glass.sh. [Online]. Available: https://github.com/net4people/bbs/issues/25

[27] C. Tang, "In-depth analysis of the Great Firewall of China," *http://www.cs.tufts.edu/comp/116/archive/fall2016/ctang.pdf*, 2016.

[28] K. Bock, A. Alaraj, Y. Fax, K. Hurley, E. Wustrow, and D. Levin, "Weaponizing middleboxes for TCP reflected amplification," in *USENIX Security Symposium*. USENIX, 2021. [Online]. Available: https://www.usenix.org/system/files/sec21-bock.pdf

[29] Sparks, Neo, Tank, Smith, and Dozer, "The collateral damage of Internet censorship by DNS injection," *SIGCOMM Computer Communication Review*, vol. 42, no. 3, pp. 21–27, 2012. [Online]. Available: https://conferences.sigcomm.org/sigcomm/2012/paper/ccr-paper266.pdf

[30] B. Marczak, N. Weaver, J. Dalek, R. Ensafi, D. Fifield, S. McKune, A. Rey, J. Scott-Railton, R. Deibert, and V. Paxson, "An analysis of China's "Great Cannon"," in *Free and Open Communications on the Internet*. USENIX, 2015. [Online]. Available: https://www.usenix.org/system/files/conference/foci15/foci15-paper-marczak.pdf

[31] P. Pearce, B. Jones, F. Li, R. Ensafi, N. Feamster, N. Weaver, and V. Paxson, "Global measurement of DNS manipulation," in *USENIX Security Symposium*. USENIX, 2017. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-pearce.pdf

[32] A. Filastò and J. Appelbaum, "OONI: Open observatory of network interference," in *Free and Open Communications on the Internet*. USENIX, 2012. [Online]. Available: https://www.usenix.org/system/files/conference/foci12/foci12-final12.pdf

[33] R. S. Raman, P. Shenoy, K. Kohls, and R. Ensafi, "Censored Planet: An Internet-wide, longitudinal censorship observatory," in *Computer and Communications Security*. ACM, 2020. [Online]. Available: https://www.ramakrishnansr.com/assets/censoredplanet.pdf

[34] A. A. Niaki, S. Cho, Z. Weinberg, N. P. Hoang, A. Razaghpanah, N. Christin, and P. Gill, "ICLab: A global, longitudinal internet censorship measurement platform," in *Symposium on Security & Privacy*. IEEE, 2020. [Online]. Available: https://people.cs.umass.edu/~phillipa/papers/oakland2020.pdf

[35] GreatFire, "GreatFire Analyzer," https://en.greatfire.org/analyzer, n.d., accessed: 2025-04-18.

[36] GreatFire, "Blocky," https://blocky.greatfire.org/, accessed: 2025-04-18.

[37] Anonymous and Amonymous. (2022, Oct.) Sharing a modified Shadowsocks as well as our thoughts on the cat-and-mouse game. [Online]. Available: https://github.com/net4people/bbs/issues/136

[38] P. Winter. (2013, Mar.) GFW actively probes obfs2bridges. [Online]. Available: https://bugs.torproject.org/8591

[39] P. Winter and S. Lindskog, "How the Great Firewall of China is blocking Tor," in *Free and Open Communications on the Internet*. USENIX, 2012. [Online]. Available: https://www.usenix.org/system/files/conference/foci12/foci12-final2.pdf

[40] T. Wilde. (2012) Knock knock knockin' on bridges' doors. [Online]. Available: https://blog.torproject.org/blog/knock-knock-knockin-bridges-doors

[41] Anonymous, Anonymous, Anonymous, D. Fifield, and A. Houmansadr. (2021, Jan.) A practical guide to defend against the GFW's latest active probing. [Online]. Available: https://github.com/net4people/bbs/issues/58

[42] Anonymous. (2021, Jan.) How to Deploy a Censorship Resistant Shadowsocks-libev Server. [Online]. Available: https://gfw.report/blog/ss_tutorial/en/

[43] S. Frolov, J. Wampler, and E. Wustrow, "Detecting probe-resistant proxies," in *Network and Distributed System Security*. The Internet Society, 2020. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/2020/02/23087.pdf

[44] S. Frolov and E. Wustrow, "HTTPT: A probe-resistant proxy," in *Free and Open Communications on the Internet*. USENIX, 2020. [Online]. Available: https://www.usenix.org/system/files/foci20-paper-frolov.pdf

[45] D. Xue, B. Mixon-Baca, ValdikSS, A. Ablove, B. Kujath, J. R. Crandall, and R. Ensafi, "TSPU: Russia's decentralized censorship system," in *Internet Measurement Conference*. ACM, 2022. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/3517745.3561461

