An Analysis of China's "Great Cannon"

Bill Marczak
UC Berkeley, Citizen Lab

Nicholas Weaver
ICSI, UC Berkeley

Jakub Dalek
Citizen Lab

Roya Ensafi
Princeton University

David Fifield
UC Berkeley

Sarah McKune
Citizen Lab

Arn Rey

John Scott-Railton
Citizen Lab

Ron Deibert
Citizen Lab

Vern Paxson
ICSI, UC Berkeley

Abstract

On March 16th, 2015, the Chinese censorship apparatus employed a new tool, the "Great Cannon", to engineer a denial-of-service attack on GreatFire.org, an organization dedicated to resisting China's censorship. We present a technical analysis of the attack and what it reveals about the Great Cannon's working, underscoring that in essence it consitutes a selective nation-state Man-in-the-Middle attack tool. Although sharing some code similarities and network locations with the Great Firewall, the Great Cannon is a distinct tool, designed to compromise foreign visitors to Chinese sites. We identify the Great Cannon's operational behavior, localize it in the network topology, verify its distinctive side-channel, and attribute the system as likely operated by the Chinese government. We also discuss the substantial policy implications raised by its use, including the potential imposition on any user whose browser might visit (even inadvertently) a Chinese web site.

1 Introduction

On March 16, 2015, GreatFire.org observed that Amazon CloudFront services they rented to make blocked websites accessible in China were targeted by a distributed denial-of-service (DDoS) attack. On March 26, two GitHub pages run by GreatFire.org also came under the same type of attack. The attacker targeted services designed to circumvent Chinese censorship. A report released by GreatFire.org fingered malicious Javascript returned by Baidu servers as the source of the attack [25]. Baidu denied that their servers were compromised [42]. Several previous technical reports [5, 36, 41] suggested that the Great Firewall of China (GFW) orchestrated these attacks by injecting malicious Javascript into Baidu connections as they transited China's network border.

In this work we show that while the attack infrastructure was co-located with the GFW, the perpetrators carried out the attack using a separate offensive system, with different capabilities and design, that we term the "Great Cannon" (GC). The Great Cannon is not simply an extension of the Great Firewall, but a distinct attack tool that hijacks traffic to (or presumably from) individual IP addresses, and can arbitrarily replace unencrypted content as a man-in-the-middle.

The operational deployment of the GC represents a significant escalation in state-level information control: the normalization of the widespread use of an attack tool to enforce censorship by weaponizing users. Specifically, the GC manipulates the traffic of "bystander" systems outside China, silently programming their browsers to create a massive DDoS attack. While in this case employed for a highly visible attack, the GC clearly has the capability for use in a manner similar to the NSA's QUANTUM system [48], affording China the opportunity to deliver exploits targeting any foreign computer that communicates with any China-based website not fully protected by HTTPS.

We begin in § 2 with a review of the Great Firewall's operation, and then describe how the GC operates in § 3. We analyze the history of the injections as observed at a medium-sized enterprise (§ 4) and the impact of the Great Cannon in the DoS attack on GreatFire.org in § 5. In § 6 we provide the technical basis for attributing the operation of the GC to the Chinese government, and discuss the consequent policy implications in § 7. We reflect on some plausible enhancements that the GC's operators could employ (§ 8) and offer final conclusions in § 9.

2 Review of The Great Firewall

In general, firewalls serve as *in-path* barriers between networks: all traffic between the networks must flow through the firewall. Thus the name is actually a misnomer for China's "Great Firewall", which instead operates as an *on-path* system. The GFW eavesdrops on traffic between China and the rest of the world (TAP in Figure 1), and terminates requests for banned content (for example, upon seeing a request for "http://www.google.com/?falun", regardless of the actual destination server) by injecting a series of forged TCP Reset (RST) packets that tell both the requester and the destination to stop communicating [15].

On-path systems have architectural advantages for censorship due to their less disruptive nature in the pres-

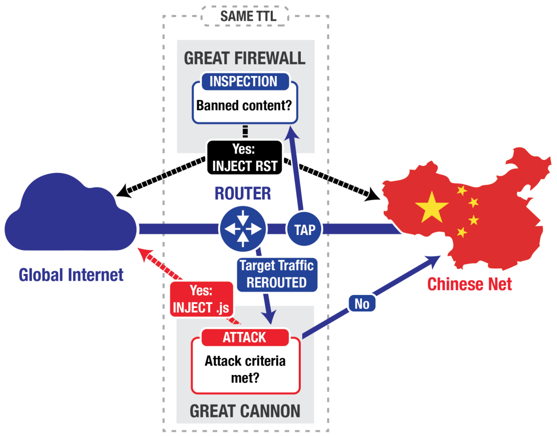

Figure 1: A simplified logical topology of the Great Cannon and Great Firewall

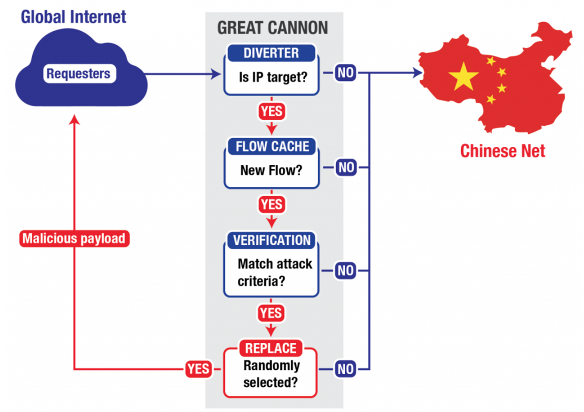

Figure 2: The Great Cannon's decision flow

ence of failure, but are less flexible and stealthy than in-path systems as attack tools, because while they can inject additional packets, they cannot prevent in-flight packets (packets already sent) from reaching their destination [49]. Thus, one generally can identify the presence of an on-path system interacting with active flows by observing anomalies resulting from the presence of both the injected and legitimate traffic.

