CCC Panel
October 24, 2017

Developing a Science of Internet Censorship Resistance: Opportunities and Challenges for Network Measurement

Phillipa Gill

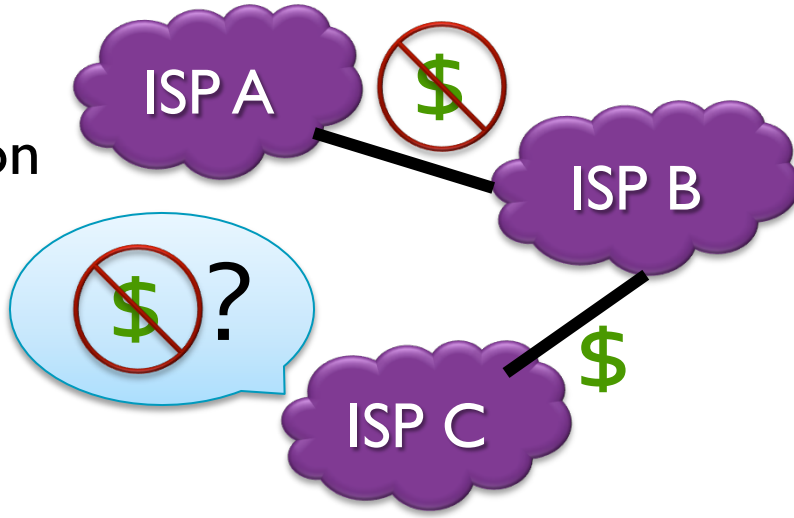
University of Massachusetts -- Amherst

The challenge of network measurement

What we want to measure != What we are able to measure

Decentralized Internet:

- No global measurement primitives
- Strong incentives to **hide** information
- How can we learn **network paths**?



Variety of edge networks:

- Broadband subscribers
- Mobile
- Data centers
- How can we get **vantage points**?

The challenge of network measurement

Variety of content

- Mobile apps + games
- Streaming video
- Social networks
- How to measure **diverse apps**?

Measuring network interference compounds the challenge!

Measuring censorship/network interference

- Networks are not always transparent in their actions
- Interference may not be consistent
- Measurements may pose a risk to users
- How to measure when the **subject may hide**?

Winning the censorship arms race

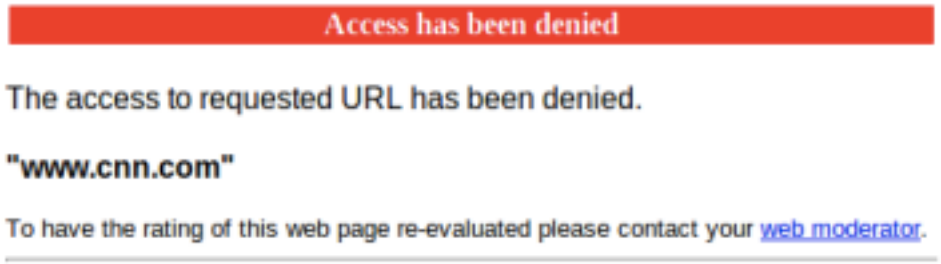
- **Collaborative approach**
  - Bringing together **political science** context and cutting edge **network measurement** techniques
- **Circumvention**
  - Improving Tor anonymity with **empirical data**
  - Designing covert channels with **adaptability** in mind
- **Measuring politically motivated actors**
  - Targeted attacks against activist groups
  - Fingerprinting filtering products

Censorship Measurement 101

- Example:

