FilterMap: Measuring Censorship Filters at Global Scale

Ram Sundara Raman¹, Adrian Stoll¹, Jakub Dalek², Reethika Ramesh¹, Will Scott³, Roya Ensafi¹

University of Michigan¹, The Citizen Lab², Independent³

24 February 2020
Content Filtering Technologies

- Filters, DPIs, middleboxes
- **Dual Use Technology**
  - Intended use - Security
  - Side effect - Censorship, surveillance
- **Commoditization of filters** - High availability, low cost, and advanced features
- Very little, but important, information on use of filters
Netsweeper and Citizen Lab

- **Netsweeper** – Canadian filter vendor - Provides carrier grade filtering, dynamic categorization of websites
- **Citizen Lab** conducted investigations of use of Netsweeper products over several years
- “**Alternative Lifestyles**” category used by UAE, others to block LGBTQ content
- Netsweeper **removed the option** to block category
Canadian Internet Filtering Company Says It's Stopped 'Alternative Lifestyles' Censorship

The UAE was found to be blocking LGBTQ content using a pre-set category in Netsweeper's software. Amid pressure from rights groups, the company says it's disabled that category.

By Jordan Pearson
Jan 21 2019, 12:25pm  Share  Tweet  Snap
Proliferation of Filters
Previous Work

- Biased towards few, well-known filters
- Significant manual effort
  - Physical access
  - In-country collaborators
Blockpages

- Filters respond with blockpages
- Rich with information
  - Trademark of the manufacturing vendor
  - Identity of the deploying actor
- Use blockpages to identify censorship filter deployments
- Identification using blockpages is consistent and scalable
Objectives

Data Collection
Collect many blockpages from filter deployments

Data Analysis
Identify filters from blockpages
Data Collection

Collect the most comprehensive database of filter blockpages
Data Collection

Censorship measurement techniques frequently observe blockpages
Data Collection

Censorship measurement techniques frequently observe blockpages

OONI
Volunteer measurement
https://ooni.org/

Challenges
- Limited scale and ethical constraints
Data Collection

Censorship measurement techniques frequently observe blockpages

Challenges
- Cannot detect filters on common Port 80/443
Data Collection

Censorship measurement techniques frequently observe blockpages

- Novel remote measurement technique
- **Web servers** running on ports 80 and 443
- Idea: Responses from web server when requesting a domain not hosted on the server is predictable

---

OONI
https://ooni.org
Remote measurement

Quack
New remote measurement

Hyperquack
New remote measurement
Hyperquack

46.43.36.222
Hyperquack

46.43.36.222
Hyperquack

Measurement Machine

46.43.36.222
Hyperquack

GET https://www.ndss-symposium.org

46.43.36.222
Hyperquack

GET https://www.ndss-symposium.org

Measurement Machine

46.43.36.222
Hyperquack

GET https://www.usenix.org

Measurement Machine

46.43.36.222
Hyperquack

GET https://www.usenix.org

Moved Permanently

The document has moved here.

Apache/2.4.25 (Debian) Server at www.usenix.org Port 443
Hyperquack

GET https://www.sigsac.org

Measurement Machine

46.43.36.222
Hyperquack

GET https://www.sigsac.org

Moved Permanently

The document has moved here.

Apache/2.4.25 (Debian) Server at www.sigsac.org Port 443
GET https://www.sigsac.org

Moved Permanently

The document has moved [here](https://www.sigsac.org).

*Apache/2.4.25 (Debian) Server at [www.sigsac.org](http://www.sigsac.org) Port 443*
Canonical Templates

<h1>Moved Permanently</h1>
<p>The document has moved</p>
<p><address>Apache/2.4.25 (Debian)
Server at www.signag.org Port 443</address></p>

- Request several bogus but benign domain patterns (<www>.example1298.<com>)
- From the response, remove commonly changing elements e.g. date, domain
- If response for all tests match, save as <strong>canonical template</strong>
Censorship Detection

- Send HTTP(S) GET requests for sensitive keywords
- If response different from canonical template, then there is censorship
- Control tests both before and after to ensure consistency
Censorship Detection

- Send HTTP(S) GET requests for sensitive keywords
- If response different from canonical template, then there is censorship
- Control tests both before and after to ensure consistency
## Censorship Detection

- Send HTTP(S) GET requests for sensitive keywords.
- If response different from canonical template, then there is censorship.
- Control tests both before and after to ensure consistency.

