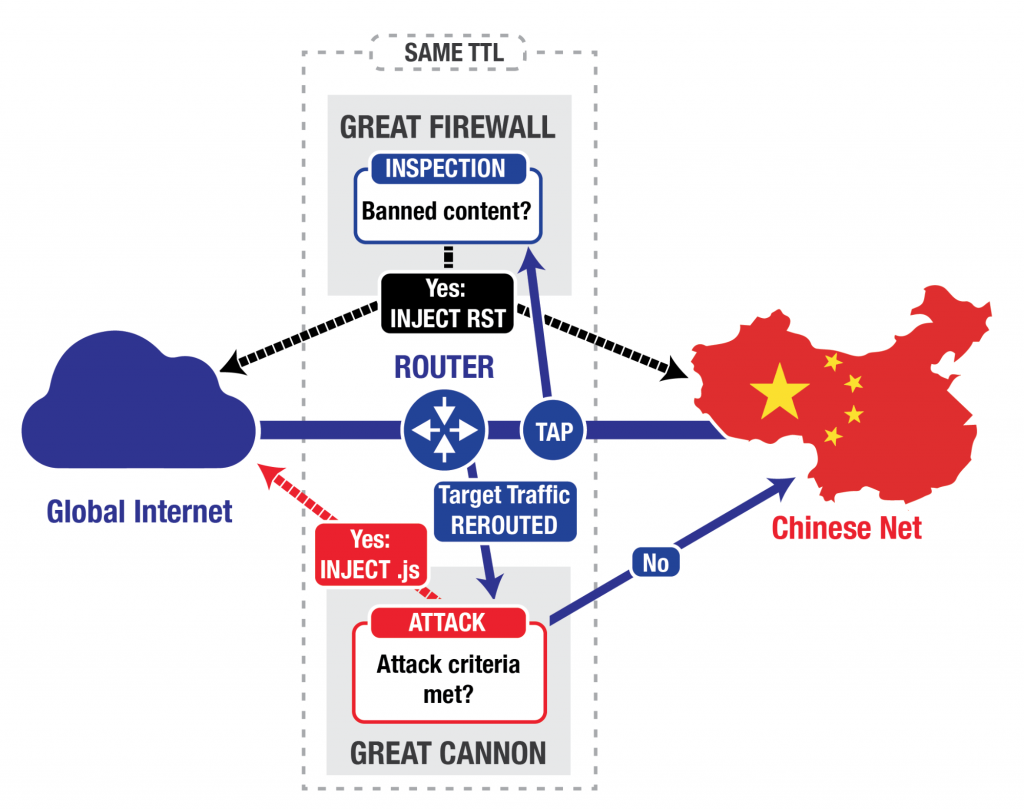
In depth analysis of the Great Firewall and Great Cannon of China

Chao Tang



Abstract: Created by the Golden Shield Project, the Great Firewall of China (GFW) and the Great Cannon (GC) together form world’s largest system of censorship. The GFW is an on-path system. As a result, it can monitor traffic and inject additional packets, but cannot stop in-flight packets from reaching its destination. Under this architecture, it achieves censorship through the following means: First, it eavesdrops on all Internet traffic between China and the rest of the world, then terminate requests for censored content by injecting forged TCP Reset packets to both ends. Second, the GFW blocks access to specific IP addresses. Third, it uses high-level DNS poisoning to prevent access to blocked URLs. On the other hand, the GC is an in-path system that is “capable of not only injecting traffic but also directly suppressing traffic”. It only examines traffic to or from targeted addresses, and hijacks traffic by dropping the original request and sending a malicious script to enlist the requesting user in DDoS attacks. While the GFW is a tool designed to censor the Internet for all Chinese Internet users, the GC is a powerful man-in-the-middle attack tool that targets high profile “enemies”

Overview: In this project, the author will provide an in-depth overview on the functions of the GFW and the GC. Then, the author will specifically explore three incidents involving their use: a large-scale internet outage in China on Jan 21, 2016 suspected to be caused by the GFW, a DDOS attack by the GC against GreatFire.org on March 16, 2015, and the active probing system used by the GFW to block the Tor network.