Measured in the lab

Measured in the field

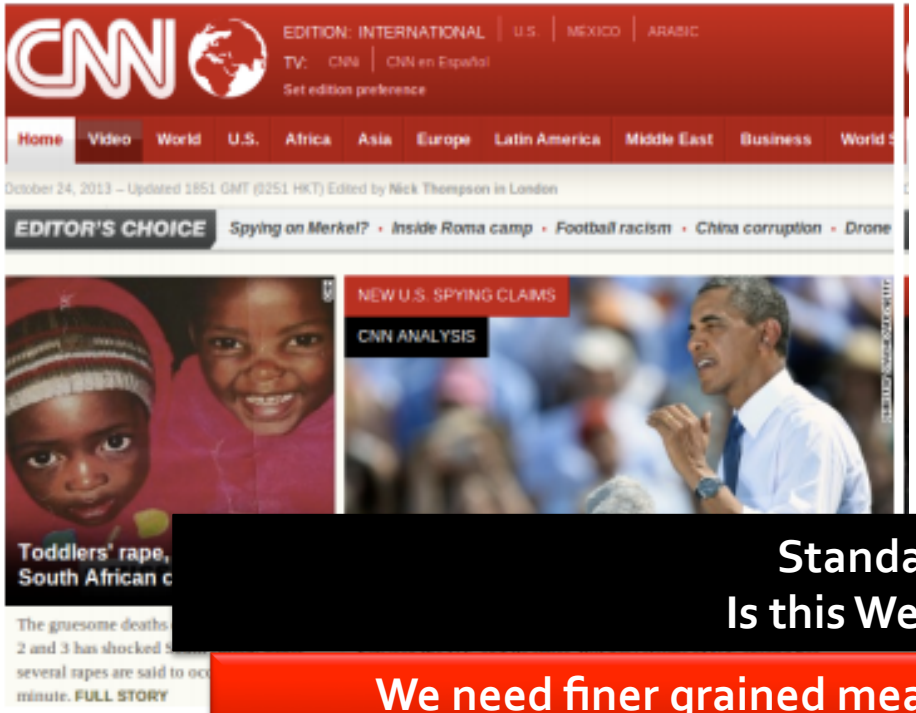


Standard question:
Is this Web site blocked?

Censorship Measurement 101

- Example:

Measured in the lab



Measured in the field

(no html page returned)

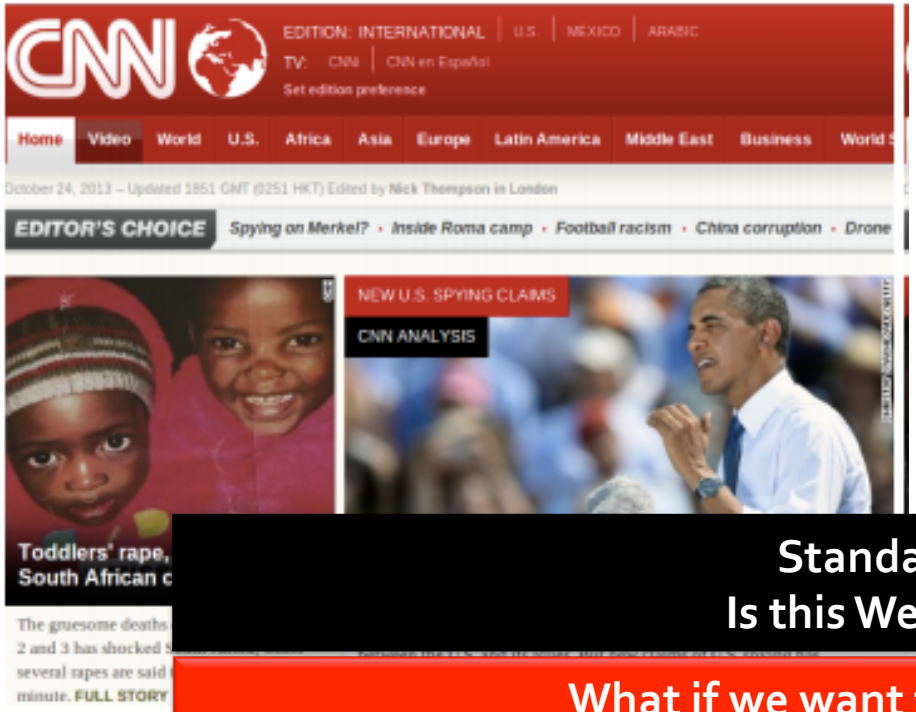
Standard question:
Is this Web site blocked?

We need finer grained measurements to answer this question!

