



# Censored Planet

An Internet-wide, Longitudinal Censorship Observatory

**Ram Sundara Raman**, Prerana Shenoy, Katharina Kohls\*, Roya Ensafi

University of Michigan, \*Ruhr University Bochum



**Internet censorship is pervasive and evolving**



**Measuring Internet  
censorship globally is  
a complex problem**

# Measuring Internet censorship is complex



Censorship  
method variance

# Measuring Internet censorship is complex

---

## Censorship method variance

DNS manipulation

TCP/IP blocking

Application layer (HTTP)  
connection

# Measuring Internet censorship is complex

---

Censorship  
method variance

Geographical and  
Network variance



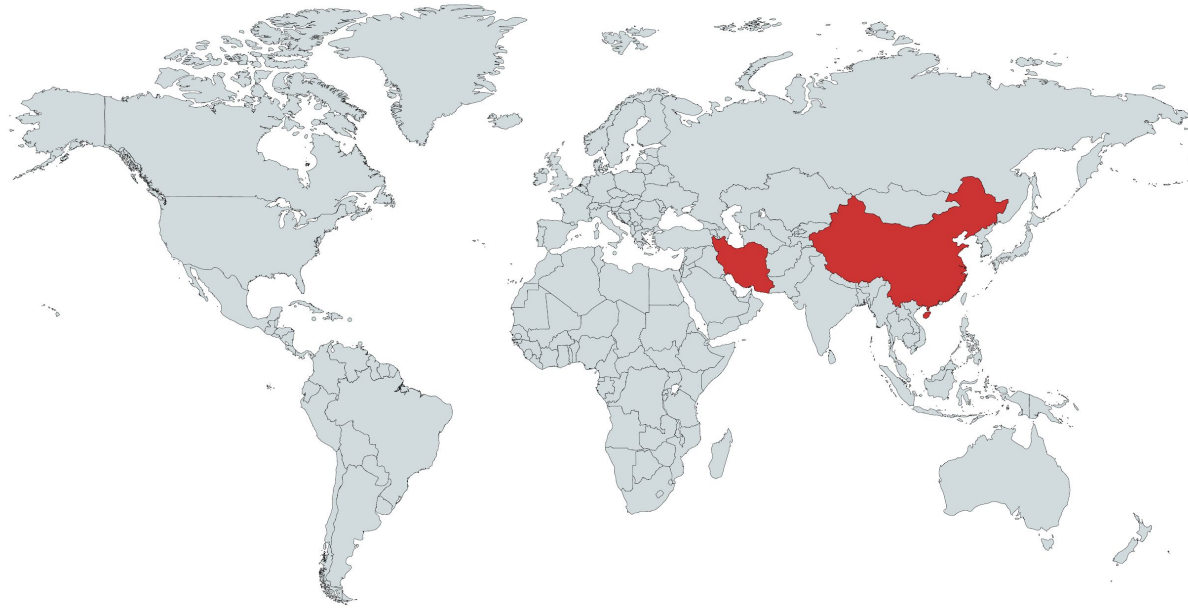
# Measuring Internet censorship is complex

---

Censorship  
method variance

Geographical and  
Network variance

Longitudinal  
variance



**Previous studies: Few countries and limited snapshots**

---

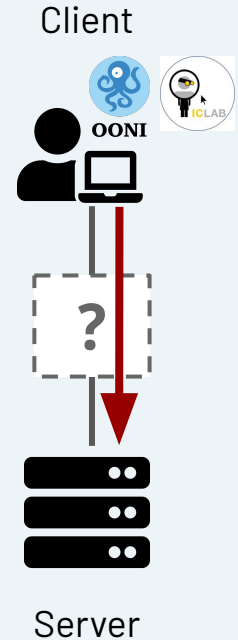


# Direct Censorship Measurement Platforms

- Ask **volunteers** on the ground, or deploy software or hardware in censored region (e.g. OONI probe<sup>[1]</sup>)
- Use **VPNs**, or **research networks** (e.g. ICLab<sup>[2]</sup>)

[1] Open Observatory of Network Interference, <https://ooni.org>

[2] A. Akhavan Niaki, S. Cho, Z. Weinberg, N. P. Hoang, A. Razaghpanah, N. Christin, and P. Gill. ICLab: A Global, Longitudinal Internet Censorship Measurement Platform. In IEEE Symposium on Security and Privacy (SP), 2020.