GreatFire replied to me and said that although they haven’t done any research papers, I can email them with questions. Also, they recommended that I use a VPS in China to test different behaviors.

Sources:

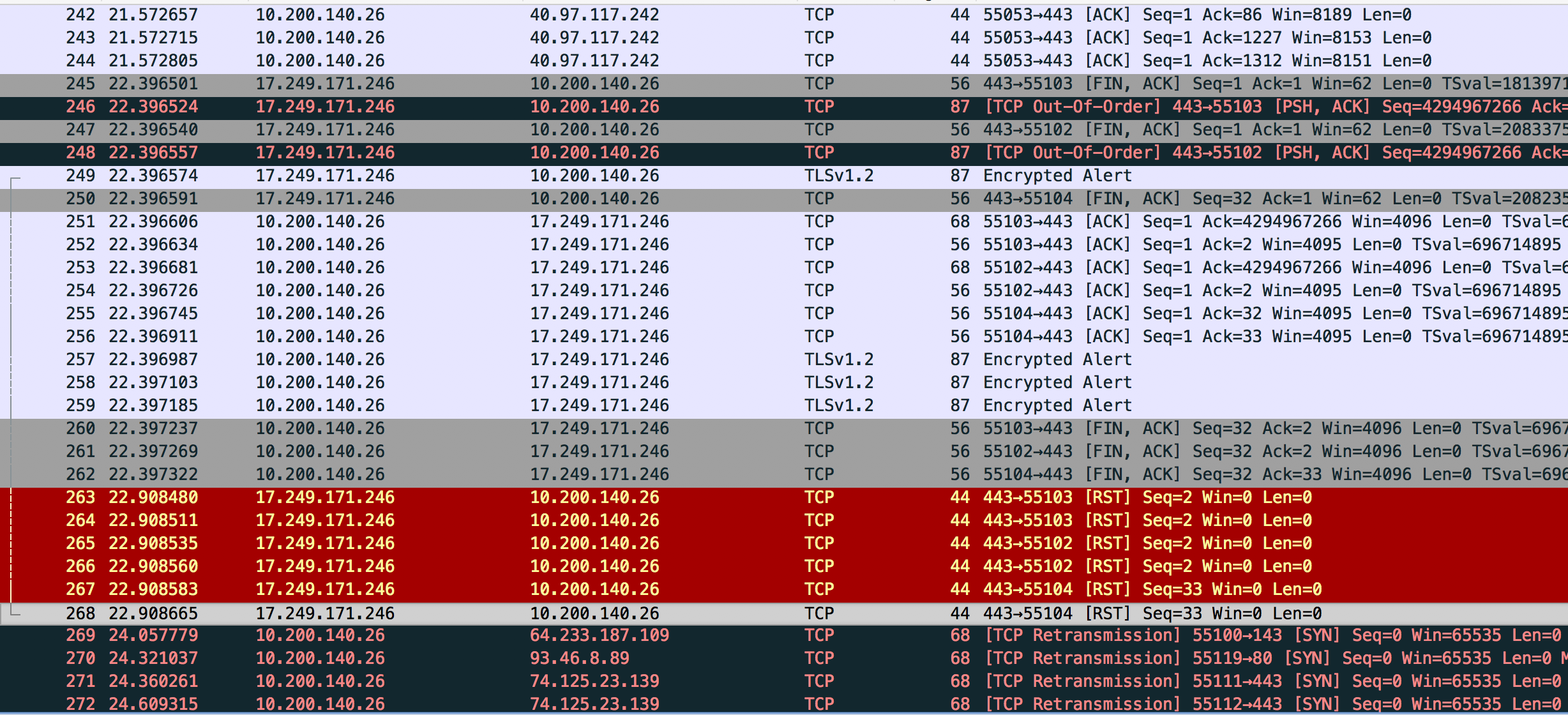
* <https://www.usenix.org/system/files/conference/foci15/foci15-paper-marczak.pdf>
* <https://www.cs.princeton.edu/~pwinter/pdf/ensafi2015a.pdf>
* <https://blog.torproject.org/category/tags/gfw>
* <https://en.greatfire.org/blog/2014/jan/internet-outage-china-jan-21>
* <http://www.cl.cam.ac.uk/~rnc1/ignoring.pdf>

