Internet censorship

Tom Ristenpart
CS 6431
Current estimates of Internet censorship

OpenNet Initiative (ONI), Reporters Without Borders

"Internet Censorship and Surveillance World Map" by Jeffrey Ogden (W163)

Magenta-colored countries are countries with pervasive censorship and surveillance of the Internet.
How would you censor web requests?

- IP filtering
- DNS filtering / redirection
- URL filtering
- Packet filtering (search keywords in TCP packets)
- Protocol filtering (detect Tor protocol)
Golden Shield Project

“If you open the window for fresh air, you have to expect some flies to blow in” – Deng Xiaoping in 1980s

Great Firewall of China:
• IP filtering
• DNS filtering / redirection
• URL filtering
• Packet filtering (search keywords in TCP packets)
• Protocol filtering
• Active probing of suspect destination IP addresses
Islamic Republic of Iran

• Every ISP must run “content-control software”
  – SmartFilter (up until 2009) made by USA company
  – Nokia Siemens deep-packet inspection (DPI) systems

• According to wikipedia: 50% of top 500 most popular websites blocked in Iran

• Occasional widespread filtering of Tor, TLS, other encrypted protocols
Censorship as two-step process

1. *Sensitive content identification*
   - DNS and IP blacklists
   - Keyword blacklists with DPI
   - Protocol identification (e.g., TLS)
   - Tool identification (e.g., Tor)

2. *Censoring action*
   - DNS poisoning
   - HTTP man-in-the-middle
   - TCP resets
   - Dropping packets
Types of packet inspection

Internet service providers need only look at IP headers. Deep packet inspection (DPI) analyzes application headers and data.
DPI technology

• From Narus’ website (http://narus.com/index.php/product/narusinsight-intercept):
  – “Target by phone number, URI, email account, user name, keyword, protocol, application and more”, “Service- and network agnostic”, “IPV 6 ready”
  – Collects at wire speeds beyond 10 Gbps

• Narus allegedly used by NSA in San Francisco AT&T office
NarusInsight™ Selected To Save Pakistan's Telecommunications Networks Millions Of Dollars Per Year

NarusInsight™ Selected to Save Pakistan's Telecommunications Networks Millions of Dollars Per Year

Narus System Chosen to Detect Rogue VoIP Traffic

MOUNTAIN VIEW, Calif.—September 21, 2007—Narus, Inc., the leader in carrier-class security for the world's largest IP networks, today announced that the company has teamed up with Inbox Business Technologies Pvt. Ltd, a leading total IT solution provider in Pakistan, to keep Pakistan's telecommunication networks clear of illegal, rogue and malicious IP traffic. NarusInsight was chosen by the Pakistan Telecommunication Authority (PTA) (the government administration responsible for regulating the establishment, operation and maintenance of telecommunication systems, and the provision of telecom services) to detect rogue VoIP traffic flowing through the telecommunications network in Pakistan.

How do we know about censorship?

• Anecdotes from people within censored regions
• More formal surveys
• Network measurements:
  – Web sites aggregating info such as GreatFire
  – Herdict tool (browser plugin to manually report blcokage)
  – Open Observatory of Network Interference (opt-in measurements of network connections)
  – Encore paper
Encore web-based measurements

• Burnett-Feamster study
• Embed in other non-blocked web pages measurement functionality of suspect censorship targets
• Use cross-site embeddings and browser side-channels
1. Origin serves page to client containing measurement task
2. Client renders page and executes measurement task
3. Task issues a cross-origin request for a resource on measurement target
4. Censor may filter request or response
Measure Web censorship

By adding a single line of code to your Web site, visitors of your site will automatically contribute data about how they experience Web censorship:

<iframe src="//encore.noise.gatech.edu/task.html" width="0" height="0" style="display: none"></iframe>

Learn more about Encore  
Read the SIGCOMM 2015 paper  
Encore settings
The results:

Confirmed blocking of:

- youtube.com in **Pakistan, Iran, China**
- Twitter.com, facebook.com in **China, Iran**
Statement from the SIGCOMM 2015 Program Committee: The SIGCOMM 2015 PC appreciated the technical contributions made in this paper, but found the paper controversial because some of the experiments the authors conducted raise ethical concerns. The controversy arose in large part because the networking research community does not yet have widely accepted guidelines or rules for the ethics of experiments that measure online censorship. In accordance with the published submission guidelines for SIGCOMM 2015, had the authors not engaged with their Institutional Review Boards (IRBs) or had their IRBs determined that their research was unethical, the PC would have rejected the paper without review. But the authors did engage with their IRBs, which did not flag the research as unethical. The PC hopes that discussion of the ethical concerns these experiments raise will advance the development of ethical guidelines in this area. It is the PC’s view that future guidelines should include as a core principle that researchers should not engage in experiments that subject users to an appreciable risk of substantial harm absent informed consent. The PC endorses neither the use of the experimental techniques this paper describes nor the experiments the authors conducted.
How would you **avoid** censorship?