# Limitations of Direct Measurements

Scale

Coverage

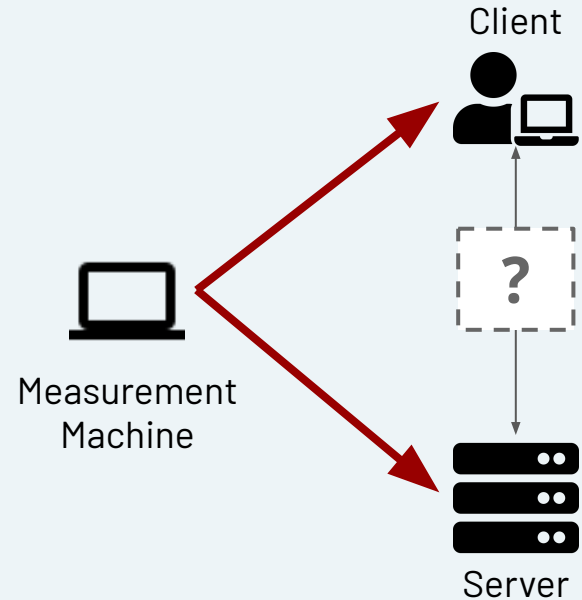
Continuity

Synchronization

Ethics

# Remote Censorship Measurements

Detect whether pairs of hosts  
around the world can talk to each other  
without controlling either endpoint.



**TCP**

## **Augur**

Augur: Internet-Wide Detection of Connectivity Disruptions, IEEE S&P 2017

**DNS**

## **Satellite/Iris**

Global Measurement of DNS Manipulation, Usenix Security 2017  
Satellite: Joint analysis of CDNs and network-level interference, Usenix ATC 2016

**Echo  
Discard**

## **Quack**

Quack: Scalable Remote Measurement of Application-Layer Censorship, Usenix Security 2018

**HTTP  
HTTPS**

## **Hyperquack**

Measuring the Deployment of Network Censorship Filters at Global Scale, NDSS 2020



## Limitations:

- Specialized techniques
- Limited snapshots
- Labor intensive
- Accuracy

## TCP

## Augur

Augur: Internet-Wide Detection of Connectivity Disruptions, IEEE S&P 2017

## DNS

## Satellite/Iris

Global Measurement of DNS Manipulation, Usenix Security 2017  
Satellite: Joint analysis of CDNs and network-level interference, Usenix ATC 2016

## Echo Discard

## Quack

Quack: Scalable Remote Measurement of Application-Layer Censorship, Usenix Security 2018

## HTTP HTTPS

## Hyperquack

Measuring the Deployment of Network Censorship Filters at Global Scale, NDSS 2020



# Censored Planet Observatory

---

- Collect global censorship measurement data continuously using remote measurement techniques (Augur, Satellite, Quack, Hyperquack)
- Analyze the data to create a more complete and accurate view of global Internet censorship
- Custom rapid focus measurements to analyze censorship events quickly



# Censored Planet Observatory

- Started in August 2018 and continuously collecting censorship data on 6 Internet protocols (TCP, DNS, Echo, Discard, HTTP, HTTPS)
- Continuous baseline of reachability data for **2000 sensitive domains and IP addresses (From Alexa and Citizen Lab) each week**
- More than **95,000 vantage points** in **221 countries and territories** (updated every week)
- Rapid focus capabilities to analyze censorship events in detail

