China’s Great Firewall is crude, slapdash, and surprisingly easy to breach. Here’s why it’s so effective anyway.

BY JAMES FALLOWS

“The Connection Has Been Reset”

Illustration by John Ritter

Any foreigners who come to China for the Olympics will use the Internet to tell people back home what they have seen and to check what else has happened in the world.

The first thing they’ll probably notice is that China’s Internet seems slow. Partly this is because of congestion in China’s internal networks, which affects domestic and international transmissions alike. Partly it is because even electrons take a detectable period of time to travel beneath the Pacific Ocean to servers in America and back again; the trip to and from Europe is even longer, because that goes through America, too. And partly it is because of the delaying cycles imposed by China’s system that monitors what people are looking for on the Internet, especially when they’re looking overseas. That’s what foreigners have heard about.

They’ll likely be surprised, then, to notice that China’s Internet seems surprisingly free and uncontrolled. Can they search for information about “Tibet independence” or “Tiananmen shooting” or other terms they have heard are taboo? Probably—and they’ll be able to click right through to the controversial sites. Even if they enter the Chinese-language term for “democracy in China,” they’ll probably get results. What about Wikipedia, famously off-limits to users in China? They will probably be able to reach it. Naturally the visitors will wonder: What’s all this I’ve heard about the “Great Firewall” and China’s tight limits on the Internet?

In reality, what the Olympic-era visitors will be discovering is not the absence of China’s electronic control but its new refinement—and a special Potemkin-style unfettered access that will be set up just for them, and just for the length of their stay. According to engineers I have spoken with at two tech organizations in China, the government bodies in charge of censoring the Internet have told them to get ready to unblock access from a list of specific Internet Protocol (IP) addresses—certain Internet cafés, access jacks in hotel rooms and conference centers where foreigners are expected to work or stay during the Olympic Games. (I am not giving names or identifying details of any Chinese citizens with whom I have discussed this topic, because they risk financial or criminal punishment for criticizing the system or even disclosing how it works. Also, I have not gone to Chinese government agencies for their side of the story, because the very existence of Internet controls is almost never...
discussed in public here, apart from vague statements about the importance of keeping online information “wholesome.”

Depending on how you look at it, the Chinese government’s attempt to rein in the Internet is crude and slapdash or ingenious and well crafted. When American technologists write about the control system, they tend to emphasize its limits. When Chinese citizens discuss it—at least with me—they tend to emphasize its strength. All of them are right, which makes the government’s approach to the Internet a nice proxy for its larger attempt to control people’s daily lives.

Disappointingly, “Great Firewall” is not really the right term for the Chinese government’s overall control strategy. China has indeed erected a firewall—a barrier to keep its Internet users from dealing easily with the outside world—but that is only one part of a larger, complex structure of monitoring and censorship. The official name for the entire approach, which is ostensibly a way to keep hackers and other rogue elements from harming Chinese Internet users, is the “Golden Shield Project.” Since that term is too creepy to bear repeating, I’ll use “the control system” for the overall strategy, which includes the “Great Firewall of China,” or GFW, as the means of screening contact with other countries.

In America, the Internet was originally designed to be free of choke points, so that each packet of information could be routed quickly around any temporary obstruction. In China, the Internet came with choke points built in. Even now, virtually all Internet contact between China and the rest of the world is routed through a very small number of fiber-optic cables that enter the country at one of three points: the Beijing-Qingdao-Tianjin area in the north, where cables come in from Japan; Shanghai on the central coast, where they also come from Japan; and Guangzhou in the south, where they come from Hong Kong. (A few places in China have Internet service via satellite, but that is both expensive and slow. Other lines run across Central Asia to Russia but carry little traffic.) In late 2006, Internet users in China were reminded just how important these choke points are when a seabed earthquake near Taiwan cut some major cables serving the country. It took months before international transmissions to and from most of China regained even their pre-quake speed, such as it was.

Thus Chinese authorities can easily do something that would be harder in most developed countries: physically monitor all traffic into or out of the country. They do so by installing at each of these few “international gateways” a device called a “tapper” or “network sniffer,” which can mirror every packet of data going in or out. This involves mirroring in both a figurative and a literal sense. “Mirroring” is the term for normal copying or backup operations, and in this case real though extremely small mirrors are employed. Information travels along fiber-optic cables as little pulses of light, and as these travel through the Chinese gateway routers, numerous tiny mirrors bounce reflections of them to a separate set of “Golden Shield” computers. Here the term’s creepiness is appropriate. As the other routers and servers (short for file servers, which are essentially very large-capacity computers) that make up the Internet do their best to get the packet where it’s supposed to go, China’s own surveillance computers are looking over the same information to see whether it should be stopped.