The GFW keeps track of connections and reassembles their packets to determine if it should block traffic [29]. As opposed to considering each packet in isolation, this reassembly process requires additional computational resources, but facilitates better blocking accuracy. While a web request usually fits within a single packet, web replies may be split across several packets, and the GFW needs to reassemble these packets to understand whether a web reply contains banned content.

On any given physical link, the GFW runs its reassembly and censorship logic in multiple parallel processes [4] (perhaps running on a cluster of many different computers). Each process handles a subset of the link's connections, with all packets on a connection going to the same process. This load-balanced architecture reflects a common design decision when a physical link carries more traffic than a single computer can track [46]. Each GFW process also exhibits a highly distinctive side-channel—it progressively increments the IP TTL field on successive packets injected into the same connection [4].

3 The Great Cannon

The Great Cannon differs from the GFW: the GC operates as an in-path system, capable of not only injecting traffic but also directly suppressing traffic. Doing so enables it to act as a full "man-in-the-middle" (MITM) for targeted flows. We show that the GC does not actively examine all traffic on a link, but only intercepts traffic to (or presumably from) a set of targeted addresses. It is plausible that this reduction of the full traffic stream to just a (likely small) set of addresses significantly aids with enabling the system to keep up with the very high volume of traffic: the GC's full processing pipeline only has to operate on the much smaller stream of traffic to or from the targeted addresses. In addition, the GC only examines individual packets in determining whether to take action, avoiding the computational costs of TCP bytestream reassembly. The GC also maintains a flow cache of connections that it uses to ignore recent connections it has deemed no longer requiring examination.

We also show that the GC however shares several features with the GFW. Like the GFW, the GC also uses a multi-process design, with different source IP addresses handled by distinct processes. The packets injected by the GC have the same peculiar TTL side-channel as those injected by the GFW, suggesting that both the GFW and the GC likely share some common code. We find the GC deployed at the same network locations as the GFW. Taken together, these findings suggest that although the GC and GFW are independent systems with different functionality, there are significant structural relationships between the two.

In the attack on GitHub and GreatFire.org, the GC intercepted traffic sent to Baidu infrastructure servers that host commonly used analytics, social, or advertising scripts. If the GC saw a request for certain Javascript files on one of these servers, it appeared to probabilistically take one of two actions: it either passed the request on to Baidu's servers unmolested (roughly 98.25% of the time), or it dropped the request before it reached Baidu and instead sent a malicious script back to the requesting user (roughly 1.75% of the time). In this case, the requesting user is an individual outside China browsing a website making use of a Baidu infrastructure server (e.g., a website with ads served by Baidu's ad network).

The malicious script enlisted the requesting user as an unwitting participant in the DDoS attack against GreatFire.org and GitHub. Figure 2 shows our overall view of the GC's decision flow.

3.1 Evaluating the GC's Functionality

We began our investigation by confirming the continued normal operation of the GFW's censorship features.[1] We did so with measurements between our test system outside of China and a Baidu server that we observed returning the malicious Javascript. We sent the Baidu server a request that the GFW would process as a query for "http://www.google.com/?falun", a URL long known to trigger the GFW to inject forged TCP Resets to terminate the connection. We confirmed that the normal, well-understood operation of the GFW continued, including both the injected TCP Reset and, later, the legitimate response (an HTTP 403 reply) from the Baidu server.

We then localized where (with respect to our measurement system) in the network topology the GFW operated. For a given measurement packet, we vary the packet's TTL value in a traceroute-style progerssion. Doing so enabled us to isolate the hop responsible for the GFW's injector.

The GC appears to act probabilistically. Censorship systems generally operate in a deterministic fashion: they aim to block all content that matches the target criteria. The GC, on the other hand—at least for this particular attack—appears to act probabilistically, and ignores most of the traffic it could act on. In one test, it completely ignored all traffic from one of four measurement IP addresses, and for the three other measurement IP addresses it only responded to 526 requests out of 30,000 from the three (1.75%).

The GC operates as a separate, in-path system. As noted previously, our traces of GFW operation showed both the injected TCP Reset, as well as the legitimate server reply. In contrast, no legitimate server reply accompanied an injected malicious reply from the GC. We ran further testing, where we retransmitted our request to Baidu over the same connection, and with the same sequence numbers, after we received a malicious response. We observed Baidu responding as normal to the retransmitted request, treating it as new (previously unseen) data. Thus, we conclude that the GC must have dropped our request before it reached Baidu, a capability not present in the GFW.

We also checked whether the GFW and GC might share the same injector device, but found no evidence that they do. In particular, from a given TCP source port, we sent one request designed to trigger GC injection, followed by a request designed to trigger GFW injection. We repeated the experiment from a number of source ports. While packets injected by both the GFW and GC exhibited a similar (peculiar) TTL side-channel (previously reported in [4]) indicative of shared code between the two systems, we found no apparent correlation between the GFW and GC TTL values themselves.