# 21.8 billion

Measurements over 20 Months

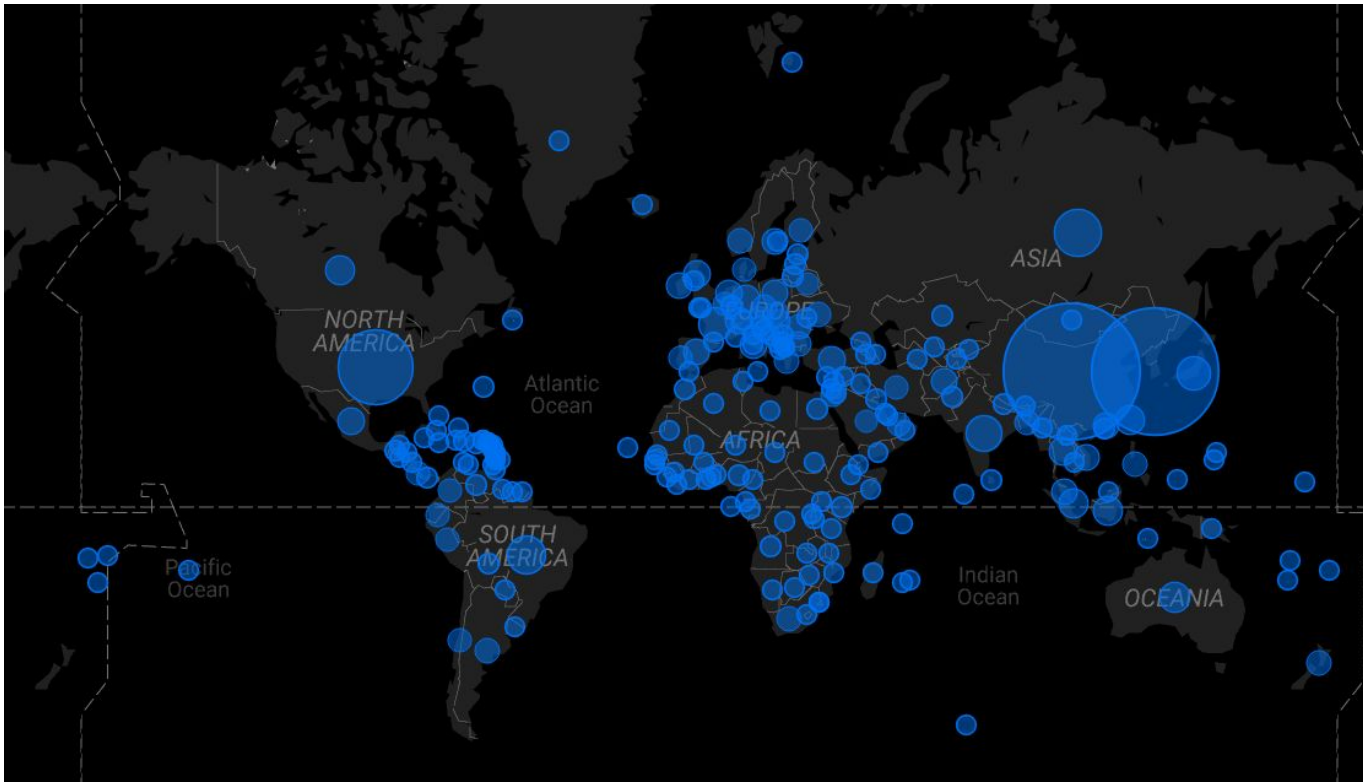
# 221 countries

42%-360% increase compared to OONI, ICLab

# 8 ASes (median)/country

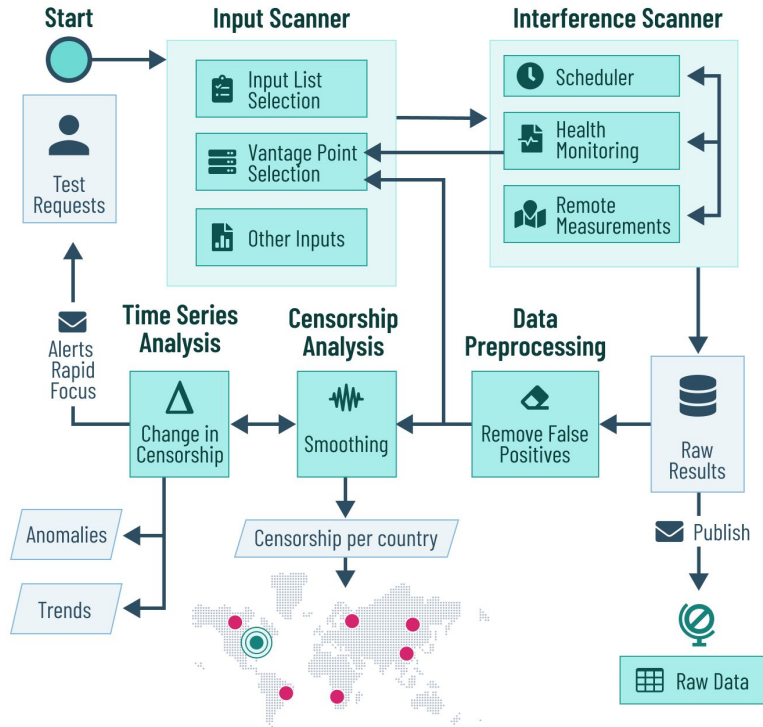
Median increase of 4-7 ASes per country





## Vantage Points in March 2020 (Scale 1-30k)

---

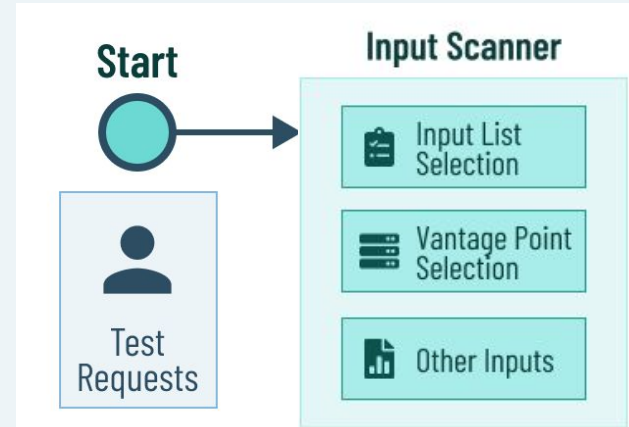


# Modular Design

---

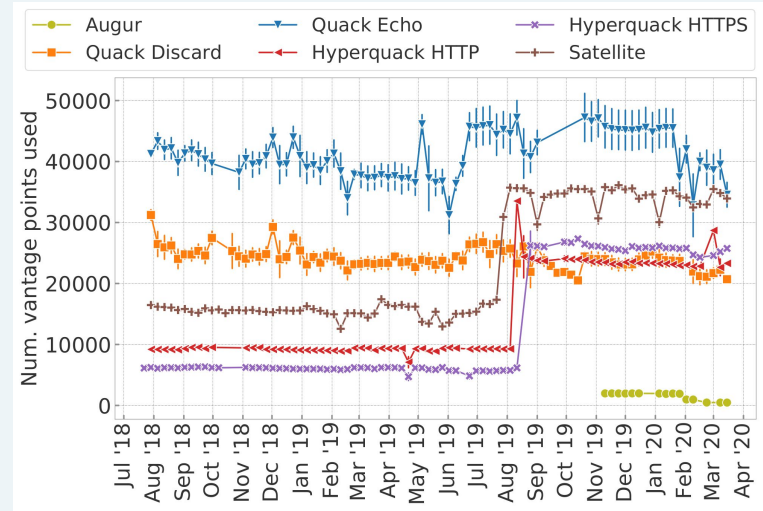
# Input Scanner

- **Vantage Point Selection**
  - Internet-wide scans for infrastructural machines
  - Consistency and diversity
- **Test List Selection**
  - Popular domains (Alexa)
  - Sensitive domains (Citizen Lab)
- Updated every week



# Input Scanner

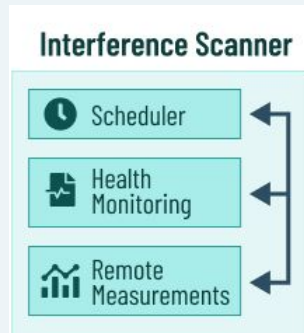
- **Vantage Point Selection**
  - Internet-wide scans for infrastructural machines
  - Consistency and diversity
- **Test List Selection**
  - Popular domains (Alexa)
  - Sensitive domains (Citizen Lab)
- Updated every week



