

acmqueue Splinternet Behind the Great Firewall of China

Once China opened its door to the world, it could not close it again

Daniel Anderson

What if you could not access YouTube, Facebook, Twitter, and Wikipedia? How would you feel if Google informed you that your connection had been reset during a search? What if Gmail was only periodically available, and Google Docs, which was used to compose this article, was completely unreachable? What a mess!

These things happen almost every day in China. If you are a foreign visitor to China, you could experience what NBA player J.R. Smith encountered: “Dear China, The fact that u won’t let me work my Skype on my desktop or Twitter is really pissing me off.”²⁰ As software developer Tony Hunt said, “That was really the most frustrating thing, as I never knew whether the connection had just dropped or if the site was being censored.”¹⁵

Most of these problems are caused by GFW (Great Firewall of China, also known as GFC), one of the most important building blocks in China’s comprehensive censorship system, and perhaps the most sophisticated Internet censorship system in the world.¹² The Chinese government can remove the “harmful information” or even punish its authors inside China. For information hosted outside China, however, the Chinese authorities can do nothing but block or filter access. Without censorship at the international gateway of the Internet, the traditional censorship systems are utterly worthless. This is why GFW is so critical for the whole system, and so important to the stability of the Chinese state.

On the other hand, economic development is equally critical (if not more so) to the stability of the Chinese government. It cannot cut off the physical link to the Internet completely lest it block business traffic (such as HTTPS or VPN). With the development of the Chinese economy, China has more network users than any other country, with a penetration rate of 38.3 percent.⁵ A censorship system that is too severe could endanger the stability of the government.

Essentially, GFW is a government-controlled attacking system, launching attacks that interfere with legitimate communications and affecting many more victims than malicious actors. Using special techniques, it successfully blocks the majority of Chinese Internet users from accessing most of the Web sites or information that the government doesn’t like. GFW is not perfect, however. Some Chinese technical professionals can bypass it with a variety of methods and/or tools. An arms race between censorship and circumvention has been going on for years, and GFW has caused collateral damage along the way.

The victims of this aggressive censorship system include not only Chinese users, but also those in other countries who have nothing to do with China. GFW threatens not only freedom of speech and the free flow of information, but also the global economy. For example, Patrick Chovanec, professor at Tsinghua University’s School of Economics and Management in Beijing, asserts that Chinese censorship of Google, Facebook, and Twitter provides business advantages to their Chinese competitors, thus also serving a role of economic protectionism.⁷

A BRIEF HISTORY

Because of Wikipedia, most people once believed that GFW was part of the Golden Shield Project, operated by the MPS (Ministry of Public Security) in China.²⁵ A widely read article appeared in 2010, however, that more accurately states that the Golden Shield Project originally aimed only to build an intranet for the police in China and had nothing to do with censoring the public Internet.¹⁴ This article seems to have been leaked by some people who participated in the construction of GFW, as it provides a detailed timeline of its development. According to the article, GFW was built in 1999 and is operated by CNCERT/CC (National Computer Network Emergency Response Technical Team/Coordination Center of China), controlled by MIIT (Ministry of Industry and Information Technology). Because most ISPs are state-owned and also controlled by MIIT, CNCERT/CC can deploy the censorship systems in the backbones of all ISPs without any obstacles.

Binxing Fang, now the president of Beijing University of Posts and Telecommunications, is the principal designer and so-called “father of GFW.” He accepted the title in an article published by a state-controlled news outlet.¹⁸ In fact, many top universities, institutes, and IT companies have contributed to the building and/or operation of GFW. As long as the project can bring in money, many professors and technical professionals in China are willing to accept it. Some are even proud of it as they believe the wall benefits the country.

CENSORSHIP TECHNOLOGIES

Over a decade of development, GFW has been deployed near the gateways of all Chinese domestic ISPs. With DPI (deep packet inspection) technology, GFW wiretaps all international links and inspects the traffic to detect any sensitive keywords going through the gateway. GFW depends mainly on three technologies to block “harmful” information: IP blocking, DNS (Domain Name System) injection, and TCP RST (Reset).