The GC appears to be co-located with the GFW. We used the same TTL technique to localize the GC on the path between our test system and the Baidu server. We found that for our path, the GC acted on traffic between hop 17 and hop 18, the same link we observed as responsible for the GFW. We also observed that unlike the GFW, we could trigger the GC using "naked" requests (i.e., requests sent in isolation, with no previous TCP SYN as required for standard communication). Acting on "naked" requests implies that the GC's content analysis is more primitive (and easily manipulated), but does offer significant performance advantages, as the GC no longer needs to maintain complex state concerning connection status and TCP bytestream reassembly.[2]

We also checked two separate servers in China (115.239.210.141 and 123.125.65.120) whose traffic the GC targets to determine whether the GC existed alongside the GFW on multiple network paths. From our measurement system outside of China, we found that for 115.239.210.141, the GFW and the GC both existed between hop 12 and 13, a the link between 144.232.12.211 and 202.97.33.37 where the traffic enters China Telecom. For 123.125.65.120, both exist between hop 17 and 18, between 219.158.101.61 and 219.158.101.49 (China Unicom). A previous report by Robert Graham [41] used the same TTL technique to conclude that on one link, the GC was located "inside China Unicom infrastructure."

The GC was aimed only at specific destination IP addresses. When we probed an IP address adjacent to the Baidu server (123.125.65.121), the GC ignored the requests completely, although the GFW acted on censorable requests to this host.

The GC only acts on the first data packet of a connection. For a given source IP address and port, the GC only examines the first data packet sent when deciding whether to inject a reply. Avoiding examination of subsequent packets requires using a flow chace to remember which connections it has examined.

We confirmed these behaviors by sending a number of probes to the Baidu server, requesting resources that trigger the GC's injection. Each probe had a different

[1] Our preliminary report, available at https://citizenlab.org/2015/04/chinas-great-cannon/, includes links to packet captures generated in this investigation.

[2] Both the well-known airpwn [45] tool and NSA's QUANTUM system are similarly stateless in their decision making, reflecting that attack tools for injecting malicious content do not require robust reassembly.

source port. We sent 500 probes, each with the triggering request split across three packets (so 1,500 packets total). The GC ignored each probe. We then sent 500 probes where the first packet's data is an invalid HTTP request, and the second packet's data is a complete, triggering request. The GC ignored each probe. Finally, we sent single-packet probes, each containing a complete triggering request; the GC acted on these in some cases, reflecting its probabilistic decision-making process.

How big is the GC flow cache? We attempted to completely fill the GC flow cache by sending packets to the Baidu server with different ports, while probing to see whether the entries that we previously added had now expired. Our attempt suggests that for this test the GC flow cache between our test system and the Baidu server supports up to around 16,000 entries for a single sending IP address.

The flow cache capacity test also provides evidence that the GC's probabilistic choice occurs on the decision to act on a particular flow. When we succeeded in completely filling the flow cache, subsequently injected packets occurred for different source ports than in the initial test. If the GC only intercepted a subset of flows to the target IP address, we would expect subsequent injections to appear for the same flows, since most schemes to probabilistically select flows for interception (such as hashing the connection 4-tuple) would select the same set of flows the second time around.

The GC have a load-balanced architecture. By sending from multiple addresses simultaneously, we determined that the GC uses a separate flow cache for different source IP addresses, and that packets injected from different source IP addresses have distinct TTL side-channels. This finding suggests a load-balanced architecture similar to the GFW, where packets are routed to GC nodes based on source IP address. We then sent traffic alternating from four measurement IP addresses in an attempt to fill a 16,000 entry cache. This attempt did not manage to fill the cache, suggesting that the GC not only processed the different source IP addresses with different injection elements, but also did so using different space in its flow cache or using distinct flow caches. As stated before, one of the four source IP addresses never received any injected replies.

Potential OpenFlow implementation. In considering the GC's need to operate at very high speeds, we note that it might employ OpenFlow functionality for efficient operation [38]. To do so it would use a default rule that diverts all payload-containing packets destined to the target to the OpenFlow controller, which examines each packet to determine whether to inject an attack, forwarding the packet onward in the case of a negative decision. The controller would then insert a rule into the flow cache instructing the router to pass all further packets directly on to the destination.

4 Assessing the GC's History of Use

Provos reported that Google's Safe Browsing project captured instances of the attack between March 3rd and April 7th, including an HTTP variant that wrapped the real page in an iframe with an additional "do not attack" `t=` parameter appended to the URL [37]. The addition to the URL will cause the browser to reload the target page, rendering it in the iframe so the user does not notice any interception. We've observed multiple instances of the `t=` flag in our own data, each with a different value. This leads us to believe that the flag is just a nonce, causing the browser to reload the page, hoping that the cannon won't reattack the newly issued page load. Safe Browsing also spotted attack tests between March 3rd and March 13th before the GC's operators launched their attack against GreatFire.org's CloudFront instances.

We built a small analyzer to process pcap files, searching archival packet captures for connections that exhibit the GC's distinctive fingerprint of both an increasing TTL and an increasing window size over a sequence of three consecutive packets. Using this detector, we examined 8 months of traces recorded at the network border of the Lawrence Berkeley National Laboratory. Although sequences of three packets with increasing TTLs in practice occur fairly often, the combination of these plus the incrementing window size proved highly reliable, with only one source in the traces producing false positives.

Aside from the false positives, all of the activity found in the traces related to the GreatFire.org DoS attack. The payloads varied only in minor ways, including later use of `/packer/` [19] to pack the Javascript. The detections revealed somewhat broader targeting than Baidu: we also observed on April 1 a single injection for a retrieval from `l.qq.com`, a non-Baidu property that belongs to Tencent QQ. We also observed evidence suggesting manual entry of the DoS targets: one of the injections directed the user's browser to attempted to attack `d2yeolxorqum8y.cloudfront.net`, a typo of the actual domain `d2ye0lxorqum8y`. (The attacker substituted the letter O for the number 0.)