The mirroring routers were first designed and supplied to the Chinese authorities by the U.S. tech firm Cisco, which is why Cisco took such heat from human-rights organizations. Cisco has always denied that it tailored its equipment to the authorities’ surveillance needs, and said it merely sold them what it would sell anyone else. The issue is now moot, since similar routers are made by companies around the world, notably including China’s own electronics giant, Huawei. The ongoing refinements are mainly in surveillance software, which the Chinese are developing themselves. Many of the surveillance engineers are thought to come from the military’s own technology institutions. Their work is good and getting better, I was told by Chinese and foreign engineers who do “oppo research” on the evolving GFW so as to design better ways to get around it.
Andrew Lih, a former journalism professor and software engineer now based in Beijing (and author of the forthcoming book *The Wikipedia Story*), laid out for me the ways in which the GFW can keep a Chinese Internet user from finding desired material on a foreign site. In the few seconds after a user enters a request at the browser, and before something new shows up on the screen, at least four things can go wrong—or be made to go wrong.

The first and bluntest is the “DNS block.” The DNS, or Domain Name System, is in effect the telephone directory of Internet sites. Each time you enter a Web address, or URL—www.yahoo.com, let’s say—the DNS looks up the IP address where the site can be found. IP addresses are numbers separated by dots—for example, TheAtlantic.com’s is 38.118.42.200. If the DNS is instructed to give back no address, or a bad address, the user can’t reach the site in question—as a phone user could not make a call if given a bad number. Typing in the URL for the BBC’s main news site often gets the no-address treatment: if you try news.bbc.co.uk, you may get a “Site not found” message on the screen. For two months in 2002, Google’s Chinese site, Google.cn, got a different kind of bad-address treatment, which shunted users to its main competitor, the dominant Chinese search engine, Baidu. Chinese academics complained that this was hampering their work. The government, which does not have to stand for reelection but still tries not to antagonize important groups needlessly, let Google.cn back online. During politically sensitive times, like last fall’s 17th Communist Party Congress, many foreign sites have been temporarily shut down this way.

Next is the perilous “connect” phase. If the DNS has looked up and provided the right IP address, your computer sends a signal requesting a connection with that remote site. While your signal is going out, and as the other system is sending a reply, the surveillance computers within China are looking over your request, which has been mirrored to them. They quickly check a list of forbidden IP sites. If you’re trying to reach one on that blacklist, the Chinese international-gateway servers will interrupt the transmission by sending an Internet “Reset” command both to your computer and to the one you’re trying to reach. Reset is a perfectly routine Internet function, which is used to repair connections that have become unsynchronized. But in this case it’s equivalent to forcing the phones on each end of a conversation to hang up. Instead of the site you want, you usually see an onscreen message beginning “The connection has been reset”; sometimes instead you get “Site not found.” Annoyingly, blogs hosted by the popular system Blogspot are on this IP blacklist. For a typical Google-type search, many of the links shown on the results page are from Wikipedia or one of these main blog sites. You will see these links when you search from inside China, but if you click on them, you won’t get what you want.

The third barrier comes with what Lih calls “URL keyword block.” The numerical Internet address you are trying to reach might not be on the blacklist. But if the words in its URL include forbidden terms, the connection will also be reset. (The Uniform Resource Locator is a site’s address in plain English—say, www.microsoft.com—rather than its all-numeric IP address.) The site FalunGong .com appears to have no active content, but even if it did, Internet users in China would not be able to see it. The forbidden list contains words in English, Chinese, and other languages, and is frequently revised—“like, with the name of the latest town with a coal mine disaster,” as Lih put it. Here the GFW’s programming technique is not a reset command but a “black-hole loop,” in which a request for a page is trapped in a sequence of delaying commands. These are the programming equivalent of the old saw about how to keep an idiot busy: you take a piece of paper and write “Please turn over” on each side. When the Firefox browser detects that it is in this kind of loop, it gives an error message saying: “The server is redirecting the request for this address in a way that will never complete.”

The final step involves the newest and most sophisticated part of the GFW: scanning the actual contents of each page—which stories The New York Times is featuring, what a China-related blog carries in its latest update—to judge its page-by-page acceptability. This again is done with mirrors. When you reach a favorite blog or news site
and ask to see particular items, the requested pages come to you—and to the surveillance system at the same
time. The GFW scanner checks the content of each item against its list of forbidden terms. If it finds something it
doesn’t like, it breaks the connection to the offending site and won’t let you download anything further from it.
The GFW then imposes a temporary blackout on further “IP1 to IP2” attempts—that is, efforts to establish
communications between the user and the offending site. Usually the first time-out is for two minutes. If the
user tries to reach the site during that time, a five-minute time-out might begin. On a third try, the time-out
might be 30 minutes or an hour—and so on through an escalating sequence of punishments.

