Measuring and Circumventing Internet Censorship and Control

Nick Feamster
Georgia Tech

http://www.cc.gatech.edu/~feamster/

Joint work with Sam Burnett, Santosh Vempala, Sathya Gunasekaran, Crisitan Lumezanu, Hans Klein, Wenke Lee, Phillipa Gill, and others
Internet Censorship is Widespread

- Practiced in **59 countries around the world**
  - Many western countries
  - Several electoral democracies (e.g., S. Korea, Turkey) have significant censorship
    - YouTube blocked in Turkey for two years
    - Many North Korean sites blocked in South Korea

- Twelve countries have **centralized infrastructure for monitoring/blocking**

Source: Open Network Initiative
Why do countries censor?

• Political stability

• National security

• Social values
Trend: Increasing Number of Users in Non-Western Regions

Internet Users by Region

* Source: International Telecommunications Union
Examples of Recent Trends

- In 23 countries, a blogger or Internet user was arrested for content posted online
  - Chinese woman sent to labor camp for satirical Twitter message
  - Indonesian woman fined for an email complaining about a local hospital

- Twelve countries instituted bans on Twitter, YouTube or some other online social media service
Conventional Internet Censorship

Censor

Alice

Punish User

Firewall

Censor

Block Traffic

Censored net

Uncensored net

Bob
Technical Enforcement: Blocking

• ISP acts on instructions from a judge, government official, etc.
  – Filtering: IP address, DNS
  – Keyword-based: search for keyword in URL
    • China, Iran, Tunisia have such systems in place

• Common: Use of centralized infrastructure (e.g., routing)

Source: Renesys
Questions

• **How widespread** is Internet censorship?
• How do countries **enforce** censorship?
  – How does it evolve over time?
  – Does it coincide with other events?

• How can citizens **circumvent** it?

• How (else) might a government (or organization) exercise **control** over its citizens?
Outline

• **Measuring** censorship
  – Censorship is widespread, but the extent and evolution of practices are unknown

• **Circumventing** censorship
  – Deniability is a key challenge
  – Bootstrapping remains significant open problem

• **Combating** manipulation
  – Analysis of Twitter behavior of propagandists
  – Measurement and illustration of filter bubbles
Monitoring Censorship

- **Herdict**: Crowdsourcing reports of Internet censorship
- **Google Transparency Report**: Monitor reachability of online services
Monitoring Censorship: Challenges

• “Censorship” is **ill-defined**
  – Personalization may be confused with censorship
  – Performance problems may be confused with censorship

• Measurement tools **can be blocked**
  – Measurements may be blocked
  – Reports may be blocked

• Measurements **tough to characterize**
  – Reports may be falsified

• Running the tool **may be incriminating**
Problems with Current Approaches

- Biased by what users choose to report
- Lack of corroborating, open measurements
- Not general (focused only on limited services)
- Not longitudinal
- Do not cover a set of ISPs or access modes within a country
- Do not run on a diversity of hardware
Design Requirements

• **Easy to install and use:** Should be easy to install and run on a variety of platforms.

• **Cross-platform:** Tests should be write once, run anywhere.

• **Flexible:** Should be capable of implementing a wide variety of experiments, including many from the test specifications from existing projects (e.g., OONI).

• **Secure:** Arbitrary remote code execution is bad.

• **Extensible:** Should be capable of incorporating new experiments.
Censorscope: Design Overview

https://github.com/projectbismark/censorscope

- User installs base software and registers with server
- Server periodically pushes upgrades
- Client sends properties
- Client downloads measurement script, written in a Lua-based DSL
- Client returns measurement results
Target Platforms

Exploit Existing Deployments

- **BISmark**: Home routers
  - 200+ home routers deployed in 20+ countries

- **Android**: Mobile devices (MySpeedTest)
  - 5,000 installations in 30+ countries

Expand to New Deployments

- **Linux/MAC OS X**: End hosts
- **Fathom**: Browsers
Tests: Planned and In-Progress

