The Great Firewall of China

What is it and how does it work?

Student brief - by Alexandra Snoy

15-744 Computer Networks

Monday 24 April 2017

Background

- Largest and most sophisticated internet censorship system in the world, filtering information flow both in and out of China
- Built at the end of the 90’s
- Operated by the Ministry of Industry and Information Technology (MIIT) of China
- Its exact mode of operation is not publicly disclosed, and the system is in constant evolution
- Some international research attempts to figure out how the system works: we will see 3 examples ranging over the last 10 years.

Where does the filtering occur?

- Mostly at the 24 border AS (i.e. peering with other countries) belonging to CHINANET and CNCGROUP, the 1st and 2nd largest ISPs in China respectively.
- Inside these AS, almost 500 routers performing filtering were found by performing probes with increasing TTL from/to different locations
- Almost 80% distributed among CHINANET's different regional routers, several hops into the country, and almost 20% on CNCGROUP's backbone routers → ISPs have different deployment strategies

- Xu X., Mao Z.M., Halderman J.A. - Internet Censorship in China: Where Does the Filtering Occur? (Univ. Michigan, 2011) https://web.eecs.umich.edu/~zmao/Papers/china-censorship-pam11.pdf

Filtering techniques used

- IP filtering
  - IP addresses get put on a blacklist and get routed to a null route via BGP
  - but difficult to maintain list of IPs, danger of leaking the null route to neighbors, and possible “overblocking” (several websites hosted on the same IP for example)
- DNS poisoning/hijacking
  - filtering routers sniff for keywords in DNS type A queries, then spoof a reply from a supposedly authoritative DNS server but give a fake IP address
  - but can cause important collateral damage for transiting traffic
- TCP packet inspection / TCP reset
  - main method, more precise with less side-effects → discussed next

TCP reset: how does it work?

- Routers have attached an Intrusion Detection System (IDS) with a list of keywords to look for in both incoming and outgoing TCP packets
  - e.g. `GET /?falun` will trigger the IDS
- The router will then forge TCP reset packets (i.e. with the RST flag set) and send them to both endpoints so that they both stop sending packets to each other
- Subsequently, for a certain amount of time any attempt to set up a new connection between the two same endpoints will trigger the RST packets even in the absence of any keywords

- Clayton, R., Murdoch, S., Watson, R. - Ignoring the Great Firewall of China (Univ. Cambridge, 2006) https://www.cl.cam.ac.uk/~rnc1/ignoring.pdf

TCP reset: limitations

- Router and IDS capacities: need to store the state of the endpoints and, have to send forged RST packets before the endpoints can exchange real packets
  - →limit blocking time: a few minutes to 1 hour, 20 minutes on average
  - →inspect only part of the traffic (2/3 approximately)

- Risk of DoS attacks: i.e. an attacker wanting to prevent a Chinese embassy to access some official site in China can spoof its IP and trigger the Firewall
  - →blocking of subsequent connections between the end points only happens if closely related port numbers are used

- Clayton, R., Murdoch, S., Watson, R. - Ignoring the Great Firewall of China (Univ. Cambridge, 2006) https://www.cl.cam.ac.uk/~rnc1/ignoring.pdf

TCP reset: circumvention

- This type of filtering does rely on end points implementing well-behaved TCP behavior: what happens if they don't abort on RST ?
- The reset packets are additional, and legitimate packets actually go through the Firewall unchanged, so a normal connection can be maintained simply by ignoring the reset packets.
- Although it is somewhat unclear whether an user in China can justify this as not being an attempt to evade the Firewall

- Clayton, R., Murdoch, S., Watson, R. - Ignoring the Great Firewall of China (Univ. Cambridge, 2006) https://www.cl.cam.ac.uk/~rnc1/ignoring.pdf