We further analyzed the LBNL data to assess whether the DoS attack had a measureable transport-layer effect on the associated CloudFront infrastructure. We took all of the `cloudfront.net` domain names present in the injections our detector located (12 total), correlated these with the site's logs of outbound DNS queries in order to determine the associated IP addresses (1,088 total, from 34 /24s and 6 /16s) and then looked in the site's connection logs to assess the fate of any connections made to those IP addresses. Here, "fate" corresponds to whether the connection (1) went unanswered (no SYN ACK seen), (2) failed to complete (either an

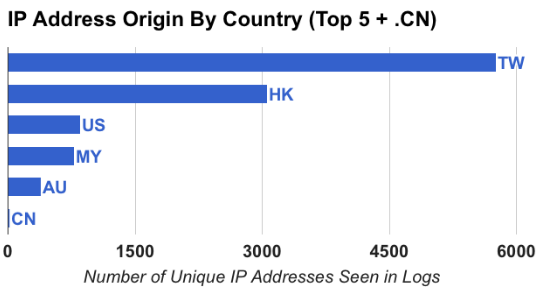

Figure 3: The distribution of countries/regions participating in the DOS attack.

initial RST or a subsequent RST seen), or (3) successfully completed with a FIN handshake.

On any given day during the course of the attack, LBNL systems attempted tens to hundreds of thousands of connections to corresponding CloudFront systems. Virtually all of these were answered (always more than 99%). On most days 2–5% of the connections failed to complete, which is also the rate we generally observed for nearby days outside the time of the attack. On three days 8–10% failed to complete, however these levels (and higher) were also occasionally observed for days outside the attack period. These results suggest that CloudFront's infrastructure suffered no network-layer disruption and at most very modest server-level disruption.

Finally, a colleague[3] contacted us to discuss their analysis of extensive logs dating back two years looking for earlier GC activity. The logs revealed multiple instances of its employment to engineer DoS attacks predating the attack on GreatFire.Org, with the earliest instance occurring a full year earlier.

5 Analysis the GreatFire.org's Logs

The staff of GreatFire.org provided us with server logs covering the period of March 18 to 28. (A report previously published by GreatFire.org uses a different sample [25].) This period captures the end of the attack.

To keep our analysis manageable, we examined a sample of the data from March 18th 11:00 GMT to March 19th 7:00 GMT, as seen from two of the three most commonly used backend servers. For each hour, we selected a random subset of approximately 30MB of compressed logs for each server. The total sample includes 16.6M web requests, with 13K unique source IP addresses. We used the MaxMind GeoIP2 Lite database [32] from March 3rd, 2015, to assign a country of origin to each source IP address. For any IP address that did not result in a definite geolocation using this tool (31 addresses), we looked up the address manually using the iplocation.net service.

67% of the IP addresses originated from Taiwan and Hong Kong, two regions where Chinese is the official language. China, however, accounted for only 18 requests. This finding aligns with our understanding of the malicious code injection occurring at the border of the Chinese Internet. Figure 3 shows a distribution of the most prominent countries.

To determine which websites had their responses altered by the injection, we extracted the domain names of the 25 most frequently seen referrers in our dataset, finding that these domains account for 55% of the total requests in the sample.

The most commonly seen domain (38% of requests) is `pos.baidu.com`, a part of Baidu's ad network. Many non-Baidu sites display ads served through Baidu's ad network, indicating that visitors to non-Baidu sites displaying ads also became targeted: due to inter-iframe isolation, advertisements set the Referrer of the advertisement network, not the final site.

We examined the remaining top 25 domains, and could link each one to Baidu: in each case, the site is either a Baidu property or uses Baidu analytics, advertisements, or static resources. This finding indicates that Baidu was a major injection target for this attack. According to Alexa statistics, Baidu itself is the fourth-most visited site globally, the highest ranking China-based site on the global list [2], and has received an estimated 4.99 million unique visitors from the US alone in the past 30 days.

We speculate that Baidu was chosen as an injection target simply because it is an effective way to target many users outside of China.

6 Attributing the Great Cannon

We find compelling evidence that the Chinese government operates the GC. In recent public statements, China deflected questions regarding whether they were behind the attack, instead emphasizing that China often finds itself a victim of cyber attacks [12]. A subsequent Chinese news article, containing an explicit denial and a denouncement of our initial public report as false [22], was itself later censored within China [13].

Where does the GC operate? We tested two international Internet links into China belonging to two different Chinese ISPs, and found that in both cases the GC was co-located with the GFW. This co-location across different ISPs strongly suggests a governmental actor.

Who built the Great Cannon? That the GFW and GC have the same type of TTL side-channel suggests that they share some source code. We are unaware of any public software library for crafting packets that introduces this type of TTL side-channel.

[3] Who requests to remain anonymous.

What is the Great Cannon's role? Our observations indicate that the GC's design does not reflect technology well-suited for performing traffic censorship. Its operation only examines the first data packet of a given connection and it only examines traffic with targeted IP addresses, which provides a weak censorship mechanism compared to the GFW. More generally, the GC's design does not, in practice, enable it to censor any traffic not already censorable by the GFW.[4] Thus, the evidence indicates that the GC's role is to inject traffic under specific targeted circumstances, not to censor traffic.

Who was the Great Cannon attacking? The DDoS attack launched by the GC using "bystander" machines directly aligns with known political concerns of the Chinese government. The Cyberspace Administration of China has previously referred to GreatFire.org as a "foreign anti-Chinese organization" (境外反华组织) [16]. The particular GreatFire.org service targeted in this attack provides proxies to bypass the GFW using encrypted connections via Amazon's CloudFront cloud service.