Overview

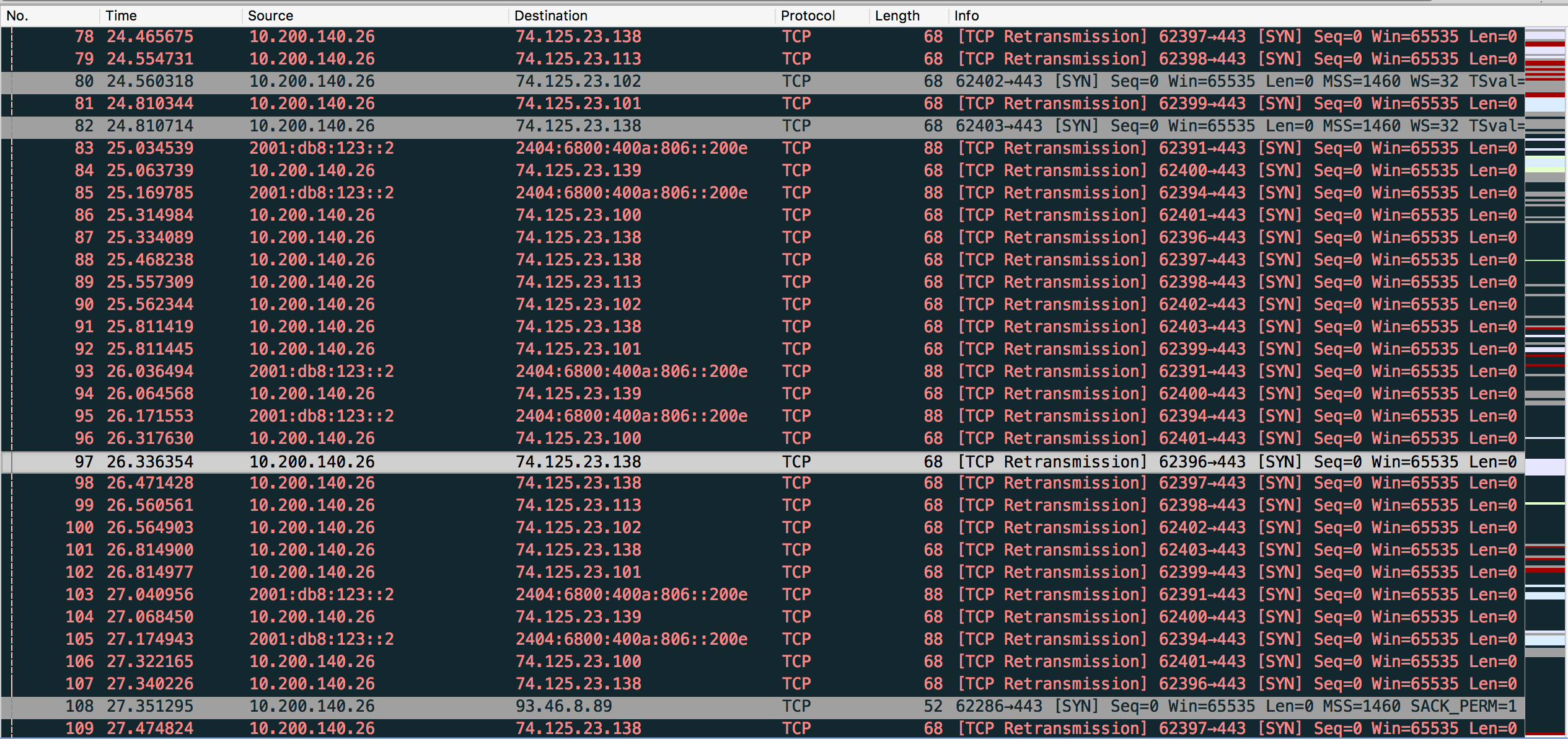
* *Discuss architecture of both systems*
* *On-path vs in-path*
* They are dynamic systems, constantly evolving in an arms race against anti-censorship technology.
* They are complementary systems, existing under one umbrella infrastructure but serving two different goals.