IP ADDRESS BLOCKING

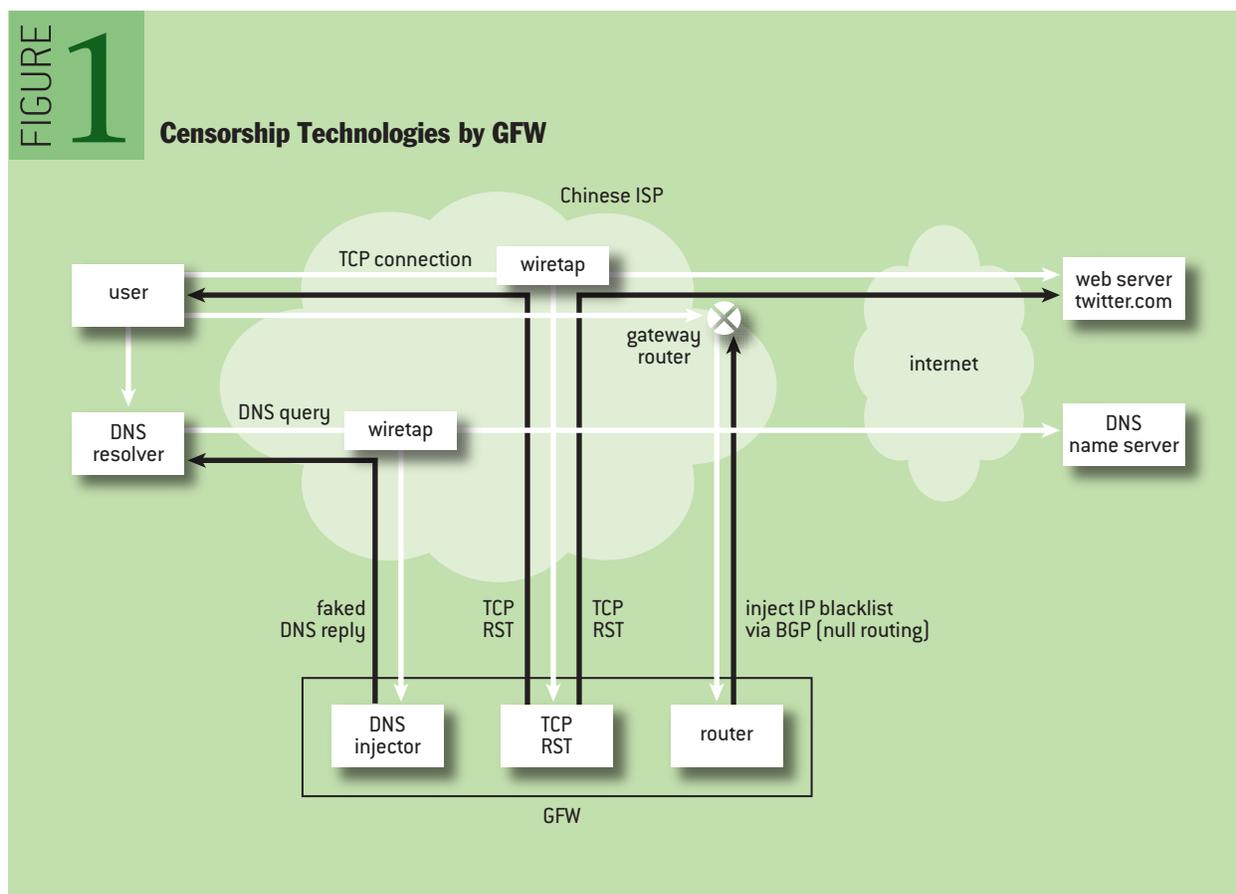
According to a paper published by its designers, GFW relies on null routing (see figure 1) to block IP packets by blacklisting destination addresses.¹⁶ By peering with the gateway routers of all Chinese ISPs, GFW injects routing information into BGP (Border Gateway Protocol) and hijacks all traffic to blocked websites—such as twitter.com. Although null routing can block only the outbound traffic from China and permits inbound traffic, it is enough to block a Web site because most current Internet communication can be established only with two-way interaction.

This is a lightweight censorship solution: the government (through GFW) maintains a centralized blacklist without much involvement from ISPs, and so without much risk of leakage; null routing adds only a tiny load to the gateway router of ISPs; and no dedicated devices are needed. IP blocking is easy to circumvent, however, by setting up a proxy outside of China or moving the Web site to another IP address. If the Web site changes its IP address and keeps its domain name unchanged, then the user will always access it—whatever IP address it uses.