<table>
<thead>
<tr>
<th>Measurement Machine</th>
<th>TCP Handshake</th>
<th>Web Server</th>
</tr>
</thead>
<tbody>
<tr>
<td>Build Canonical template of server response</td>
<td>GET <a href="https://example%7B1,2,3%7D.com">https://example{1,2,3}.com</a></td>
<td>HTTPS reply (e.g., Status Code: 301 Moved)</td>
</tr>
<tr>
<td>GET <a href="https://example%7B1,2,3%7D.com">https://example{1,2,3}.com</a></td>
<td>HTTPS reply (e.g., Status Code: 301 Moved)</td>
<td></td>
</tr>
<tr>
<td>GET <a href="https://blocked.com">https://blocked.com</a></td>
<td>Inject</td>
<td></td>
</tr>
<tr>
<td>GET <a href="https://example%7B1,2,3%7D.com">https://example{1,2,3}.com</a></td>
<td>HTTPS reply (e.g., Status Code: 301 Moved)</td>
<td></td>
</tr>
</tbody>
</table>
Censorship Detection

- Send HTTP(S) GET requests for sensitive keywords
- If response different from canonical template, then there is censorship
- Control tests both before and after to ensure consistency
Hyperquack increases scale to millions of vantage points!

53 million public HTTP hosts

Source - censys.io
Vantage Point Selection

- We use **infrastructural servers** to reduce risk
- **PeeringDB** - list of official websites of Internet service providers
- Use servers hosting the website for measurement ~10,000
Vantage Point Selection

- We use infrastructural servers to reduce risk
- PeeringDB - list of official websites of Internet service providers
- Use servers hosting the website for measurement ~10,000

https://corporate.comcast.com/
Vantage Point Selection

- We use infrastructural servers to reduce risk
- PeeringDB - list of official websites of Internet service providers
- Use servers hosting the website for measurement ~10,000

https://corporate.comcast.com/

23.219.228.121
Ethics

- Followed all the ethical recommendations made in Quack
- Made it clear that we are running measurements on our website
- Rate limit and close connections
- Make only one measurement at a time to a server
- OONI obtains informed consent
Measurements

- **Latitudinal Measurements:**
  - 3 weeks in October 2018
  - HyperQuack - 9,223 VPs
  - Quack - 33,602 VPs
  - 18,736 domains - Citizen Lab Test List
  - Added OONI data

- **Longitudinal Measurements:**
  - HyperQuack and Quack twice a week - November 2018 to January 2019
  - Citizen Lab Global List (~1200 domains) + Alexa Top 1000 domains
Data Analysis

Automate the identification of filters from more than a million disrupted responses
Iterative Classification

- **Insight:** Filters often send the same blockpage regardless of the test domain
- Recursively finds large groups of HTML pages with the same content
- Blockpage clusters are labeled with signatures, a unique subset of the HTML page or header
- Example: `<th>Barracuda NextGen Firewall:</th>`
Image Clustering

- Cluster pages with **dynamic content** - DBSCAN algorithm
- **Tremendously reduce the manual effort** - 1 page in 200 groups
FilterMap enables continuous, sustainable, data-driven view of filter deployment
Results

FilterMap creates a map of filter deployments based on the vantage points measured.
FilterMap Results

- FilterMap found **90 blockpage clusters** (Clusters indicate either vendors or actors)
- Filters are deployed in many locations in **103 countries**
- Filter types found - Commercial products, national firewalls, ISP and organizational deployments
Commercial Filters
Commercial Filters

- 15 commercial filters used in 102 countries
- Sold by companies in the US
- Filters found in **36 out of 48 countries** labelled as “Not Free” or “Partly Free” by Freedom House
- Pornography, gambling, provocative attire and anonymization tools most commonly blocked
FilterMap Results

- 4 National Firewalls - Iran, Saudi Arabia, Bahrain and South Korea
FilterMap Results

- 4 National Firewalls - Iran, Saudi Arabia, Bahrain and South Korea
- Large number of filters in ISPs, especially in Russia
FilterMap Results