Number of vantage points used over time

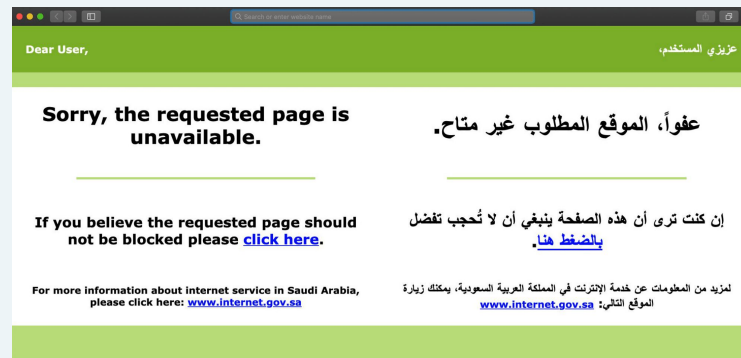
# Interference Scanner

- **Scheduler**
  - Manages vantage points and synchronizes measurements
- **Health Monitoring**
  - Vantage point health
  - Measurement errors
- Remote Measurements performed every week - Augur, Satellite, Quack, Hyperquack
- Raw data published



# Data Pre-processing

- Aggregating to common data schema
- **Confirming censorship**
  - Use clustering techniques in previous work<sup>[1]</sup> to find and group blockpages.
  - Consider only blockpage responses as censorship

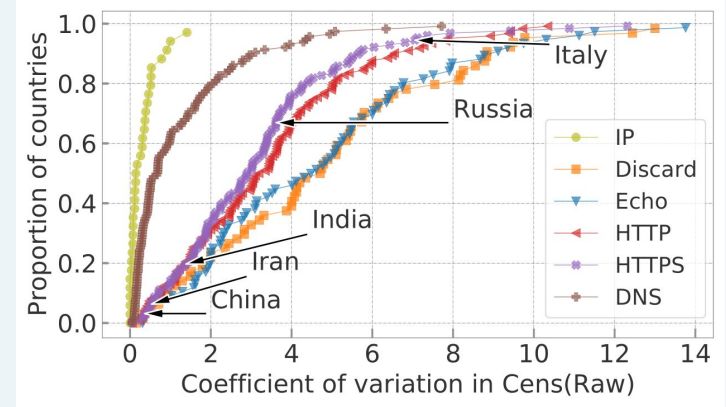


Example blockpage in Saudi Arabia

[1] Measuring the Deployment of Network Censorship Filters at Global Scale, NDSS 2020

# Censorship Analysis

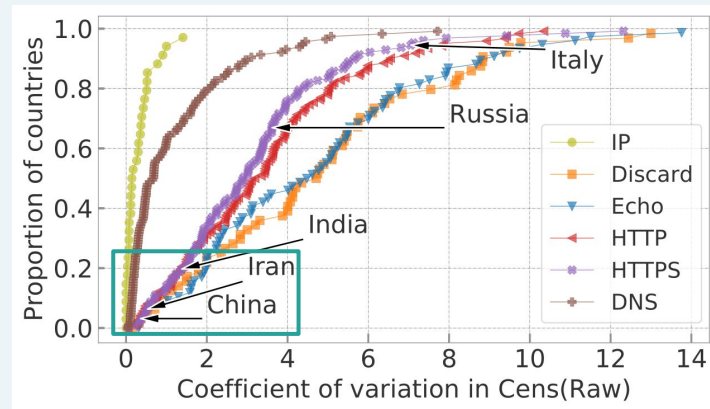
- Censorship values also vary *within* countries
- Countries with heterogeneous censorship policies have high variance in raw censorship values



Coefficient of variation in raw  
censorship metric

# Censorship Analysis

- Censorship values also vary *within* countries
- Countries with heterogeneous censorship policies have high variance in raw censorship values