GreatFire.org also uses two GitHub repositories that provide technology for users who wish to circumvent Chinese government censorship [24]. The attack on GitHub specifically targeted these repositories, possibly in an attempt to compel GitHub to remove these resources. GitHub encrypts all traffic using TLS, preventing a censor from only blocking access to specific GitHub pages. In the past, China attempted to block GitHub, but the block was lifted within two days following significant negative reaction from local programmers [34].

7 Policy Contexts and Implications

Deploying the Great Cannon is a major shift in tactics, and has a highly visible impact. It is likely that this attack, with its potential for political backlash,[5] would require the approval of high-level authorities within the Chinese government. These authorities may include the State Internet Information Office (SIIO),[6] which is responsible for Internet censorship. It is also possible that the top body for cybersecurity coordination in China, the Cybersecurity and Informatization Leading Group (CILG) [31, 1], would have been involved. The CILG is chaired by Xi Jinping, the General Secretary of the Communist Party's Central Committee and President of the People's Republic of China, and includes as members senior leaders from across the government; its administrative office is housed within the SIIO [31].

The government's reasoning for deploying the GC here is unclear, but it may wish to confront the threat presented to the Communist Party of China's (CPC) ideological control by the "collateral freedom" strategy advanced by GreatFire.org and others. The attack was exceptionally costly to GreatFire.org in terms of potential monetary costs according to their public statements [11]. This also disrupted GitHub which hosted GreatFire.org's repositories (although it did not appear to disrupt Amazon Cloudfront). Such a disruption could be both an attempt to block the operations of an undesirable resource, and a signal sent to other organizations of the potential price tag for this kind of activity. Deployment of the GC may also reflect a desire to counter what the Chinese government perceives as US hegemony in cyberspace.

This approach would be consistent with the CPC's paramount focus on protecting "domestic stability" (and its own authority) against entities it has identified as "foreign hostile forces," including not only governments but also Western media outlets (such as the New York Times) and NGOs or other civil society actors (such as GreatFire.org) [33, 47, 10, 14]. According to such a world view, the collateral freedom strategy is a provocative, hostile act that threatens China's security.

7.1 Implications of Using Traffic to Baidu

The incorporation of Baidu in this attack suggests that the Chinese authorities are willing to pursue domestic stability and security aims at the expense of other goals, including fostering economic growth in the tech sector. Selecting Baidu's international traffic may appear counterproductive given the importance of Baidu to the Chinese economy: the company enjoys stature as one of China's "big three" Internet firms, alongside Alibaba and Tencent [40], and currently ranks as the top site in China [3]. While its shares came under pressure after the February release of its Q4 and fiscal year 2014 reports [52], its total revenue in 2014 was USD $7.906 billion, with online marketing revenues for that period valued at USD $7.816 billion [7].

Baidu has denied involvement in the attack and asserted its internal security was not compromised. Yet the targeting of international visitors trying to reach sites that are Baidu properties, or that use Baidu analytics, advertisements, or static resources, could undermine the company's reputation and its appeal to overseas users and advertisers, although the actual loss of revenue in this particular attack was almost certainly negligible.

Baidu writes in its SEC filings that it was the target

[4] The only exception would be an attempt to censor TLS traffic using forged certificates when it is infeasible to block the entire censored site based on the TLS Server Name Indication field. However, should such censorship be detected, this would likely result in the termination of the forging root within many browsers [50, 17].

[5] Particularly after the Snowden disclosures, and the public / state outcry associated with the NSA's QUANTUM system and other programs, the Chinese government would presumably be aware of the significant international political ramifications of a decision to use the GC to target overseas entities, and escalate the matter accordingly.

[6] Also referred to as the "Cyberspace Administration of China" [51].

of legal action in the United States in 2011 [8] for complying with Chinese censorship. Baidu explicitly notes that cooperation and coordination with Chinese censorship authorities could be costly in terms of brand image, profit, and stockholder confidence: "our compliance with PRC regulations governing Internet access and distribution of news and other information over the Internet may subject us to negative publicity or even legal actions outside of China." [8]

Moreover, exploiting Baidu's international reach as a means for conducting digital attacks belies the government's recent commitment to enhance the global presence of Internet companies. At the meeting of the National People's Congress on March 5, 2015, Premier Li Keqiang (who is also Vice-Chair of the CILG) announced:

> We will develop the "Internet Plus" action plan to integrate the mobile Internet, cloud computing, big data, and the Internet of Things with modern manufacturing, to encourage the healthy development of e-commerce, industrial networks, and Internet banking, and to guide Internet-based companies to increase their presence in the international market. [44]

This goal—which closely echoes that contained in a draft declaration presented (but not passed) at the November 2014 Wuzhen World Internet Conference [39, 43, 20]—may not come to fruition if Chinese domestic companies appear unreliable, their business objectives secondary to other objectives of the Chinese Government.

Chinese authorities may, however, be betting that their use of Baidu traffic to mount this DDoS attack will ultimately be perceived as an isolated occurrence, a sort of "force majeure," with limited impact on Baidu's long-term economic prospects—particularly given Baidu's apparent status as unwitting victim and its strong market position.

Additionally, Baidu's CEO Robin Li is a member of the Chinese People's Political Consultative Conference [9] and well-positioned for lucrative government contracts going forward—such as his artificial intelligence project "China Brain," for which he has sought military support [26]. He may have little personal incentive (let alone opportunity, given the existing legal and regulatory framework applicable to Internet companies in China [28]) to challenge this action by the government.

7.2 Authorities Possibly Responsible

Even for the GFW, it is difficult to pinpoint the precise authorities behind its deployment, or its operators and origins. This makes understanding the origins of the GC equally challenging. However, some clues are available. For example, the shared source code and co-location between the GFW and GC suggest that the GC could have been developed within the same institutional framework as the GFW. We might therefore draw further insight into the GC by assessing what we know about the GFW.