[46] A. Ortwein, K. Bock, and D. Levin, "Towards a comprehensive understanding of Russian transit censorship," in *Free and Open Communications on the Internet*, 2023. [Online]. Available: https://www.petsymposium.org/foci/2023/foci-2023-0012.pdf

[47] R. Ramesh, R. S. Raman, M. Bernhard, V. Ongkowijaya, L. Evdokimov, A. Edmundson, S. Sprecher, M. Ikram, and R. Ensafi, "Decentralized control: A case study of Russia," in *Network and Distributed System Security*. The Internet Society, 2020. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/2020/02/23098.pdf

[48] T. K. Yadav, A. Sinha, D. Gosain, P. K. Sharma, and S. Chakravarty, "Where the light gets in: Analyzing web censorship mechanisms in India," in *Internet Measurement Conference*. ACM, 2018. [Online]. Available: https://delivery.acm.org/10.1145/3280000/3278555/p252-Yadav.pdf

[49] B. Liu, C. Lu, H. Duan, Y. Liu, Z. Li, S. Hao, and M. Yang, "Who is answering my queries: Understanding and characterizing interception of the DNS resolution path," in *USENIX Security Symposium*, Aug. 2018. [Online]. Available: https://www.usenix.org/system/files/conference/usenixsecurity18/sec18-liu_0.pdf

[50] Net4People, "Net4People BBS Issues." [Online]. Available: https://github.com/net4people/bbs/issues

[51] NTC Community, "NTC Party: "No Thought is a Crime" — Internet Censorship Circumvention Forum." [Online]. Available: https://ntc.party/

[52] XTLS, "Xray-core Project Issue Tracker." [Online]. Available: https://github.com/XTLS/Xray-core/issues

[53] V2Fly, "V2Ray Core Project Issue Tracker." [Online]. Available: https://github.com/v2fly/v2ray-core/issues

[54] SagerNet, "sing-box Project Issue Tracker." [Online]. Available: https://github.com/SagerNet/sing-box/issues

[55] Hysteria, "Hysteria Proxy Project Issue Tracker." [Online]. Available: https://github.com/apernet/hysteria/issues

[56] ThEWiZaRd0fBsoD, "在启用TCP Timestamp（TCP时间戳）后，GFW对obfs4的审查无效/ After enabling TCP Timestamp, GFW's censorship of obfs4 is rendered ineffective," https://github.com/net4people/bbs/issues/442, Jan 2025, (Accessed on April 18, 2025).

[57] ——, "The operators in Henan Province, China, seem to have less stringent censorship regarding IPV6," https://github.com/net4people/bbs/issues/416, Nov 2024, (Accessed on April 18, 2025).

[58] ghost, ""河南新上的SNI/HOST黑名单墙" (GitHub Discussion #3601)," https://github.com/XTLS/Xray-core/discussions/3601#discussioncomment-10293992, Aug 2024, (Accessed on April 18, 2025).

[59] H. Lee, "启用TCP Timestamps解决SNI阻断," https://blog.tsinbei.com/archives/1361/, Sep 2023, (Accessed on April 18, 2025).

[60] V. L. Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen, "Tranco: A research-oriented top sites ranking hardened against manipulation," in *Network and Distributed System Security Symposium 2019*, ser. NDSS '19, 2019.

[61] X. Mi, X. Feng, X. Liao, B. Liu, X. Wang, F. Qian, Z. Li, S. Alrwais, L. Sun, and Y. Liu, "Resident Evil: Understanding residential IP proxy as a dark service," in *2019 IEEE Symposium on Security and Privacy (SP)*, 2019, pp. 1185–1201. [Online]. Available: https://ieeexplore.ieee.org/document/8835239

[62] "IP2Location LITE IP address geolocation database." [Online]. Available: https://www.ip2location.com/database/ip2location

[63] K. Bock, G. Hughey, L.-H. Merino, T. Arya, D. Liscinsky, R. Pogosian, and D. Levin, "Come as you are: Helping unmodified clients bypass censorship with server-side evasion," in *SIGCOMM*. ACM, 2020. [Online]. Available: https://geneva.cs.umd.edu/papers/come-as-you-are.pdf

[64] 5e2t, "After enabling TCP timestamp, GFW's censorship of obfs4 is rendered ineffective," https://github.com/net4people/bbs/issues/442#issuecomment-2566913190, Jan 2025, (Accessed on April 7, 2025).