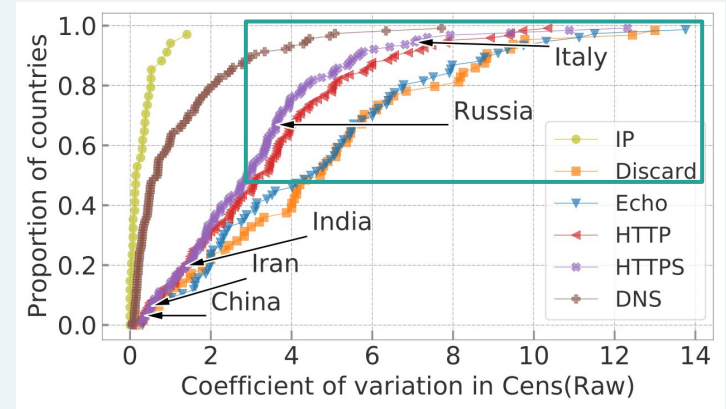


Coefficient of variation in raw  
censorship metric



# Censorship Analysis

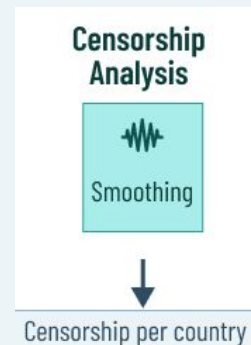
- Censorship values also vary *within* countries
- Countries with heterogeneous censorship policies have high variance in raw censorship values



Coefficient of variation in raw censorship metric

# Censorship Analysis

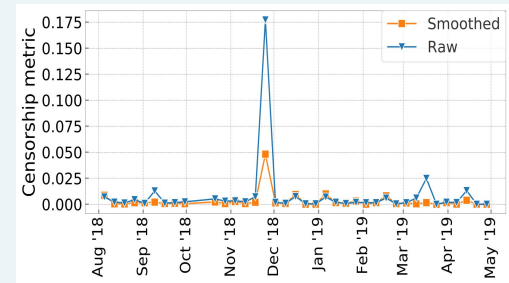
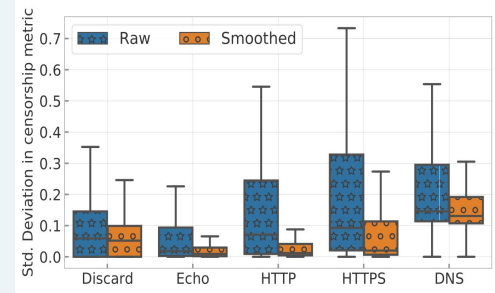
- Objective: Obtain a representative metric of censorship within a country that is not affected by outlier vantage points
- Apply an **optimization model** (Nelder-Mead) to obtain a weight for each Autonomous System that smooths the metric.



# Censorship Analysis

- Objective: Obtain a representative metric of censorship within a country that is not affected by outlier vantage points
- Apply an **optimization model** (Nelder-Mead) to obtain a weight for each Autonomous System that smooths the metric.

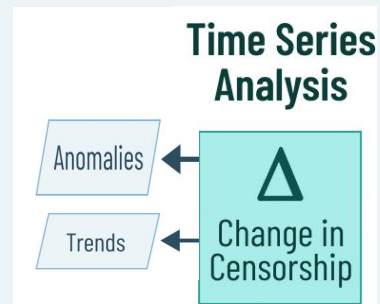
Standard deviation in censorship metrics



Censorship over time (Pakistan)

# Time Series Analysis

- **Anomaly Detection** - Bitmap-based detection
- **Trend Analysis** - Mann-Kendall test



## Findings

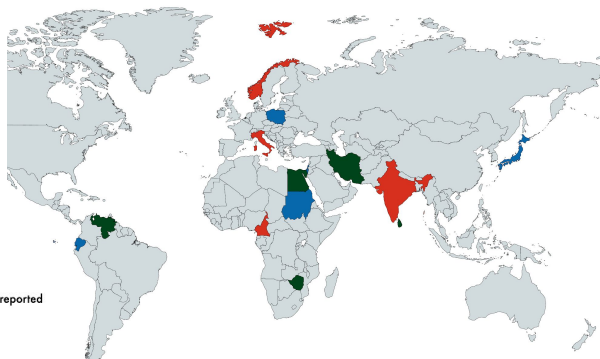
- Censorship Events
- Censorship Trends
- Case Studies