● 4 National Firewalls - Iran, Saudi Arabia, Bahrain and South Korea
● Large number of filters in ISPs, especially in Russia
● Of the 90 blockpage clusters -
  ○ 70 - Latitudinal
  ○ 20 additional - Longitudinal
● FilterMap can continuously track filter proliferation
Limitations and Future Work

- Blockpages as a source
  - Future work - Certificate, TCP/IP header
- Evasion - Possible but unlikely
- Exact filter location in network is unknown
Implications

- Unrestricted transfer - Easier to deploy and harder to circumvent
- Million-dollar fines and increased regulation
- FilterMap is maintained as source of longitudinal data
- Accountability to filter manufacturers
Summary

● Crucial to collect information about the use of dual-use technologies for censorship

● FilterMap - Framework for semi-automatically measuring filter deployments continuously and sustainably

● Found widespread use of filters for blocking access to content

● Data and Results available at https://censoredplanet.org/filtermap
Measuring the Deployment of Network Censorship Filters at Global Scale

Ram Sundara Raman¹, Adrian Stoll¹, Jakub Dalek², Reethika Ramesh¹, Will Scott³, Roya Ensafi¹

University of Michigan¹, The Citizen Lab², Independent³

Thank you

https://censoredplanet.org/filtermap
Backup Slides
Netsweeper

Canadian Filter Vendor

Enterprise Web Filtering
Protect the Network. Boost Productivity.

Country-wide Filtering and Regulatory Compliance
Protect citizens from harmful online content and ensure regulatory compliance within country borders

Dynamic Categorisation
Dynamic categorisation of web content, in real-time, with billions of URL already categorized into 90+ categories.

SSL Decryption
High-performance SSL decryption, that enables logging, reporting, and policy management of HTTPS traffic.
<table>
<thead>
<tr>
<th></th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OONI</strong></td>
<td>In-depth measurements close to the user (Volunteer -&gt; Site)</td>
<td>Scale, Continuity, Ethics</td>
</tr>
<tr>
<td><strong>Quack</strong></td>
<td>Scale - 33,000 vantage points</td>
<td>Only Port 7 measurements</td>
</tr>
<tr>
<td><strong>Hyperquack</strong></td>
<td>Port 80 and Port 443 measurements</td>
<td>Can only detect filter if it acts in both directions (MM -&gt; VP)</td>
</tr>
</tbody>
</table>
Blockpages as Identifiers

- Goes against the purpose of the censor to remove blockpages
- Vendors rarely have any incentive to remove trademarks
- Modified blockpages can still be detected
- Identification using blockpages is scalable
- Work can be extended to include other identifiers such as TCP/IP headers, DNS records, certificates
Unexpected Responses

- Observation - Disrupted measurements could either be filter **blockpages** or **unexpected responses** - Server not found errors, DDoS checks
- Similar to blockpages, Analysis also identified groups of unexpected responses
The page length metric

- **Fraction of HTML pages**
- **Page length**

Graph showing the fraction of HTML pages against page length for Blockpage and Unexpected Responses.
Data Collection

Censorship measurement techniques frequently observe blockpages

- **OONI**
  - Volunteer measurement
  - [https://ooni.org/](https://ooni.org/)

- **Quack**
  - Remote measurement
  - VanderSloot et al. [USENIX 2018]

- **Hyperquack**
  - New remote measurement
**OONI**

Direct measurement technique

**Pros**
- In-depth, user view

**Challenges**
- Limited scale
- Ethical constraints

---

TCP Handshake

GET https://blocked.com

Inject

Volunteer

Server
Quack

Remote measurement - TCP port 7 (Echo)

Pros
- 33,000 usable Echo servers

Challenges
- Cannot detect filters on common Port 80/443
Hyperquack

- Novel remote measurement technique introduced in this study
- Uses web servers running on port 80 and port 443
- Idea: Responses from web server when requesting a domain not hosted on the server is predictable
Ethics

- OONI provides good summary of risk and obtains informed consent
- Only use organizational servers in Quack and Hyperquack
  - Servers of ISPs
  - Echo servers having NMap labels such as routers, switches etc.
- Discussed the study with colleagues inside and outside the community
Ethics