GFW

* TCP Reset: The GFW can’t stop in-flight packets, so it sends TCP RST packets to both ends to terminate the connection
  + Counter-measure: ignore the RST packets if it didn’t originate from the destination IP

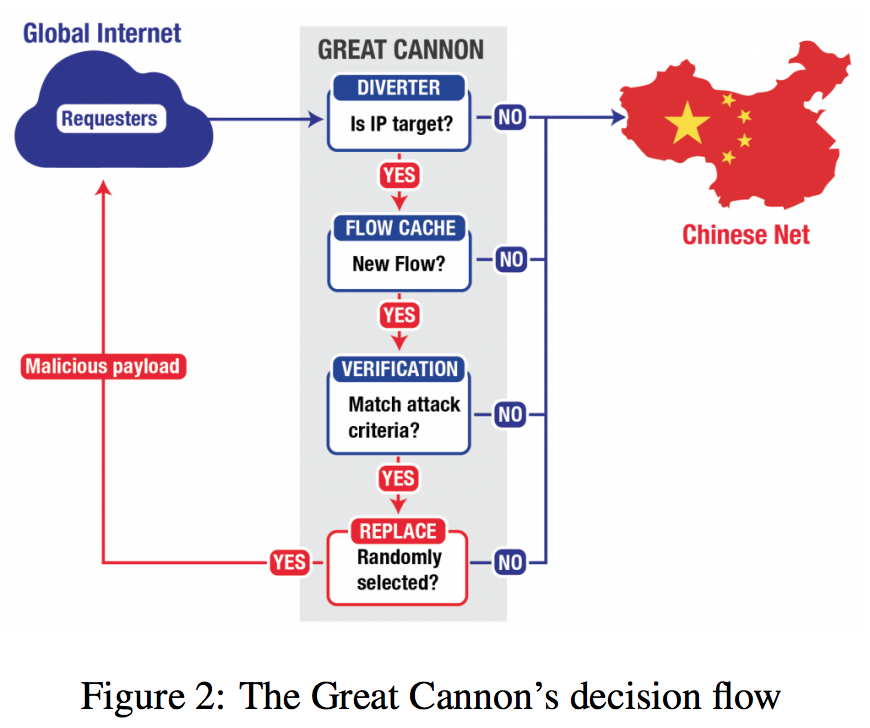


* IP address blocking: The GFW blocks known IP addresses of those against the party or the censorship rules (example: twitter)
  + Counter-measure: get a new IP



* DNS poisoning: The GFW redirects users who try to access blocked sites to random real IP addresses (There are hilarious reports of random servers getting hit with absurd number of requests from GFW redirection: http://furbo.org/2015/01/22/fear-china/)
  + Counter-measure: not possible
* General countermeasures
  + VPNs: Encrypted traffic goes out to a foreign server, which executes the original request and sends back the response also encrypted
    - Actively being targeted by the GFW. Popular connection options like IPsec, L2TP/IPsec, PPTP and even SSTP connections are now being detected and blocked within minutes of connecting
    - Uses machine learning to identify VPN connections! (Fascinating. Will focus more on this in the final paper.)
  + Tor: When someone uses the Onion routing protocol, it encrypts the message, strips away parts of the identifying header, and goes through the Tor relay to disguise the origin and content of the traffic. The vanilla Onion protocol no longer works though, as the GFW employs an active probing system () to reactively probe and block proxies like Tor. (This will be one of the focuses in my paper, and will be talked about in much greater detail.)

Great Cannon

* Man-in-the-middle system for targeted flows, capable of injecting and suppressing traffic.
* Used to compromise foreign visitors to Chinese sites for use in large-scale DDOS attacks
* In the attack on GitHub and GreatFire.org, the GC intercepted traffic sent to Baidu (standard search engine used in China) infrastructure servers that host analytics, social, or advertising scripts. It then probabilistically passes the request unaltered or sends back a malicious script (roughly 1.7% of the time). 
* Specifics:
  + In-path system co-located with the GFW
  + Only acts on the first data packet of a connection
  + Load-balanced architecture

Supporting Material: I will make a poster to showcase what I’ve learned. I intend it to be very graph oriented. One idea I had was to make a line listing the OSI models, and structuring the different measures used by the GFW and GC around that.