Some reports characterize the GFW as an element of China's "Golden Shield" project [27] under the authority of the Ministry of Public Security. However, unverified insider information 'leaked' online suggests that the GFW was developed within a separate entity: the "National Computer Network and Information Security Management Center" (国家计算机网络与信息安全管理中心) (hereafter, "the Center") [21, 18, 6]. Little is publicly known about the Center. It appears to bear close relationship [21] to the National Computer Network Emergency Response Technical Team/Coordination Center of China (CNCERT/CC) run by the Ministry of Industry and Information Technology (MIIT)—indeed, the listed address for CNCERT/CC is the same as that of the Center as indicated on its patent applications—and the former National Network and Information Security Coordination Team,[7] a subcommittee of the State Informatization Leading Group subsumed by the CILG in 2014 [31]. Notably, "MIIT also regulates China's six Internet service providers (ISPs), which in turn are expected to monitor and filter content on their networks according to censorship guidelines established by the State Council Information Office and the SIIO" [31]. Those ISPs include China Telecom and China Unicom, on whose links we co-located the GFW and GC.

It is unknown whether the GFW and/or the GC are in fact maintained (or may have been developed in whole or in part) by the Center. However, patent applications filed by this entity, taken together, appear to indicate a mandate for large scale network surveillance, filtering, and defense. The Center has filed nearly 100 patent applications,[8] for areas including topic and dialect recognition, website-classification, and network monitoring. Moreover, according to state media, during the time of the GFW's development the so-called "father of the Great Firewall," Fang Binxing, was employed at CNCERT/CC [23], an entity that appears closely tied to

[7] This entity was also known in English as the State Network and Information Security Coordination Group. Its responsibilities included: "researching and enacting strategy and policy of national information security safeguard[;] organizing and coordinating related departments of government to protect critical information infrastructure[;] mobilizing and directing computer emergency response[;] improving information sharing and notification." [30]

[8] It is important to note that patent applications do not necessarily reflect current capacities or actual deployment of a technology. They do, however, provide insight into the designs, focus, and goals of the filing entities.

the Center.[9] Fang is likewise listed as an inventor on a 2008 patent application by the Center, indicating some collaboration with the Center prior to that point.

While we cannot determine the exact role played by the Center, the patent documentation and the Center itself require further research and analysis to determine whether they are relevant to operation of the GC, or present other human rights-related concerns.

8 Potential Enhancements

A technically simple change in the GC's configuration—switching to operating on traffic *from* a specific IP address rather than *to* a specific address—would enable its operator to deliver malware to targeted individuals who communicates with any Chinese server not employing cryptographic protections. The GC operator first needs to discover the target's IP address and identify a suitable exploit. The operator then tasks the GC to intercept traffic from the target's IP address, and replace certain responses with malicious content. If the target ever made a single request to a server inside China not employing encryption (e.g., Baidu's ad network), the GC could deliver a malicious payload to the target. A target might not necessarily realize that their computer was communicating with a Chinese server, as a non-Chinese website located outside China could (for example) serve ads ultimately sourced from Chinese servers.

Although China can launch such attacks with the GFW, the MITM nature of the GC does offer a potential stealth advantage. The GC's current distinctive side-channel is easily recognized, but its operators could easily correct this implementation artifact.

Since the GC operates as a full MITM, it would also be straightforward to have it intercept unencrypted email to or from a target IP address and undetectably replace any legitimate attachments with malicious payloads, manipulating email sent from China to outside destinations. Even email transmission protected by standard channel confidentiality (STARTTLS) can be undermined because the GC is in a position to launch a downgrade attack, steering the transmission to only use legacy, unencrypted communication. This could enable "perfect spearphishing": the substitution for an existing legitimate attachment with a malicious one, with no plausible way for the recipient to distinguish the alteration as suspicious. However, unlike injecting malicious scripts, this attack will usually require a fully stateful MITM proxy to execute.

9 Concluding Remarks

The attack launched by the Great Cannon appears relatively obvious and coarse: a denial-of-service attack on services objectionable to the Chinese government. Yet the attack itself underscores a far more significant capability: an ability to "exploit by IP address". This possibility, not yet observed but a feature of its architecture, represents a potent cyberattack capability.

Our findings in China add another documented case to at least two other known instances of governments tampering with unencrypted Internet traffic to control information or launch attacks—the other two being the use of QUANTUM by the US NSA and UK's GCHQ.[10] In addition, product literature from two companies, FinFisher and Hacking Team, indicate that they sell similar "attack from the Internet" tools to governments around the world [35]. These latest findings emphasize the urgency of replacing legacy web protocols like HTTP with their cryptographically strong counterparts, such as HTTPS.

We remain puzzled as to why the GC's operator chose to employ its capabilities in a highly visible fashion. Conducting such a widespread attack clearly demonstrates the weaponization of the Chinese Internet to co-opt arbitrary computers across the web and outside of China to achieve China's policy ends. The repurposing of the devices of unwitting users in foreign jurisdictions for covert attacks in the interests of one country's national priorities is a dangerous precedent—contrary to international norms and in violation of widespread domestic laws prohibiting the unauthorized use of computing and networked systems.

10 Acknowledgments

Special thanks to: Adam Senft (Citizen Lab) and Paul Pearce (UC Berkeley, ICSI). We wish to acknowledge GreatFire for making their logs available to us for analysis. We also express our deep gratitude to several individuals, anonymous or pseudonymous at their request, including "Jack B," who aided us in understanding elements of the attack in the early days, and others who helped us formulate and develop this report.