DNS INJECTION

The first step in surfing the Web is to query the DNS for the IP address of the domain name (for example, www.facebook.com). GFW disturbs the DNS resolution by DNS injection. With its DPI devices deployed near all the international gateways, GFW can monitor each DNS query originating

FIGURE 1 Censorship Technologies by GFW



from any DNS server (resolver) or end computer inside China. If GFW sees any sensitive query, such as “www.facebook.com,” it will inject a faked DNS reply with an invalid IP address. In most cases, the faked reply arrives much earlier than the legitimate one, and the DNS server will accept the first one and forward it to the user. Because GFW spoofs the IP addresses of the legitimate DNS name servers outside China, the DNS server in China cannot distinguish the faked answers from the legitimate ones.²⁷

GFW manipulates DNS traffic at a few international gateways of ISPs, so the government doesn’t need to tamper with the thousands of DNS-resolving servers distributed at the edges of ISPs. Almost all the DNS resolvers in China are polluted.¹⁷

TCP RST

GFW works as an IPS (intrusion protection system) that inspects all traffic and stops “bad” communication if it detects a sensitive keyword. To terminate a TCP connection, GFW injects a series of TCP reset packets with the spoofed source address, port number, and sequence number, without suppressing the original request or legitimate response. With TCP RST, GFW can dynamically terminate any TCP connection with a sensitive keyword—without the knowledge of the address or domain name where the “bad” information originates. It simply relies on a list of sensitive keywords.⁸ (Jedidiah R. Crandall of the University of New Mexico et al. published a paper on how to detect these keywords.^{9,10})

Unlike a firewall that has to take on the burden of forwarding all the “good” traffic, as well as adding another single point of failure, GFW is lightweight in nature. According to papers published by the designers of GFW,^{4,28} its architecture is scalable for high-bandwidth backbones.

With the encrypted traffic of HTTPS (e.g., Gmail), GFW cannot see any keywords, but it can kill the encrypted connections blindly by the injection of TCP RST and degrades the service of these Web sites. The Chinese government insists this instability is caused by the poor service of the Web sites, but few Chinese network users believe them. Anyway, this degraded service forces them to give up services outside China (such as Google) and switch to domestic ones (such as Baidu).

COLLATERAL DAMAGE

The censorship technologies just mentioned—IP blocking, DNS injection, and TCP RST—also cause collateral damage (i.e., unintended blocking of Web sites or content).

TCP RST is one of the most effective methods, as it kills any connections on a blacklist of keywords. For example, in March 2010 Google’s search engine in Hong Kong was blocked because the string “rfa” appeared as a part of search parameters. For GFW, “rfa” means Radio Free Asia.¹⁹ Similarly, because the names of party leaders (such as Hu, Xi, and Wen) are often sensitive words and forbidden in search engines, searches for the Chinese terms *Xue Xi* (study), *Hu Luo Bo* (carrot), and *Wen Du Ji* (thermometer) are also likely to be banned.

IP blocking causes collateral damage because multiple Web sites can be hosted on the same IP address. The adversary of GFW, the censored Web site, can also deliberately introduce collateral damage. For example, GFW blocked the IP address of www.mit.edu because www.falundafa.org, a Web site banned by Chinese government, resolved to the same address.²⁶ Because of this, MIT’s OpenCourseWare site was also blocked. This led to such a major outcry that GFW revoked the block.

The collateral damage caused by GFW is not limited to China. Even if you live in the United States and access a Web site outside China, you could be censored. Two factors contribute to such collateral damage: BGP routing among ISPs and the iteration of DNS resolutions.

PREFIX HIJACKING

GFW hijacks all traffic to blocked Web sites by announcing the routing prefixes (networks) to Chinese ISPs via BGP. This technology is the same as that used by Pakistan to block YouTube.com, which caused blockage of YouTube.com worldwide in 2007.²⁹ If a Chinese ISP—for example, China Telecom—re-announces these prefixes to its neighbor ISPs outside China, and its neighbors accept these prefixes, then the neighbor ISPs could redirect the traffic meant for these blacklisted Web sites to GFW. Some reports have shown that in early 2010 Chinese ISPs hijacked a large amount of Internet traffic, intentionally or incidentally, demonstrating that GFW has the capability to censor at least some access by American users.^{21,24}