- DNS lookups
- TCP connectivity
- HTTP requests
- DNS spoofing
- DNS tampering
- HTTP host tampering
- Bridget
- Block page detection
- Web performance measurement

Seeking help developing tests for a variety of platforms.
Outline

• **Measuring** censorship
  – Censorship is widespread, but the extent and evolution of practices are unknown

• **Circumventing** censorship
  – Deniability is a key challenge
  – Bootstraping remains significant open problem

• **Combating** manipulation
  – Analysis of Twitter behavior of propagandists
  – Measurement and illustration of filter bubbles
General Approach: Use a Helper

The helper sends messages to and from blocked hosts on your behalf.
Circumvention Systems

- Anonymous routing systems
- Community wireless networks
- Distributed services
Significant Challenge: Deniability

- Easy to hide what you are getting
  - E.g., just use SSL or some other confidential channel
- And sometimes easy to “get through” censors
  - Reflection (e.g., Tor)
- But hard to hide that you are doing it!
Design Principles

• **Redundancy and hiding** to thwart disruption
  – Erasure coding, steganography
    (from coding, message hiding)

• **Disguise content retrieval** as innocuous activity
  – Distributed hash table lookup
    (from distributed systems)

• **Decouple** sending and receiving of messages
  – User-generated content sites as drop sites
    (from the “real world”)
Collage: Let User-Generated Content Help Defeat Censorship

- Robust by using redundancy
- Users generate innocuous-looking traffic
- No dedicated infrastructure required

Collage in Detail

Collage steps:
1. Obtain message
2. Pick message identifier
3. Obtain cover media
4. Embed message in cover
5. Upload UGC to content host
6. Find and download UGC
7. Decode message from UGC

Step 1: Digital imageUid
- Application specific
- Only intended recipient should know it
Collage: Challenges

• Determining **how to embed** the message
  – Discovery should be difficult
  – Disruption should be difficult

• Agreeing on **where to embed** the message
  – Alice and Bob must agree on a message identifier

• Designing the process to be **deniable**
  – Alice’s process of retrieval should look “normal”
How to Embed the Message

- **Encrypt** the message using the identifier
- Generate chunks using **erasure coding**
  - Generate many chunks, recover from any $k$-subset
  - Allows splitting among many vectors, robustness
- **Embed** chunks into vectors

Steganography: hard to detect
Watermarking: hard to remove

Collage steps:
1. Obtain message
2. Pick message identifier
3. Obtain cover media
4. **Embed message in cover**
5. Upload UGC to content host
6. Find and download UGC
7. Decode message from UGC
Where to Embed the Message

- Crawling all of Flickr is not an option
- Must agree on a subset of content on user-generated content sites without any immediate communication

Collage steps:
1. Obtain message
2. Pick message identifier
3. Obtain cover media
4. Embed message in cover
5. **Upload UGC to content host**
6. Find and download UGC
7. Decode message from UGC

**Solution:** A predictable way of mapping message identifiers to subsets of content hosts.
Making the Embedding Deniable

Tasks

1. Hash the identifier
2. Hash the tasks
3. Map identifier to closest tasks

Receivers perform these tasks to get vectors
- Senders publish vectors so that when receivers perform tasks, they get the sender’s vectors

Collage steps:
1. Obtain message
2. Pick message identifier
3. Obtain cover media
4. Embed message in cover
5. Upload UGC to content host
6. Find and download UGC
7. Decode message from UGC

Message Identifier

http://nytimes.com

Search for blue flowers on Flickr

Look at JohnDoe’s videos on YouTube
## Feasibility Case Study

<table>
<thead>
<tr>
<th></th>
<th>News Articles</th>
<th>Covert Tweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content host</td>
<td>Flickr</td>
<td>Twitter</td>
</tr>
<tr>
<td>Message size</td>
<td>30 KB</td>
<td>140 Bytes</td>
</tr>
<tr>
<td>Vectors needed</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>Storage needed</td>
<td>600 KB</td>
<td>4 KB</td>
</tr>
<tr>
<td>Sending traffic</td>
<td>1,200 KB</td>
<td>1,100 KB</td>
</tr>
<tr>
<td>Sending time</td>
<td>5 minutes</td>
<td>60 minutes</td>
</tr>
<tr>
<td>Receiving traffic</td>
<td>6,000 KB</td>
<td>600 KB</td>
</tr>
<tr>
<td>Receiving time</td>
<td>2 minutes</td>
<td>½ minute</td>
</tr>
</tbody>
</table>