Users who try hard enough or often enough to reach the wrong sites might attract the attention of the authorities.
At least in principle, Chinese Internet users must sign in with their real names whenever they go online, even in
Internet cafés. When the surveillance system flags an IP address from which a lot of “bad” searches originate, the
authorities have a good chance of knowing who is sitting at that machine.

All of this adds a note of unpredictability to each attempt to get news from outside China. One day you go to the
NPR site and cruise around with no problem. The next time, NPR happens to have done a feature on Tibet. The
GFW immobilizes the site. If you try to refresh the page or click through to a new story, you’ll get nothing—and
the time-out clock will start.

This approach is considered a subtler and more refined form of censorship, since big foreign sites no longer need
be blocked wholesale. In principle they’re in trouble only when they cover the wrong things. Xiao Qiang, an
expert on Chinese media at the University of California at Berkeley journalism school, told me that the
authorities have recently begun applying this kind of filtering in reverse. As Chinese-speaking people outside the
country, perhaps academics or exiled dissidents, look for data on Chinese sites—say, public-health figures or
news about a local protest—the GFW computers can monitor what they’re asking for and censor what they find.

Taken together, the components of the control system share several traits. They’re constantly evolving and
changing in their emphasis, as new surveillance techniques become practical and as words go on and off the
sensitive list. They leave the Chinese Internet public unsure about where the off-limits line will be drawn on any
given day. Andrew Lih points out that other countries that also censor Internet content—Singapore, for instance,
or the United Arab Emirates—provide explanations whenever they do so. Someone who clicks on a pornographic
or “anti-Islamic” site in the U.A.E. gets the following message, in Arabic and English: “We apologize the site you
are attempting to visit has been blocked due to its content being inconsistent with the religious, cultural,
political, and moral values of the United Arab Emirates.” In China, the connection just times out. Is it your
computer’s problem? The firewall? Or maybe your local Internet provider, which has decided to do some filtering
on its own? You don’t know. “The unpredictability of the firewall actually makes it more effective,” another
Chinese software engineer told me. “It becomes much harder to know what the system is looking for, and you
always have to be on guard.”

There is one more similarity among the components of the firewall: they are all easy to thwart.

As a practical matter, anyone in China who wants to get around the firewall can choose between two well-known
and dependable alternatives: the proxy server and the VPN. A proxy server is a way of connecting your computer
inside China with another one somewhere else—or usually to a series of foreign computers, automatically
passing signals along to conceal where they really came from. You initiate a Web request, and the proxy system
takes over, sending it to a computer in America or Finland or Brazil. Eventually the system finds what you want
and sends it back. The main drawback is that it makes Internet operations very, very slow. But because most
proxies cost nothing to install and operate, this is the favorite of students and hackers in China.

A VPN, or virtual private network, is a faster, fancier, and more elegant way to achieve the same result.
Essentially a VPN creates your own private, encrypted channel that runs alongside the normal Internet. From within China, a VPN connects you with an Internet server somewhere else. You pass your browsing and downloading requests to that American or Finnish or Japanese server, and it finds and sends back what you’re looking for. The GFW doesn’t stop you, because it can’t read the encrypted messages you’re sending. Every foreign business operating in China uses such a network. VPNs are freely advertised in China, so individuals can sign up, too. I use one that costs $40 per year. (An expat in China thinks: that’s a little over a dime a day. A Chinese factory worker thinks: it’s a week’s take-home pay. Even for a young academic, it’s a couple days’ work.)

As a technical matter, China could crack down on the proxies and VPNs whenever it pleased. Today the policy is: if a message comes through that the surveillance system cannot read because it’s encrypted, let’s wave it on through! Obviously the system’s behavior could be reversed. But everyone I spoke with said that China could simply not afford to crack down that way. “Every bank, every foreign manufacturing company, every retailer, every software vendor needs VPNs to exist,” a Chinese professor told me. “They would have to shut down the next day if asked to send their commercial information through the regular Chinese Internet and the Great Firewall.” Closing down the free, easy-to-use proxy servers would create a milder version of the same problem. Encrypted e-mail, too, passes through the GFW without scrutiny, and users of many Web-based mail systems can establish a secure session simply by typing “https:” rather than the usual “http:” in a site’s address—for instance, https://mail.yahoo.com. To keep China in business, then, the government has to allow some exceptions to its control efforts—even knowing that many Chinese citizens will exploit the resulting loopholes.

Because the Chinese government can’t plug every gap in the Great Firewall, many American observers have concluded that its larger efforts to control electronic discussion, and the democratization and grass-roots organizing it might nurture, are ultimately doomed. A recent item on an influential American tech Web site had the headline “Chinese National Firewall Isn’t All That Effective.” In October, Wired ran a story under the headline “The Great Firewall: China’s Misguided—and Futile—Attempt to Control What Happens Online.”