- Set up WHOIS records and web page
- Spread our requests over many servers, make a single request at a time, add delays, and use a round-robin schedule
- Fresh TCP connections and close all states
- Average - triggered filters 99 times a day
# Vantage Point Characterization

<table>
<thead>
<tr>
<th></th>
<th>HTTP</th>
<th>HTTPS</th>
<th>Quack</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Set</td>
<td>9223</td>
<td>6200</td>
<td>36000</td>
</tr>
<tr>
<td>Experiment Set</td>
<td>9063</td>
<td>6070</td>
<td>33602</td>
</tr>
<tr>
<td>Number of Countries</td>
<td>215</td>
<td>204</td>
<td>75</td>
</tr>
<tr>
<td>Median / Country</td>
<td>11</td>
<td>13</td>
<td>151</td>
</tr>
<tr>
<td>Number of AS</td>
<td>4558</td>
<td>3442</td>
<td>3463</td>
</tr>
</tbody>
</table>
## Iterative Classification Evaluation

<table>
<thead>
<tr>
<th></th>
<th>BP (%)</th>
<th>#</th>
<th>UR (%)</th>
<th>#</th>
<th>UC (%)</th>
<th># of Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTTP</td>
<td>56.51%</td>
<td>27</td>
<td>39.39%</td>
<td>105</td>
<td>4.10%</td>
<td>3</td>
</tr>
<tr>
<td>HTTPS</td>
<td>3.48%</td>
<td>5</td>
<td>83.83%</td>
<td>67</td>
<td>12.70%</td>
<td>1</td>
</tr>
<tr>
<td>Quack</td>
<td>93.08%</td>
<td>34</td>
<td>4.8%</td>
<td>116</td>
<td>2.12%</td>
<td>2</td>
</tr>
<tr>
<td>OONI</td>
<td>13.02%</td>
<td>16</td>
<td>43.27%</td>
<td>44</td>
<td>43.71%</td>
<td>2</td>
</tr>
</tbody>
</table>
FilterMap Results - Data Collection

- Hyperquack - 38 signatures - Mostly commercial products
- Quack - 49 signatures - Mostly ISP deployments
- OONI - 21 signatures - Mostly ISP and organizational deployments
- Hyperquack detected deployments in three times as many countries as Quack and OONI
FilterMap Results - Blockpages

- Blockpages in 14 languages - Majority of blockpages were in English
- Most blockpages cited a legal concern for blocking access to content
- Many blockpages were served from redirects
## FilterMap Results - Manufacturing Country

<table>
<thead>
<tr>
<th>Country of Origin</th>
<th>Commercial filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Israel</td>
<td>Allot, Senhua</td>
</tr>
<tr>
<td>China</td>
<td>SmartxFilter, VAS Experts</td>
</tr>
<tr>
<td>Republic of Korea</td>
<td>Barracuda, CacheFlow, Cisco, Fortinet, IBM QRadar, Juniper, Palo Alto, SonicWall, Squid, Sucuri, WatchGuard</td>
</tr>
<tr>
<td>Russia</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td></td>
</tr>
</tbody>
</table>
FilterMap Results - Categories
FilterMap Results - Longitudinal
FilterMap Results - Censys

<table>
<thead>
<tr>
<th>Filter</th>
<th># of IPs</th>
<th># of countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barracuda</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>Fortinet</td>
<td>10,748</td>
<td>151</td>
</tr>
<tr>
<td>Juniper</td>
<td>41</td>
<td>2</td>
</tr>
<tr>
<td>Palo Alto</td>
<td>3,087</td>
<td>72</td>
</tr>
<tr>
<td>Watchguard</td>
<td>211</td>
<td>28</td>
</tr>
<tr>
<td>Cisco</td>
<td>1,434</td>
<td>63</td>
</tr>
<tr>
<td>IBM QRadar</td>
<td>22</td>
<td>5</td>
</tr>
<tr>
<td>SmartxFilter</td>
<td>33,639</td>
<td>2</td>
</tr>
<tr>
<td>Sucuri</td>
<td>24</td>
<td>8</td>
</tr>
<tr>
<td>Squid</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>