Analyzing China’s Blocking of Unpublished Tor Bridges

Arun Dunna, Ciarán O’Brien, Phillipa Gill
University of Massachusetts Amherst
The Digital Arms Race
The Digital Arms Race

Tor and the Censorship Arms Race: Lessons Learned
Russia, China vow to kill off VPNs, Tor browser

New laws needed because today's censorship not good enough, apparently
Russia, China

In China, an entire generation is growing up with censored internet
The Digital Arms Race

THE ATTACK ON GLOBAL PRIVACY LEAVES FEW PLACES TO TURN up with censored internet
Internet activists are finding ways around China’s Great Firewall.
Prior Work

<table>
<thead>
<tr>
<th></th>
<th>2012</th>
<th>2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Block Method</td>
<td>(IP, Port)</td>
<td>(IP, Port)</td>
</tr>
<tr>
<td>Block Duration</td>
<td>12 hours</td>
<td>12 hours</td>
</tr>
<tr>
<td>Scanning Queues</td>
<td>15 minutes</td>
<td>Instantly</td>
</tr>
<tr>
<td>Scanner Distribution</td>
<td>Non-Uniform</td>
<td>Uniform</td>
</tr>
<tr>
<td>Most Common Scanner</td>
<td>202.108.181.70</td>
<td>202.108.181.70</td>
</tr>
<tr>
<td>Common Scanner ASes</td>
<td>4134, 4837, 17622</td>
<td>4134, 4837, 7497</td>
</tr>
</tbody>
</table>

- 2012 FOCI – P. Winter and S. Lindskog
- 2015 IMC – Ensafi et al.
- Now 2018 – what’s changed?
Setup
Setup

Tor Bridge
Setup

Tor Bridge | Tor Relay | Tor Client
Public Network Reachability

Tor Client

Consensus
Public Network Reachability

- Connect to consensus relays
Public Network Reachability

- Connect to consensus relays  **Blocked**

Tor Client  

Consensus
Public Network Reachability

- Connect to consensus relays **Blocked**
- Ping public relays from client

Tor Client ➔ Consensus
Public Network Reachability

- Connect to consensus relays  **Blocked**
- Ping public relays from client  **Blocked**
Public Network Reachability

- Connect to consensus relays **Blocked**
- Ping public relays from client **Blocked**
Public Network Reachability

- Connect to consensus relays **Blocked**
- Ping public relays from client **Blocked**
- Published relays blocked after 10 minutes
Public Network Reachability

- Connect to consensus relays **Blocked**
- Ping public relays from client **Blocked**
- Published relays blocked after 10 minutes
- Blacklist purged after 12 hrs if Tor is stopped

Tor Client ----> Public Relay

PING EVERY 5 SECONDS
Relay (Bridge) Blocking
Relay (Bridge) Blocking
Relay (Bridge) Blocking
Relay (Bridge) Blocking
Relay (Bridge) Blocking
Relay (Bridge) Blocking

- Initial connection attempt blocked – RST
Relay (Bridge) Blocking

- Initial connection attempt blocked – RST
Relay (Bridge) Blocking

- Initial connection attempt blocked – RST
- Blocks relays at IP level (not port like 2012, 2015) – unidirectional (bridge to client blocked)
Blocking Duration
Blocking Duration

RUNNING TOR
Blocking Duration

NOT RUNNING TOR
Blocking Duration

NOT RUNNING TOR
Blocking Duration

- Blocked for 12 hours, scanned in 12 hour intervals
Blocking Duration

- Blocked for 12 hours, scanned in 12 hour intervals
- GFW scans every 12 hours and IP blacklists – resource conservation?

NOT RUNNING TOR
Fingerprinting - Setup
Fingerprinting - Setup

- Blocks IP, not ports, so cannot use one IP with multiple Tor services across multiple ports
- A method to attract scanners without getting blocked?
Fingerprinting - Setup

- Blocks IP, not ports, so cannot use one IP with multiple Tor services across multiple ports
- A method to attract scanners without getting blocked?
Fingerprinting - Setup

- Blocks IP, not ports, so cannot use one IP with multiple Tor services across multiple ports
- A method to attract scanners without getting blocked?
  - Drop scanner packets – “we’re not a relay!”
  - Over 44 hours... collected 934 unique IPs
## Fingerprinting - Comparison