DNS POLLUTION

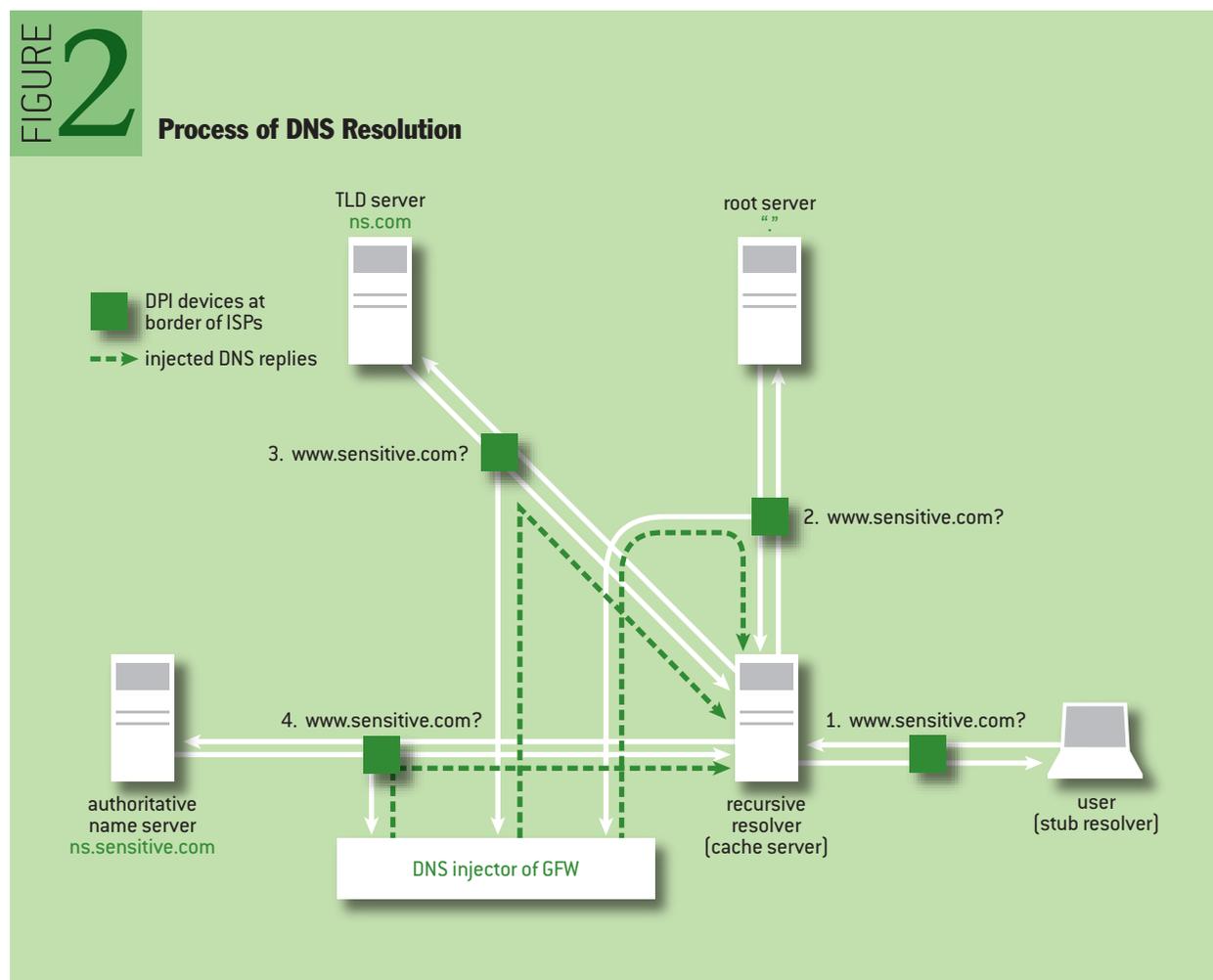
It has been proven that the DNS injection used by GFW has caused a lot of collateral damage. This means that an access point whose source and destination are both outside of China could be blocked because of DNS injection. Three factors cause this:

- Some Chinese ISPs are transit autonomous systems that provide connections for other ISPs and relay traffic among them, especially those in East Asia and Europe.