Experiments performed on a 768/128 Kbps DSL connection
Ongoing Work: New Tor “Pluggable Transports”

- Collage and Infranet: Slow performance
  - ...and strong adversary model
- What about an adversary that can examine but has limited storage capability?
Outline

• **Measuring** censorship
  – Censorship is widespread, but the extent and evolution of practices are unknown

• **Circumventing** censorship
  – Deniability is a key challenge
  – Bootstrapping remains significant open problem

• **Combating** manipulation
  – Analysis of Twitter behavior of **propagandists**
  – Measurement and illustration of **filter bubbles**
Manipulation and Propaganda

- **Sock-puppeting**: False appearance of independent speakers

- **Astroturfing**: False appearance of a grassroots movement
Detecting Propaganda

• How can Twitter be used to affect public opinion?
• Can we detect when Twitter is being used to spread propaganda?

Nevada Senate Race

Debt Ceiling Debate
Four Properties of Propagandists

• Higher fraction of retweets
• More bursty tweeting volumes
• Higher daily volumes
• Quick retweeting

bias: Measuring the Tweeting Behavior of Propagandists by Cristian Lumezanu, Nick Feamster, and Hans Klein. In the Sixth International AAAI Conference on Weblogs and Social Media (ICWSM), 2012.
Personalization as “Filter Bubble”

- Online personalization is creating situations where we only see things that already suit our own tastes.
- Personalization can also be exploited.
- **Goal:** “Burst the filter bubble.” Show the user information that might otherwise be hidden.

“A squirrel dying in front of your house may be more relevant to your interests right now than people dying in Africa”
– Mark Zuckerberg
Bobble: Bursting the Filter Bubble

- Execute query
  - As different users
  - From different vantage points
  - With different history (e.g., cookies)

- Compare differences in results
  - What shows up on the first page?
  - Where does it show up?
  - When it doesn’t appear, what are the possible explanations?
Bursting the Filter Bubble

Scott Shenker | EECS at UC Berkeley
www.eecs.berkeley.edu/Faculty/Homepages/shenker.html


Scott Shenker | ICSI
www.icsi.berkeley.edu/icsi/about/board/shenker

Leader, New Initiatives Group Chief Scientist shenker. Shenker is a professor in the Electrical Engineering and

Networking Leader Scott Shenker Appointed chief executive officer, who has led ICSI's Networking Group since 1998, will I

Scott Shenker - Research Projects | EECS
www.eecs.berkeley.edu/Research/Projects/.../shenker

Scott Shenker, Vern Paxson and Mark Allman ... And The Networked World: The Impact of the Internet, Ion Stoica

Things you didn’t see!

Scott Shenker - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Scott_Shenker

Scott Shenker is a professor of computer science at UC Berkeley. He is also the leader of the Networking Group and the Chief Scientist of the International ...
Summary

• **Measuring** censorship
  – Extent and evolution of practices are unknown
  – Come help us measure censorship!
  – [https://github.com/projectbismark/censorscope](https://github.com/projectbismark/censorscope)

• **Circumventing** censorship
  – Deniability is a key challenge
  – Covert channels exist (Collage, Infranet)
  – Bootstrapping remains significant open problem

• **Combating** manipulation
  – Analysis of Twitter behavior of propagandists
  – Measurement and illustration of filter bubbles
Other Challenges: Self-Censorship

• Censoring oneself for fear of backlash or retribution

• Occurs in many countries

• Essentially undocumented