This work is sponsored in part by The John D. and Catherine T. MacArthur Foundation, the Open Technology Fund's Information Controls Fellowship Program, and the National Science Foundation grants NSF-1237265, NSF-1213157, and NSF-1223717. All opinions are those of the authors, not the funding institutions.

[9] See [53] for an official diagram mapping the relationship between CNCERT/CC, MIIT, and the National Network and Information Security Coordination Team.

[10] Strictly speaking, QUANTUM is a "man-on-the-side" exploitation tool similar in architecture to the GFW rather than the GC, although its primary use is to compromise computers with attacks directed from the Internet.

References

[1] Adam Segal. China's New Small Leading Group on Cybersecurity and Internet Management. Council on Foregn Relations: Asia Unbound, http://blogs.cfr.org/asia/2014/02/27/chinas-new-small-leading-group-on-cybersecurity-and-internet-management/.

[2] Alexa Site Overview for baidu.com. http://www.alexa.com/siteinfo/baidu.com.

[3] Alexa. Top Sites in China. http://www.alexa.com/topsites/countries/CN.

[4] Anonymous. Towards a Comprehensive Picture of the Great Firewall's DNS Censorship. In *4th USENIX Workshop on Free and Open Communications on the Internet (FOCI 14)*, San Diego, CA, August 2014. USENIX Association.

[5] Anthr@x. Baidu's Traffic Hijacked to DDoS GitHub. Insight Labs, http://insight-labs.org/?p=1682.

[6] Australian Centre on China in the World. Fang Binxing and the Great Firewall. The China Story, https://www.thechinastory.org/yearbooks/yearbook-2013/chapter-6-chinas-internet-a-civilising-process/fang-binxing-and-the-great-firewall/.

[7] Baidu Announces Fourth Quarter and Fiscal Year 2014 Results. PR Newswire, http://www.prnewswire.com/news-releases/baidu-announces-fourth-quarter-and-fiscal-year-2014-results-300034622.html.

[8] Baidu Inc. SEC Form 20-F, FY 2011. http://www.sec.gov/Archives/edgar/data/1329099/000119312512139789/d243699d20f.htm.

[9] Baidu Founder Li and Politburo's Yu Join Top China Advisory Body. Bloomberg, http://www.bloomberg.com/news/articles/2013-02-02/baidu-chief-li-politburo-s-yu-join-china-s-top-advisory-body.

[10] Central Committee of the Communist Party of China's General Office. Communiqué on the Current State of the Ideological Sphere. English translation by ChinaFile, https://www.chinafile.com/document-9-chinafile-translation.

[11] Charlie. We are Under Attack. Greatfire.org, https://en.greatfire.org/blog/2015/mar/we-are-under-attack.

[12] Ministry of Foreign Affairs of the People's Republic of China. Foreign Ministry Spokesperson Hua Chunying's Regular Press Conference, http://www.fmprc.gov.cn/mfa_eng/xwfw_665399/s2510_665401/2511_665403/t1250354.shtml, March 30, 2015.

[13] China Digital Times. Minitrue: Cease Fire on "Great Cannon". http://chinadigitaltimes.net/2015/04/minitrue-do-not-republish-great-cannon-report/.

[14] Chris Buckley. China's New Leadership Takes Hard Line in Secret Memo. New York Times, http://cn.nytimes.com/china/20130820/c20document/dual/.

[15] Jedidiah R Crandall, Daniel Zinn, Michael Byrd, Earl T Barr, and Rich East. ConceptDoppler: a weather tracker for Internet censorship. In *ACM Conference on Computer and Communications Security*, pages 352–365, 2007.

[16] Cyberspace Administration of China. 国家网信办发言人："Outlook受中国攻击"的说法纯属污蔑. http://www.cac.gov.cn/2015-01/22/c_1114097853.htm.

[17] Dan Goodin. Google Chrome will Banish Chinese Certificate Authority for Breach of Trust. http://arstechnica.com/security/2015/04/google-chrome-will-banish-chinese-certificate-authority-for-breach-of-trust/.

[18] Daniel Anderson. Splinternet Behind the Great Firewall of China. *ACM Queue*. http://queue.acm.org/detail.cfm?id=2405036.

[19] Dean Edwards. /packer/. http://dean.edwards.name/packer/.

[20] Franz-Stefan Gady. The Wuzhen Summit and Chinese Internet Sovereignty. Huffington Post, http://www.huffingtonpost.com/franzstefan-gady/the-wuzhen-summit-and-chi_b_6287040.html.

[21] GFW的前世今生，一部GFW之父方滨兴的发家史 ["GFW Past and Present, Family History of the Father of the GFW Fang Binxing"]. https://fangbinxing.appspot.com/2010/08/10/fangbingxing.html.

[22] Global Times. "Foreign Media Grabs Chance to Hype China's 'Great Cannon'; May be American Effort to Shift Blame", Mirrored by China Digital Times: http://chinadigitaltimes.net/chinese/2015/04/【真理部】外媒借机炒中国启用网络大炮-或是被美/.

[23] Great Firewall father speaks out. Global Times, http://english.sina.com/china/p/2011/0217/360411.html.

[24] GreatFire Github Repositories: https://gitub.com/greatfire and https://github.com/cn-nytimes.

[25] GreatFire. Using Baidu 百度 to steer millions of computers to launch denial of service attacks. https://drive.google.com/file/d/0ByrxblDXR_yqeUNZYU5WcjFCbXM/view?pli=1.

[26] Hsu Chang-ping. Baidu welcomes China's military to join China Brain project on AI systems. WantChinaTimes, http://www.wantchinatimes.com/news-subclass-cnt.aspx?id=20150307000015&cid=1101.