[65] N. Niere, S. Hebrok, J. Somorovsky, and R. Merget, "Poster: Circumventing the GFW with TLS record fragmentation," in *Computer and Communications Security*. ACM, 2023. [Online]. Available: https://nerd2.nrw/wp-content/uploads/2024/05/3576915.3624372.pdf

[66] G. Wan, F. Gong, T. Barbette, and Z. Durumeric, "Retina: analyzing 100GbE traffic on commodity hardware," in *ACM SIGCOMM*, 2022. [Online]. Available: https://zakird.com/papers/retina.pdf

[67] K. Bock, P. Bharadwaj, J. Singh, and D. Levin, "Your censor is my censor: Weaponizing censorship infrastructure for availability attacks," in *Workshop on Offensive Technologies*. IEEE, 2021. [Online]. Available: https://www.cs.umd.edu/~dml/papers/weaponizing_woot21.pdf

[68] klzgrad, "GFW技术报告：入侵防御系统的评测和问题," https://www.chinagfw.org/2009/09/gfw_21.html, Aug 2009, accessed: 2025-04-11.

[69] gfwrev, "HTTP URL/深度关键词检测," https://gfwrev.blogspot.com/2010/03/http-url.html, Mar 2010, accessed: 2025-04-07.

[70] N. Weaver, R. Sommer, and V. Paxson, "Detecting forged TCP reset packets," in *Network and Distributed System Security*. The Internet Society, 2009. [Online]. Available: https://www.ndss-symposium.org/wp-content/uploads/2017/09/weav.pdf

[71] W. Eddy, "Transmission Control Protocol (TCP)," RFC 9293, Aug. 2022. [Online]. Available: https://www.rfc-editor.org/info/rfc9293

[72] R. S. Raman, M. Wang, J. Dalek, J. Mayer, and R. Ensafi, "Network measurement methods for locating and examining censorship devices," in *Emerging Networking Experiments and Technologies*. ACM, 2022. [Online]. Available: https://dl.acm.org/doi/pdf/10.1145/3555050.3569133

[73] "Centralized Zone Data Service," https://czds.icann.org/home, (Accessed on 01/31/2024).

[74] "Website categorization api | domain category check | whoisxml api," https://website-categorization.whoisxmlapi.com/api, (Accessed on 04/25/2024).

[75] "Security forces in china attack protesters seeking frozen funds - the new york times," https://www.nytimes.com/2022/07/11/business/china-bank-protest.html, (Accessed on 04/25/2024).

[76] XRay developers. XRay. [Online]. Available: https://github.com/XTLS/Xray-core

[77] GoodbyeDPI developers. GoodbyeDPI. [Online]. Available: https://github.com/ValdikSS/GoodbyeDPI

[78] ShadowRocket developers. ShadowRocket. [Online]. Available: https://apps.apple.com/us/app/shadowrocket/id932747118

[79] S. Frolov and E. Wustrow, "The use of TLS in censorship circumvention," in *Network and Distributed System Security*. The Internet Society, 2019. [Online]. Available: https://tlsfingerprint.io/static/frolov2019.pdf

[80] XRay developers. XRay pull request. [Online]. Available: https://github.com/XTLS/Xray-core/pull/3660

[81] M. Bailey, D. Dittrich, E. Kenneally, and D. Maughan, "The menlo report," *IEEE Security and Privacy*, vol. 10, no. 2, p. 71–75, mar 2012. [Online]. Available: https://doi.org/10.1109/MSP.2012.52

Appendix A.
Meta-Review

The following meta-review was prepared by the program committee for the 2025 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

A.1. Summary

This paper empirically confirms anecdotal evidence that the Henan province in China had started to deploy regional censorship mechanisms, purposely more stringent than those employed by the great firewall itself. The paper delivers a comprehensive analysis of censorship carried out by the Henan Firewall, both on in/out and out/in directions, shedding light on its functioning, blocking policies, and residual censorship mechanisms. In addition, the paper inspects whether similar regional censorship is occurring in other Chinese provinces, finding no evidence of additional interference beyond that of the great firewall.

A.2. Scientific Contributions

- Independent Confirmation of Important Results with Limited Prior Research
- Provides a Valuable Step Forward in an Established Fieldtext

A.3. Reasons for Acceptance

1) This paper provides an independent confirmation of important results with limited prior research. The paper builds a measurement apparatus to confirm (and expand on the information provided by) anecdotal reports of regional censorship within the Henan province in China. Besides providing a systematic understanding of the existence and blocking behavior of this new type of regional censorship, the paper also independently confirms the asymmetry in blocking behavior of the great firewall.
2) The paper provides a valuable step forward in an established field by applying different known methodologies to exhaustively study a new phenomenon of censorship in a part of China. The measurements carried out in the presented study are sound and rely on true-and-tested measurement methodologies, providing insights into a form of censorship that had not been further analyzed.