Let’s not stop to discuss why the vision of democracy-through-communications-technology is so convincing to so many Americans. (Samizdat, fax machines, and the Voice of America eventually helped bring down the Soviet system. Therefore proxy servers and online chat rooms must erode the power of the Chinese state. Right?) Instead, let me emphasize how unconvincing this vision is to most people who deal with China’s system of extensive, if imperfect, Internet controls.

Think again of the real importance of the Great Firewall. Does the Chinese government really care if a citizen can look up the Tiananmen Square entry on Wikipedia? Of course not. Anyone who wants that information will get it—by using a proxy server or VPN, by e-mailing to a friend overseas, even by looking at the surprisingly broad array of foreign magazines that arrive, uncensored, in Chinese public libraries.

What the government cares about is making the quest for information just enough of a nuisance that people generally won’t bother. Most Chinese people, like most Americans, are interested mainly in their own country. All around them is more information about China and things Chinese than they could possibly take in. The newsstands are bulging with papers and countless glossy magazines. The bookstores are big, well stocked, and full of patrons, and so are the public libraries. Video stores, with pirated versions of anything. Lots of TV channels. And of course the Internet, where sites in Chinese and about China constantly proliferate. When this much is available inside the Great Firewall, why go to the expense and bother, or incur the possible risk, of trying to look outside?

All the technology employed by the Golden Shield, all the marvelous mirrors that help build the Great Firewall—these and other modern achievements matter mainly for an old-fashioned and pre-technological reason. By
making the search for external information a nuisance, they drive Chinese people back to an environment in which familiar tools of social control come into play.

Chinese bloggers have learned that if they want to be read in China, they must operate within China, on the same side of the firewall as their potential audience. Sure, they could put up exactly the same information outside the Chinese mainland. But according to Rebecca MacKinnon, a former Beijing correspondent for CNN now at the Journalism and Media Studies Center of the University of Hong Kong, their readers won't make the effort to cross the GFW and find them. “If you want to have traction in China, you have to be in China,” she told me. And being inside China means operating under the sweeping rules that govern all forms of media here: guidance from the authorities; the threat of financial ruin or time in jail; the unavoidable self-censorship as the cost of defiance sinks in.

Most blogs in China are hosted by big Internet companies. Those companies know that the government will hold them responsible if a blogger says something bad. Thus the companies, for their own survival, are dragooned into service as auxiliary censors.

Large teams of paid government censors delete offensive comments and warn errant bloggers. (No official figures are available, but the censor workforce is widely assumed to number in the tens of thousands.) Members of the public at large are encouraged to speak up when they see subversive material. The propaganda ministries send out frequent instructions about what can and cannot be discussed. In October, the group Reporters Without Borders, based in Paris, released an astonishing report by a Chinese Internet technician writing under the pseudonym “Mr. Tao.” He collected dozens of the messages he and other Internet operators had received from the central government. Here is just one, from the summer of 2006:

17 June 2006, 18:35

From: Chen Hua, deputy director of the Beijing Internet Information Administrative Bureau

Dear colleagues, the Internet has of late been full of articles and messages about the death of a Shenzhen engineer, Hu Xinyu, as a result of overwork. All sites must stop posting articles on this subject, those that have already been posted about it must be removed from the site and, finally, forums and blogs must withdraw all articles and messages about this case.

“Domestic censorship is the real issue, and it is about social control, human surveillance, peer pressure, and self-censorship,” Xiao Qiang of Berkeley says. Last fall, a team of computer scientists from the University of California at Davis and the University of New Mexico published an exhaustive technical analysis of the GFW’s operation and of the ways it could be foiled. But they stressed a nontechnical factor: “The presence of censorship, even if easy to evade, promotes self-censorship.”

It would be wrong to portray China as a tightly buttoned mind-control state. It is too wide-open in too many ways for that. “Most people in China feel freer than any Chinese people have been in the country’s history, ever,” a Chinese software engineer who earned a doctorate in the United States told me. “There has never been a space for any kind of discussion before, and the government is clever about continuing to expand space for anything that doesn’t threaten its survival.” But it would also be wrong to ignore the cumulative effect of topics people are not allowed to discuss. “Whether or not Americans supported George W. Bush, they could not avoid learning about Abu Ghraib,” Rebecca MacKinnon says. In China, “the controls mean that whole topics inconvenient for
the regime simply don’t exist in public discussion.” Most Chinese people remain wholly unaware of internationally noticed issues like, for instance, the controversy over the Three Gorges Dam.

Countless questions about today’s China boil down to: How long can this go on? How long can the industrial growth continue before the natural environment is destroyed? How long can the super-rich get richer, without the poor getting mad? And so on through a familiar list. The Great Firewall poses the question in another form: How long can the regime control what people are allowed to know, without the people caring enough to object? On current evidence, for quite a while.