- Several root servers (F, I, J) are hosted in China. (The list of current DNS root servers is available at <http://www.root-servers.org>.) The ISPs hosting the mirrors of root servers announce their prefixes to neighbor ISPs, such as those in KR (South Korea) or DE (Germany), so the DNS resolvers in these ISPs will direct their DNS queries to the root servers in China.
- One DNS query is composed of a series of iterative subqueries, as shown in figure 2. As long as any of these subqueries go through Chinese ISPs, GFW can inject a forged reply and block or redirect this access.

For example, if a user in South Korea (KR) wants to access the Web site `www.sensitive.de`, where *sensitive* is a blocked domain name by GFW, then the user's DNS server (recursive resolver) will send out a series of queries to the root server ("."), TLD server (".de"), and authoritative name server ("sensitive.de"), with a full domain name ("www.sensitive.de"). If the user's ISP selects one of the root servers in China or routes the query to a TLD (top-level domain) server or to an authoritative name server through China, then GFW will censor this access.

The root DNS server pollution was first detected by a Chilean DNS operator.¹³ Martin A. Brown et al. of Renesys Corporation analyzed this incident and determined that this kind of pollution could affect many countries, because three root DNS server nodes (F, I, and J) have anycast instances in China.³ They believed that after Netnod withdrew the anycast routes for the Chinese I-root



name server from CNNIC, the collateral damage should disappear, but the collateral damage is still going on. An anonymous author measured more than 43,000 open DNS resolvers in 173 countries, excluding China, and found that 26 percent of them in 109 countries were polluted (table 1) for querying some domain names with a substring of blocked keywords such as `www.facebook.com.de`.² Most of the damage arises from censored transit paths to TLD servers (particularly the authorities for DE and KR) instead of root DNS servers.

TABLE 1 Polluted Resolvers in Different Countries

Rank	Region	Affected Resolvers	Affected Rate
1	IR	157	88.20%
2	MY	163	85.34%
3	KR	198	79.20%
4	HK	403	74.63%
5	TW	1146	66.13%
6	IN	250	60.10%
10	IT	392	37.23%
14	JP	1437	29.39%
16	RU	835	25.26%
18	US	3032	24.22%
20	CA	272	23.65%
25	DE	470	20.04%

Total 109 Affected Regions

THE FIGHT AGAINST GFW

Excepting collateral damage, GFW is very successful in blocking the majority of network users from access to sensitive political information. Nobody knows exactly how many Web sites have been blocked or filtered, but most Chinese network users do know the world's most popular Web sites are partially or completely unavailable to them, including Google, Facebook, Twitter, and YouTube.

Chinese network users are angry over Internet censorship by the government because the Internet is the only remaining medium by which they can get uncensored information and express their ideas. With no legal way to express their anger, Chinese network users attacked “the father of GFW” to release their rage,⁶ throwing eggs and shoes at him at Wuhan University¹ and defacing his Web site by replacing its content with the “Angry Shoes” image¹¹ shown in figure 3.

Because GFW blocks the destination (IP address or domain name) and inspects the channel, the basic strategy for bypassing GFW is to find some proxy nodes and encrypt the traffic. Most circumvention tools combine these two mechanisms, because just using a simple open proxy (HTTP or SOCKS) or an encrypted tunnel (such as HTTPS) does little to circumvent the sophisticated censors.

FreeGate, Ultrasurf, and Psiphon (version 3) are popular in China because they are free and intended for nontechnical users. They depend on a range of proxy servers outside China and encrypt all the HTTP traffic in SSL (Secure Sockets Layer) tunnels to these servers. Although most of this software is not open source, the government can analyze it by reverse engineering and then block all

FIGURE
3



the proxy servers. Thus, the software has to be updated now and then. Because it is not open source, some people worry that some of the software could contain Trojan horses.

Tor, the famous anonymous communication tool, was once used widely in China. The complicated encrypted protocols (also based on SSL) and thousands of proxies make Tor an ideal tool for bypassing the blocking and surveillance of GFW. The centralized directory server by which users get the list of proxy nodes, however, is the fatal defect in terms of anti-censorship. After GFW blocked the IP address of the Tor directory in 2008, Tor lost most of its users in China.