- Comparison of captured TCP SYN packets

<table>
<thead>
<tr>
<th></th>
<th>2015</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTL Range</td>
<td>46 - 51</td>
<td>48 - 50</td>
</tr>
<tr>
<td>MSS Values</td>
<td>1400, 1460</td>
<td>1368, 1400</td>
</tr>
<tr>
<td>Window Scaling</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Permit Selective ACKs</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>TCP Timestamp</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>No Operation Flag</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
## Fingerprinting - Comparison

- Comparison of captured TCP SYN packets

<table>
<thead>
<tr>
<th></th>
<th>2015</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTL Range</td>
<td>46 - 51</td>
<td>48 - 50</td>
</tr>
<tr>
<td>MSS Values</td>
<td>1400, 1460</td>
<td><strong>1368, 1400</strong></td>
</tr>
<tr>
<td>Window Scaling</td>
<td><strong>7</strong></td>
<td><strong>7</strong></td>
</tr>
<tr>
<td>Permit Selective ACKs</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>TCP Timestamp</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>No Operation Flag</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

- 111.202.242.93 is origin of 5% of SYN packets
- All of them have MSS of 1368, and we respond to them normally, but we’re not blocked - different scanner?
Fingerprinting - Infrastructure

- All scanners in China, geolocation using GeoLite2
Fingerprinting - IPs

- Out of 934 IPs seen, 908 IPs conducted one scan, and 26 conducted two scans
- Each IP does one/two scans – avoiding detection?
Fingerprinting – IPID & TTL

- No trend in IPIDs observed… but similarity in TTL indicate colocation
- Cannot determine whether RST injectors are collocated with scanners (RSTs TTL 44, vs. 48-50)
Circumvention

- Pluggable Transports!
Circumvention

- Pluggable Transports!
  - Obfs4: obfuscation, not recommended due to bridge distribution
Circumvention

- Pluggable Transports!
  - Obfs4: obfuscation, not recommended due to bridge distribution
    - Meek: uses domain fronting, recommended for Chinese users… but how long?
Amazon blocks domain fronting, threatens to shut down Signal’s account
Google disables “domain fronting” capability used to evade censors

A "long-planned" change happens to coincide with a new wave of state censorship in Russia.
Signal says Amazon, Google will no longer help it evade censorship

Hot on the heels of Google's move last week, Amazon Web Services said it will also switch off functionality for domain fronting.

A "long-planned" change happens to coincide with a new wave of state censorship in Russia.
Signal says Amazon, Google will no longer help it evade censorship

Amazon Web Services starts blocking domain-fronting, following Google’s lead
Is there a better way?
Rejecting Scanners

- Bridge relays drop packets from GFW scanners – “I’m not running Tor!”
Rejecting Scanners

- Bridge relays drop packets from GFW scanners – “I’m not running Tor!”
Rejecting Scanners

- Bridge relays drop packets from GFW scanners – “I’m not running Tor!”
Rejecting Scanners

- Bridge relays drop packets from GFW scanners – “I’m not running Tor!”
- **EX:** `iptables -A INPUT -p tcp --destination-port 7412 -m tcpmss --mss 1400 -j DROP`

![Graph showing the impact of different MSS values on packet distribution](image)
Conclusion

- GFW now blocks by IP instead of (IP, Port), which makes measuring more challenging
Conclusion

- GFW now blocks by IP instead of (IP, Port), which makes measuring more challenging
- Dropping scanner traffic results in bridge not getting blocked - how long until GFW responds?
Conclusion

- GFW now blocks by IP instead of (IP, Port), which makes measuring more challenging
- Dropping scanner traffic results in bridge not getting blocked - how long until GFW responds?
- Censor behavior changes over time, so it’s important to continue revisiting results
Conclusion

- GFW now blocks by IP instead of (IP, Port), which makes measuring more challenging
- Dropping scanner traffic results in bridge not getting blocked - how long until GFW responds?
- Censor behavior changes over time, so it’s important to continue revisiting results
Conclusion

- GFW now blocks by IP instead of (IP, Port), which makes measuring more challenging
- Dropping scanner traffic results in bridge not getting blocked - how long until GFW responds?
- Censor behavior changes over time, so it’s important to continue revisiting results
Thanks!

Arun Dunna – CS MS Student
adunna@cs.umass.edu

Calipr Lab, with Prof. Phillipa Gill
https://calipr.cs.umass.edu

We open source all of our datasets and code, available at https://calipr.cs.umass.edu, to maintain reproducibility.
Q/A
References

- Images
  - https://github.com/hjnilsson/country-flags
  - http://www.opensecurityarchitecture.org/cms/library/icon-library
  - https://www.flaticon.com/authors/smashicons