- IP filtering
- DNS filtering / redirection
- URL filtering
- Packet filtering (search keywords in TCP packets)
- Protocol filtering (detect Tor protocol)
Preventing censorship

• End-to-end encryption (HTTPS, SSH)

• What does this protect?
• What does it leak?
Tor
(The Onion Router)

Other major backbone

Tor Node
7.8.9.1

Tor Node
9.1.1.2

Tor Node
8.9.1.1

Censorship equipment

IP: 1.2.3.4

IP: 5.6.7.8

National internet

Other major backbone
Directly connecting users from the Islamic Republic of Iran

The Tor Project - https://metrics.torproject.org/
Iran DPI blocking of Tor

- Tor point-to-point connections use TLS

<table>
<thead>
<tr>
<th>No.</th>
<th>Time</th>
<th>Source</th>
<th>Destination</th>
<th>Protocol</th>
<th>Length</th>
<th>Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>359</td>
<td>2.138821000</td>
<td>128.105.35.160</td>
<td>173.194.46.114</td>
<td>TLSv1.2</td>
<td>583</td>
<td>Client Hello</td>
</tr>
<tr>
<td>362</td>
<td>2.140902000</td>
<td>128.105.35.160</td>
<td>173.194.46.122</td>
<td>TLSv1.2</td>
<td>290</td>
<td>Client Hello</td>
</tr>
<tr>
<td>369</td>
<td>2.154594000</td>
<td>128.105.35.160</td>
<td>173.194.121.33</td>
<td>TLSv1.2</td>
<td>285</td>
<td>Client Hello</td>
</tr>
<tr>
<td>371</td>
<td>2.155001000</td>
<td>128.105.35.160</td>
<td>173.194.121.42</td>
<td>TLSv1.2</td>
<td>291</td>
<td>Client Hello</td>
</tr>
</tbody>
</table>

Censorship equipment

IP: 1.2.3.4

National internet

Tor Node 7.8.9.1
TLS Handshake

Client

Pick random Nc

Check CERT using CA public verification key

Pick random PMS

C <- E(pk, PMS)

Bracket notation means contents encrypted

Server

Pick random Ns

CERT = (pk of server, signature over it)

ChangeCipherSpec,
{ Finished, PRF(MS, “Client finished” || H(transcript)) }

ChangeCipherSpec,
{ Finished, PRF(MS, “Server finished” || H(transcript’)) }

MS <- PRF(PMS, “master secret” || Nc || Ns )
Iran DPI blocking of Tor

• Tor point-to-point connections use TLS
• Use DPI to filter Tor connections:
  – Tor certificates have short expiration date
  – Most websites have long expiration date
  – Shut down those connections with short expiration dates
  – https://blog.torproject.org/blog/update-internet-censorship-iran
• Tor fixed via longer expiration dates
• Later in 2012: blocking/degrading all TLS connections
Great Firewall targeting of Tor (circa 2011 and before)

• Enumerate Tor relays and filter them
Number of directory requests to directory mirror trusted

https://torproject.org

China
Chinese Tor users via bridges
Great Firewall targeting of Tor (circa 2011 – today)

Admin noticed weird connections from China 2011

TLS connections with particular ciphersuites flagged for active probing

If remote server speaks Tor then add its IP address to blacklist

From [Winter, Lindskog 2012]

https://gist.github.com/da3c7a9af01d74cd7de7
**TLS Handshake**

1. **Client**
   - Pick random Nc
   - Check CERT using CA public verification key
   - Pick random PMS C <- E(pk, PMS)

2. **Server**
   - Pick random Ns
   - CERT = (pk of server, signature over it)

3. **Client**
   - ClientHello, MaxVer, Nc, Ciphers/CompMethods
   - ServerHello, Ver, Ns, SessionID, Cipher/CompMethod

4. **Change Cipher Spec**
   - C
     - ChangeCipherSpec, { Finished, PRF(MS, “Client finished” || H(transcript)) }
     - ChangeCipherSpec, { Finished, PRF(MS, “Server finished” || H(transcript’)) }

5. **MS**
   - MS <- PRF(PMS, “master secret” || Nc || Ns )

Bracket notation means contents encrypted
Great Firewall targeting of Tor (circa 2011 – today)

Admin noticed weird connections from China 2011

TLS connections with particular ciphersuites flagged for active probing

If remote server speaks Tor then add its IP address to blacklist

From [Winter, Lindskog 2012]

https://gist.github.com/da3c7a9af01d74cd7de7
How would you defeat DPI-based tool / protocol identification?

Make hard to detect as Tor:
• randomizers (obfsproxy)
• protocol mimicry
• tunneling

Mimicry & tunneling related to steganography
Great Firewall targeting of Tor (circa 2011 – today)

Ensafi et al. IMC 2015 follow-up study

• Active measurements, log file analysis, etc.
• China is checking obfsproxy3 bridges
• Hijack IP addresses to perform active probing
• DPI is stateful but does not reconstruct TCP streams
Directly connecting users from Egypt

The Tor Project - https://metrics.torproject.org/
From BlueCoat:

- Our awareness of the presence of these ProxySG appliances in Syria came from reviewing online posts made by so-called “hacktivists” that contained logs of internet usage which appear to be generated by ProxySG appliances. We believe that these logs were obtained by hacking into one or more unsecured third-party servers where the log files were exported and stored. We have verified that the logs likely were generated by ProxySG appliances and that these appliances have IP addresses generally assigned to Syria. We do not know who is using the appliances or exactly how they are being used. We currently are conducting an internal review and also are working directly with appropriate government agencies to provide information on this unlawful diversion.
Directly connecting users from the Syrian Arab Republic

The Tor Project - https://metrics.torproject.org/
"Twitter, mwitter!"

Directly connecting users from Turkey

The Tor Project - https://metrics.torproject.org/