VPN (virtual private network) and SSH (secure shell) are the most powerful and stable tools for bypassing all surveillance technologies, although the basic ideas are the same as with the aforementioned tools: proxies and encrypted channels. The only difference is that VPN and SSH depend on a private host (or virtual host) or an account outside of China, instead of open, free proxies. Only technical professionals are able to set up such hosts or accounts, and most of them are not free. Commercial or public VPN services will be blocked by IP address and/or domain names if they are popular enough. In fact, the domain names *vpn.* are all blocked (such as vpn.com, vpn.net, vpn.org, vpn.info, vpn.me, vpn.us, vpn.co).

Although no perfect solution exists for all users and no single solution is a guarantee, the circumvention tools do cut a hole through the great wall.

THE TUSSLE CONTINUES

The arms race between GFW and circumvention continues. GFW has an advantage over individuals because it controls all Chinese resources in cyberspace, as well as in society: computing and storage resources, ISP backbones, DNS servers, and police and/or legal punishment. GFW is not powerful enough, however, to control Internet infrastructure and communities in the broader world.

The battle between Tor and GFW exemplified this arms race. After the directory server was blocked by GFW, Tor developed some “bridge” nodes that are not listed by the directory. Users can set up their own private bridge nodes and thereby connect to the Tor network again. In 2011,

however, the updated GFW was able to detect the hidden private bridge nodes by signature and block them dynamically, so Chinese users lost the Tor network again.²² In early 2012, Tor released obfsproxy, which can obfuscate the traffic between the Tor client and the bridge nodes. In this way, GFW could see only innocent-looking transformed traffic and could not detect any signature. Now the number of Tor users in China has increased again.²³

Cloud computing brings new challenges to GFW. Powered by virtualization, cloud computing provides more computing power at lower prices. More individuals can afford a VPS (virtual private server) or SSH account with Amazon or DreamHost. With more restrictions or crackdowns on domestic Web sites, many small businesses host their Web sites outside of China. Theoretically, cloud-computing infrastructure could change the IP addresses for its customers' Web sites, and thereby lessen the chance of censorship.

IPv6 is another challenge to GFW. The Chinese government encourages IPv6 networks at least for academic purposes, so hundreds of universities in China have established their IPv6 networks with government sponsorship. Now thousands of students can access the IPv6-enabled services (such as youtube.com). The only obstacle is that DNS infrastructure is still running over IPv4, although it can issue some IPv6 addresses. The DNS injection deployed in IPv4 still works for blocking IPv6-enabled Web sites. From November 2012, when the Chinese Communist Party 18th Congress was held, some censors on IPv6 (using techniques similar to the IPv4 censors) were reported by Chinese netizens.

Besides the tussles with circumvention, the Chinese government itself is always in a dilemma: whether to encourage the development of the Internet and related new technologies that are coupled with its economy, or to restrict the development of these technologies because of challenges to the traditional censorship systems. The free flow of information could undermine the authoritarian state's control; however, a too-severe crackdown on the Internet (such as cutting fiber), could hurt China's economic development and potentially the control of the state.

Once China opened its door to the world, it could not close it again. The dramatic series of events that took place recently during the Arab Spring demonstrated the power of the Internet to change autocratic societies, and spurred the Chinese government to spare no effort to improve Internet censorship systems. It is naive to expect that the Chinese government will give up its censorship activities in the near future; however, it is impossible to shut down the Internet now. The tussle and the arm race will continue at least into the near future.

REFERENCES

1. "Angry Shoes." Flickr; <http://www.flickr.com/photos/isaacmao/5738596950/>.
2. Anonymous. 2012. The collateral damage of Internet censorship by DNS injection. *ACM SIGCOMM Computer Communication Review* 42(3): 21-27.
3. Brown, M. A., Madory, D., Popescu, A., Zmijewski, E. 2010. DNS tampering and root servers; <http://www.renysys.com/tech/presentations/pdf/DNS-Tampering-and-Root-Servers.pdf>.
4. Chen, X., Fang, B., Li, L. Architecture of intrusion detection for high-speed networks. 2004. *Journal of Computer Research and Development* 41 (September): 1481-1487.
5. China Internet Network Information Center. 2012. 29th Statistical Report on Internet Development in China; http://www.apira.org/data/upload/The29thStatisticalReportonInternetDevelopmentinChina_P9G97q.pdf.
6. China's great firewall designer "hit by shoe." 2011. BBC News Asia-Pacific; <http://www.bbc.co.uk/>