[27] Immigration and Refugee Board of Canada. China: The Public Security Bureau (PSB) Golden Shield Project, including implementation and effectiveness; Policenet, including areas of operation; level and effectiveness of information sharing by the authorities. CHN104762.E, http://www.refworld.org/docid/543ba3824.html.

[28] Jedidiah R. Crandall and Masashi Crete–Nishihata and Jeffrey Knockel and Sarah McKune and Adam Senft and Diana Tseng and Greg Wiseman. China Chats: Tracking Surveillance and Censorship in TOM-Skype and Sina UC. First Monday 18, no 7, http://firstmonday.org/ojs/index.php/fm/article/view/4628/3727, 2013.

[29] Sheharbano Khattak, Mobin Javed, Philip D Anderson, and Vern Paxson. Towards illuminating a censorship monitor's model to facilitate evasion. In *3rd USENIX Workshop on Free and Open Communications on the Internet*, page 2, 2013.

[30] Li Jingjing. Trends and Tactics in Cyber-Terrorism. Information Security Supervision Bureau, Ministry of Public Security, p. 13, http://www.asean.org/archive/arf/13ARF/2nd-Cyber-Terrorism/Doc-7.PDF.

[31] Jon Lindsay. *China and Cybersecurity: Espionage, Strategy and Politics in the Digital Domain*, chapter Introduction-China and Cybersecurity: Controversy and Context. Oxford University Press, 2015.

[32] MaxMind GeoLite2. https://dev.maxmind.com/geoip/geoip2/geolite2/.

[33] Sarah McKune. *China and Cybersecurity: Espionage, Strategy and Politics in the Digital Domain*, chapter 'Foreign Hostile Forces': The Human Rights Dimension of China's Cyber Campaigns. Oxford University Press, 2015.

[34] Michael Kan. GitHub unblocked in China after former Google head slams its censorship. http://www.computerworld.com/article/2493478/internet/github-unblocked-in-china-after-former-google-head-slams-its-censorship.html.

[35] Morgan Marquis-Boire. Schrodinger's Cat Video and the Death of Clear-Text. CitizenLab, https://citizenlab.org/2014/08/cat-video-and-the-death-of-clear-text/.

[36] Netresec. China's Man-on-the-Side Attack on GitHub. http://www.netresec.com/?month=2015-03&page=blog&post=china%27s-man-on-the-side-attack-on-github.

[37] Niels Provos. A JavaScript DDoS Attack as seen by Safe Browsing. Google Online Security Blog, http://googleonlinesecurity.blogspot.com/2015/04/a-javascript-based-ddos-attack-as-seen.html.

[38] Open Networking Foundation. https://www.opennetworking.org/sdn-resources/openflow/57-sdn-resources/onf-specifications/openflow?layout=blog.

[39] Organizing Committee of the World Internet Conference. Wuzhen Declaration. http://www.scribd.com/doc/247566581/World-Internet-Conference-Draft-Declaration.

[40] Shuli Ren. China Internet: Alibaba, Tencent, Baidu To Continue Buying Spree, R&D. Barron's Asia, http://blogs.barrons.com/asiastocks/2015/01/07/china-internet-alibaba-tencent-baidu-to-continue-buying-spree-rd/.

[41] Robert Graham. Pinpointing China's Attack on GitHub. Errata Security, http://blog.erratasec.com/2015/04/pin-pointing-chinas-attack-against.html.

[42] Russell Brandom. Last night, GitHub was hit with a massive denial-of-service attack from China. http://www.theverge.com/2015/3/27/8299555/github-china-ddos-censorship-great-firewall, March 2015.

[43] Catherine Shu. China Tried To Get World Internet Conference Attendees To Ratify This Ridiculous Draft Declaration. TechCrunch, http://techcrunch.com/2014/11/20/worldinternetconference-declaration/.

[44] State Council of the People's Republic of China. Report on the Work of the Government (2015), delivered by Li Keqiang, Premier

of the State Council, to the National People's Congress. http://english.gov.cn/archive/publications/2015/03/05/content_281475066179954.htm.

[45] toast@sourceforge,net. Airpwn. SourceForge, http://airpwn.sourceforge.net/Airpwn.html.

[46] Matthias Vallentin, Robin Sommer, Jason Lee, Craig Leres, Vern Paxson, and Brian Tierney. The NIDS cluster: Scalable, stateful network intrusion detection on commodity hardware. In *Recent Advances in Intrusion Detection*, pages 107–126. Springer, 2007.

[47] Wang Chen. Concerning the development and administration of our country's internet. Translated by Human Rights in China, China Richts Forum: "China's Internet": Staking Digital Ground no 2, http://www.hrichina.org/en/content/3241.

[48] Nicholas Weaver. Our Government Has Weaponized the Internet. Wired Opinion, http://www.wired.com/2013/11/this-is-how-the-internet-backbone-has-been-turned-into-a-weapon/.

[49] Nicholas Weaver, Robin Sommer, and Vern Paxson. Detecting Forged TCP Reset Packets. In *Networking and Distributed Security Symposium*, 2009.

[50] Wikipedia: DigiNotar. https://en.wikipedia.org/wiki/DigiNotar.

[51] Xinhua. China sets up State Internet information office. http://www.chinadaily.com.cn/china/2011-05/04/content_12440782.htm.

[52] Doug Young. Investors Burn Out On Baidu Mobile Story. Forbes Asia, http://www.forbes.com/sites/dougyoung/2015/02/12/investors-burn-out-on-baidu-mobile-story/.

[53] Zhou Yonglin. Introduction on Chinese Network Emergency Response System & CNCERT/CC's Activities. http://unpan1.un.org/intradoc/groups/public/documents/apcity/unpan016388.pdf.