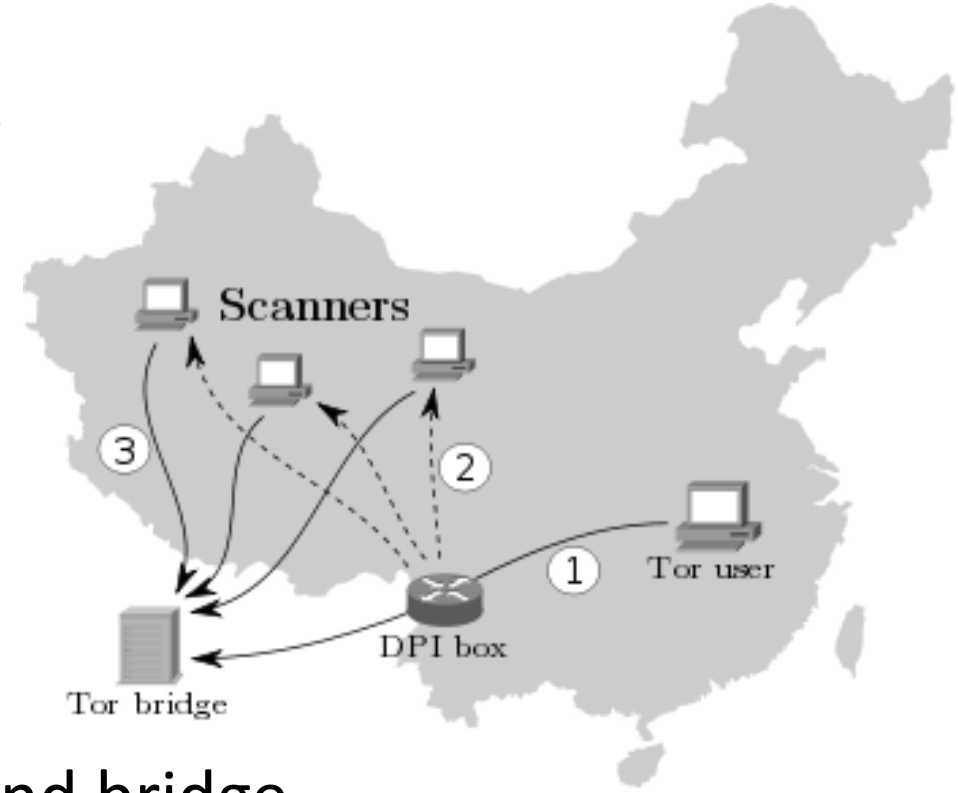
Blocking Tor: past efforts

- IP blacklisting of directory authorities → circumvented using bridges, i.e. unpublished Tor relays allowing censored users to access the Tor network
- HTTP header filtering → circumvented using HTTPS

- But since October 2011: unpublished Tor bridges were reported to get blocked after only a few minutes → How did the Firewall find them?

- Winter P., Lindskog S. - How the Great Firewall of China is Blocking Tor (Karlstad Univ., 2012) https://www.usenix.org/system/files/conference/foci12/foci12-final2.pdf

Blocking Tor: DPI & scanners



- The Firewall uses Deep Packet Inspection (DPI) to detect Tor packets.
- DPI identifies the cipher list contained in the TLS handshake particular to Tor
- The IP:port destination is then relayed to distributed scanner machines that try to establish Tor connections to the newly found bridge.
- A successful connection results in that IP:port pair getting blocked.

- Winter P., Lindskog S. - How the Great Firewall of China is Blocking Tor (Karlstad Univ., 2012) https://www.usenix.org/system/files/conference/foci12/foci12-final2.pdf

Blocking Tor: evading the DPI detection

- Filtering scanner machines: difficult because they are too hard to tell apart from legitimate Tor users
- Some software has been developed to obfuscate the traffic so that the TLS cipher list cannot be identified anymore, but needs to run on both sides
- Authors propose a server side tool that rewrites the TCP window size announced by the bridge during the TCP handshake, forcing the client to split its cipher list across two TCP segments, which seems to be sufficient to bypass the DPI detection (apparently doesn't perform packet reassembly)

- Winter P., Lindskog S. - How the Great Firewall of China is Blocking Tor (Karlstad Univ., 2012) https://www.usenix.org/system/files/conference/foci12/foci12-final2.pdf