How the Great Firewall of China is Blocking Tor

Philipp Winter and Stefan Lindskog
Karlstad University

Aug. 6, 2012
In a nutshell

1. Investigated how Tor is being **blocked**

2. Speculated about the blocking **infrastructure**

3. Looked at **countermeasures**

Significant prior work done by Tim Wilde from Team Cymru!
What Tim found out

Tor TLS Client

Hello

Tor bridge
Experimental setup

- **China**
  - VPS (full root access)
  - Found 32 open SOCKS proxies via Google
  - PlanetLab

- **Russia**
  - Middle relay

- **Singapore**
  - Bridge in Amazon EC2 cloud

- **Sweden**
  - Several bridges
Meet Alice!
Alice wants to use Tor!
HTTP mostly does not work
But HTTPS is fine!
Now, Alice needs the consensus.

7/8 directory authorities were blocked.
SYN/ACK from relays and bridges swallowed
Most public relays in consensus blocked

- Downloaded consensus containing 2819 relays at the time
- Could establish TCP connection to only 1.6% of all relays
- After three days: Only one of them still reachable
Where does the fingerprinting happen?

- Outside China
- Open SOCKS
- PlanetLab
- VPS
- Tor TLS
- Client Hello
Bridges can be unblocked!

- Made GFC block 2 private bridges:
  - **1st bridge**: Blocked Chinese address space but whitelisted VPS in China
  - **2nd bridge**: Unmodified
- After \(\sim 12\) hours: First bridge became reachable again
So what about the scanners?
We now have our data!

- After 2.5 weeks: 3295 scans!

- Have a look yourself:
  http://www.cs.kau.se/philwint/static/gfc/
When are the scanners connecting?
There is a daily pattern!
Where are the scanners coming from?

▶ 50% from 202.108.181.70.

▶ 50% from random IP addresses.

▶ All IP addresses part of AS\{4837, 4134, 17622\}.
What about 202.108.181.70?

netname: BJ-GD-TECH-C0
descr: Beijing Guanda Technology Co.Ltd
country: CN
admin-c: CH455-AP
tech-c: SY21-AP
mnt-by: MAINT-CNCGROUP-BJ
changed: suny@publicf.bta.net.cn 20020524
status: ASSIGNED NON-PORTABLE
source: APNIC

[...]
IP spoofing?

- **No** communication with scanners possible
- Sometimes, several minutes after scan, host starts **replying** to pings
- **Suspicious**: TTL differs!
- **Conjecture**: GFC is spoofing random IP addresses for scanning
So how can we help Alice?
Two dimensions to the problem

Censorship devices can identify Tor by:

1. **Protocol** — "the TLS client hello looks like Tor!"
2. **Destination** — "that guy is connecting to a bridge!"

China is currently breaking both dimensions.
Protocol obfuscation

- Makes it hard to break the first dimension of the problem

- Most censorship devices recognize Tor by looking at the TLS client/server hello

- Solution: Wildly obfuscate the entire protocol or make it look like smth. else

- [https://www.torproject.org/docs/pluggable-transports](https://www.torproject.org/docs/pluggable-transports)
Packet fragmentation

- Experiments with fragroute showed that the GFC does no packet reassembly

- Developed small tool for server-side packet fragmentation
  https://github.com/NullHypothesis/brdgrd

- Transparently rewrites first announced TCP window size

- Makes Tor client split its cipher list into two parts
It's looking better for us

- Flash proxies to tackle bridge distribution problem (Fifield et al., PETS’12)

- Many pluggable transports (SkypeMorph, Stegotorus, ...)

- https://bridges.torproject.org asks for CAPTCHA now
Thanks to

- Anonymous reviewers
- Tor developers
- Fabio Pietrosanti
- Simone Fischer-Hübner
- Rose-Mharie Åhlfeldt
- Harald Lampesberger

Contact: philipp.winter@kau.se (4096R/2D081E16)

Code/Data/Paper:  
http://www.cs.kau.se/philwint/static/gfc/