- news/world-asia-pacific-13455819.
7. Chovanec, P. 2011. Al Jazeera: Internet censorship in China; <http://chovanec.wordpress.com/2010/01/02/al-jazeera-internet-censorship-in-china/>.
 8. Clayton, R., Murdoch, S., Watson, R. 2006. Ignoring the Great Firewall of China. In *Privacy Enhancing Technologies*. Springer: 20-35.
 9. ConceptDoppler; <http://www.conceptdoppler.org>.
 10. Crandall, J., Zinn, D., Byrd, M., Barr, E., East, R. 2007. ConceptDoppler: a weather tracker for Internet censorship. In *Proceedings of the 14th ACM Conference on Computer and Communications Security*: 1–4.
 11. Defaced Web site in Beijing University of Posts and Communications; <http://yilee.info/media/fang-xiao-zhang/1.jpg>.
 12. Deibert, R., Palfrey, J. G., Initiative, O., Rohozinski, R., Zittrain, J. 2010. Access Controlled. The Shaping of Power, Rights, and Rule in Cyberspace. MIT Press: 264.
 13. Ereche, M. V. 2010. Odd behaviour on one node in I root-server; <https://lists.dns-oarc.net/pipermail/dns-operations/2010-March/005260.html>.
 14. History of GFW and Binxing Fang. (in Chinese) 2010; <https://fangbinxing.appspot.com/2010/08/10/fangbinxing.html>.
 15. Hunt, T. 2012. Browsing the broken Web: a software developer behind the Great Firewall of China; <http://www.troyhunt.com/2012/03/browsing-broken-web-software-developer.html>.
 16. Liu, G., Yun, X., Fang, B., Hu, M. 2003. A control method for large-scale network based on routing diffusion. *Journal of China Institute of Communications*: 10.
 17. Lowe, G., Winters, P., Marcus, M. L. 2007. The Great DNS Wall of China; <http://cs.nyu.edu/%7Epcw216/work/nds/final.pdf>.
 18. People's Posts and Telecommunications News (PPTN). Report about President Fang: be careful about Chinese input software. (in Chinese) 2010; <http://www.cnii.com.cn/20080623/ca615907.htm>.
 19. Schonfeld, E. 2010. Update: China's firewall mistakes Google for Radio Free Asia (or not). TechCrunch; <http://techcrunch.com/2010/03/30/china-firewall-google-radio-free-asia/>.
 20. Smith, J. R. 2011; <https://twitter.com/#!/TheRealJRSmith/statuses/12818166222794752>.
 21. Toonk, A. 2010. Chinese ISP hijacks the Internet. BGPmon; <http://bgpmon.net/blog/?p=282>.
 22. Tor. 2011. Bridge easily detected by GFW. Ticket #4185. Tor Bug Tracker Wiki (October); <https://trac.torproject.org/projects/tor/ticket/4185>.
 23. Tor obfsproxy; <https://www.torproject.org/projects/obfsproxy.html.en>.
 24. U.S.-China Economic and Security Review Commission. 2010. Report to Congress: 241; http://www.uscc.gov/annual_report/2010/annual_report_full_10.pdf.
 25. Wikipedia. Great Firewall of China; http://en.wikipedia.org/wiki/Great_Firewall_of_China.
 26. Winstein, K. J. 2002. China blocks MIT Web addresses. *The Tech* 122(58); <http://tech.mit.edu/V122/N58/58web.58n.html>.
 27. Yan, B., Fang, B., Li, B., Wang, Y. 2006. Detection and defense of DNS spoofing attack. *Computer Engineering* 32(21).
 28. Yang, W., Fang, B., Yun, X., Zhang, H. 2004. A parallel cluster intrusion-detection system for backbone network. *Journal of Harbin Institute of Technology* 3.
 29. YouTube hijacking: RIPE NCC RIS case study. 2008. RIPE Network Coordination Centre; <http://>

www.ripe.net/internet-coordination/news/industry-developments/youtube-hijacking-a-ripe-nccris-case-study.

LOVE IT, HATE IT? LET US KNOW

feedback@queue.acm.org

AUTHOR BIO

Daniel Anderson is a researcher on Internet technologies and policies. You can contact Daniel by email: daniel.anderson.us@gmail.com.

© 2012 ACM 1542-7730/11/1000 \$10.00