# Censorship Events

- Identified 15 key censorship events

- 5 previously reported

- 10 unreported

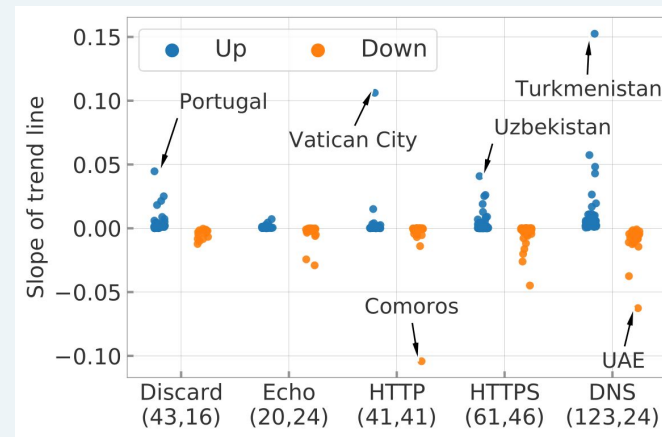


- Event known and reported
  - Event known
  - Event unknown

Country	Period	Method	Category or Domain blocked	Event
Egypt	26 Sep 2019	HTTP, HTTPS	News Media	Protests
Iran	Mar 2020	HTTP, Echo	wikimedia.com, wikia.com	Policy
Sri Lanka	21 Apr–12 May 2019	HTTP, HTTPS	Social	Terrorism
Venezuela	12–29 Jan 2019	HTTP, HTTPS	Networking	Unrest
Zimbabwe	20 Jan 2019	HTTP, HTTPS	Networking, wikipedia.org	Protests
Ecuador	8 Oct 2019	DNS	Social	Protests
India	6 Sep 2018	DNS	Networking	Law
Israel	May 2019–Jun 2019	DNS	Online Dating	Conflict
Japan	28 Jun 2019	DNS, Echo	Foreign	Relations and Military
Poland	22 Jul 2019	DNS, HTTP, HTTPS	News Media	Summit
Sudan	11 Apr 2019	HTTP, HTTPS	Govt., News	Unrest
Cameroon	25 Nov 2018	HTTP	Media, Human Rights	Unrest
India	Feb–Mar 2020	Echo, HTTPS	Social	Networking
Italy	22 Dec 2019	Discard	Human Rights	Unknown
Norway	Dec 2019–Mar 2020	DNS	Multiple	Unknown

# Censorship Trends

- Increasing levels of DNS censorship in more than 100 countries.
- HTTPS censorship also showing increasing trend.
- 11 categories of domains increasingly blocked - News Media, Provocative Attire, Human Rights, Gaming.

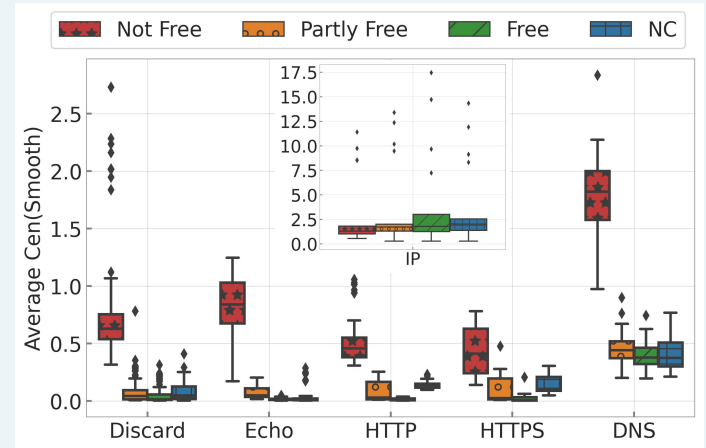


Upward and downward trends in censorship methods

# Freedom on the Net Report

- Observe significant censorship in countries labelled as “Free” or Not Considered by qualitative reports like Freedom on the Net<sup>[1]</sup>.
- Data-driven insights from Censored Plane can significantly complement qualitative reports.

[1] <https://freedomhouse.org/countries/freedom-world/scores>, 2019.

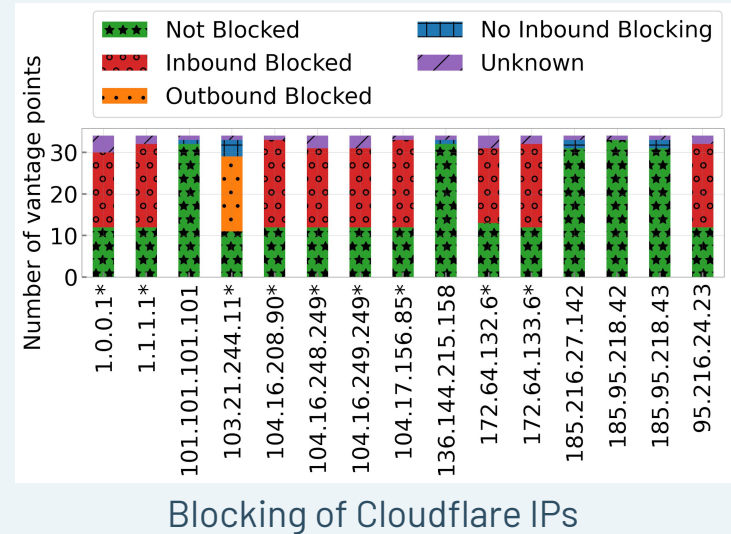


Censorship in Freedom on the Net categories



# Rapid Focus Study - Turkmenistan

- Very strict censorship policies
- Conducted rapid focus Augur measurements to DoH server IPs and Cloudflare IPs in April 2020.
- 52.9% of vantage points in Turkmenistan block all Cloudflare IPs, restricting access to thousands of services.
- Other DoH servers (eg. Snopyta) also blocked.



## Large-scale, continuous censorship measurement is essential!

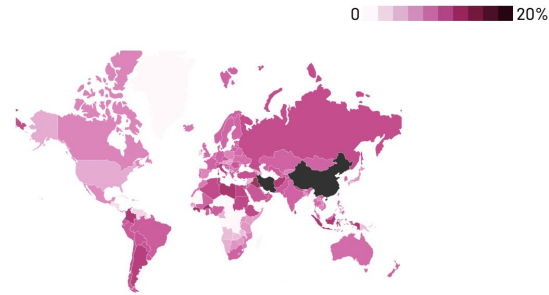
Our study finds:

- **Censorship varies over time** → Continuous censorship measurement is crucial
- **Censorship in 'free' countries** → Censorship measurements should cover all countries
- **Censorship increase in encrypted communication** → Censorship circumvention needs to apply obfuscation
- **Censorship missed by qualitative reports** → Data-driven measurements can complement reports

# Website

<https://censoredplanet.org>

## Percentage of resolvers facing interference by country



## Top disrupted domains by country

Country	Domain	Disrupted percentage
Brazil	www.date.com	76.65
Brazil	www.agentprovocateur.com	76.42
Brazil	www.hrw.org	75.94
Brazil	www.163.com	71.46
Brazil	creditkarma.com	65.68

Date and Time of Scan	File Name	Scan Tool	Scan Type	Size of File in MB
2020-06-24T06:01:03	CP_Quack-echo-2020-06-24-06-01-03.tar.gz	Quack - echo	Application Layer	621.177
2020-06-23T00:08:31	CP_Quack-https-2020-06-23-00-08-31.tar.gz	Quack - https	Application Layer	3940.94
2020-06-22T14:45:38	CP_Quack-https-2020-06-22-14-45-38.tar.gz	Quack - https	Application Layer	3340.128
2020-06-22T01:02:10	CP_Quack-http-2020-06-22-01-02-10.tar.gz	Quack - http	Application Layer	1580.374
2020-06-21T12:00:01	CP_Satellite-2020-06-21-12-00-01.tar.gz	Satellite	DNS Layer	7137.384



# Censored Planet

## Thank you!

<https://censoredplanet.org>



# Backup Slides

---

# Ethics

- Censorship research frequently raises ethical considerations e.g., under what conditions is it safe enough to use remote vantage points?
- IRBs are often not positioned to help. We turn to authorities such as the **Belmont and Menlo Reports** to guide ethical thinking.
- Frequently consult with colleagues to check our reasoning and conclusions.
- Ensure suitable protections in place, including technical practices to minimize risk to individuals
- Use hosts that are a part of Internet infrastructure
- Follow the ethical scanning guidelines, coordinate with our network administrators and our upstream ISP, host web pages indicating purpose of probes, set rate limits and more

# OOONI's approach

- “Provide as much informed choice to the user as possible”
  - Choose websites to test
  - Choose whether to upload measurement
  - Choose what type of data to submit
- Establish relationships with local & civil society
- Keep the community of volunteer involved

# Coverage

---

Platforms	# AS	# Country	Median ASes / country
ICLab	56	48	1
OONI	1,915	155	4
Censored Planet Observatory	9,014	221	8