Censorship Measurement 101

- Example:

Measured in the lab



Measured in the field

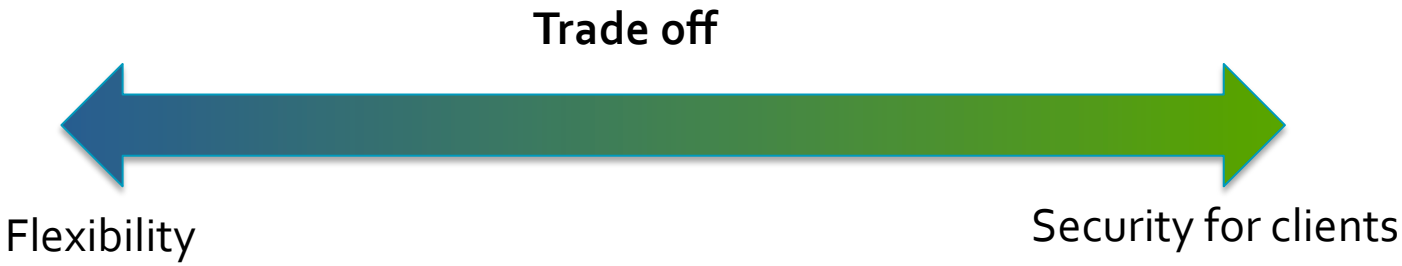
(no html page returned)

Standard question:
Is this Web site blocked?

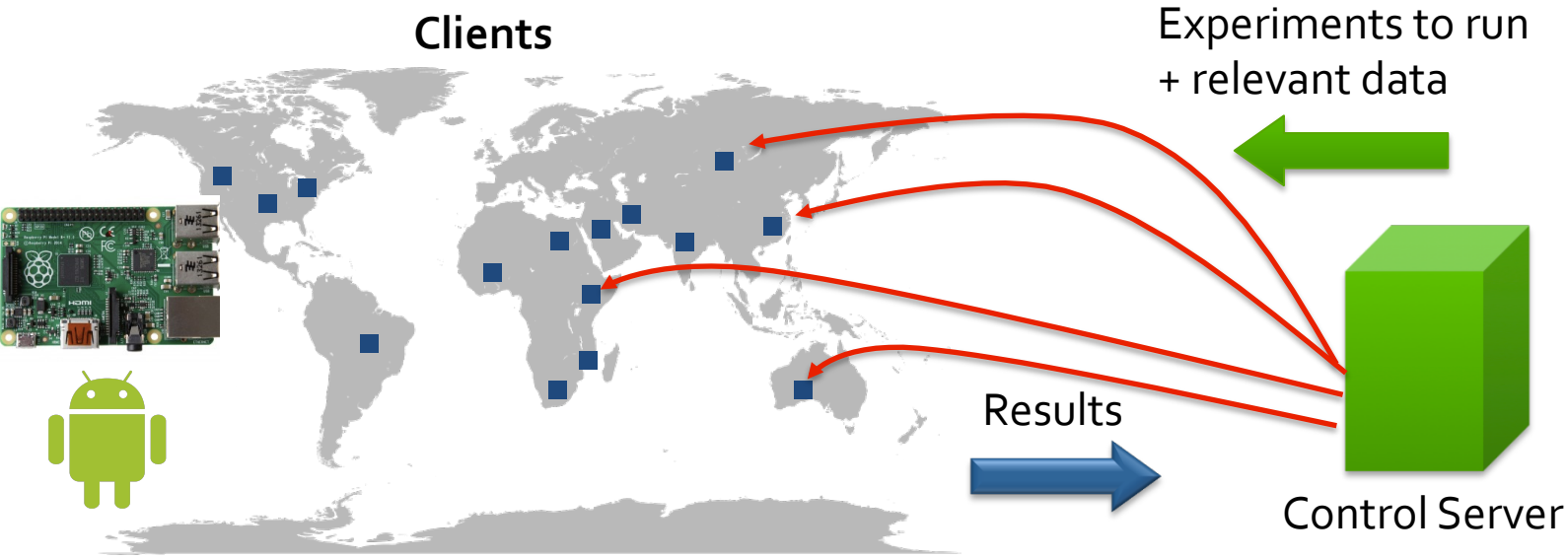
What if we want to ask more questions:
How was this site blocked?
What product was used to block it?
Who is blocking it?

What does this mean for measurement?

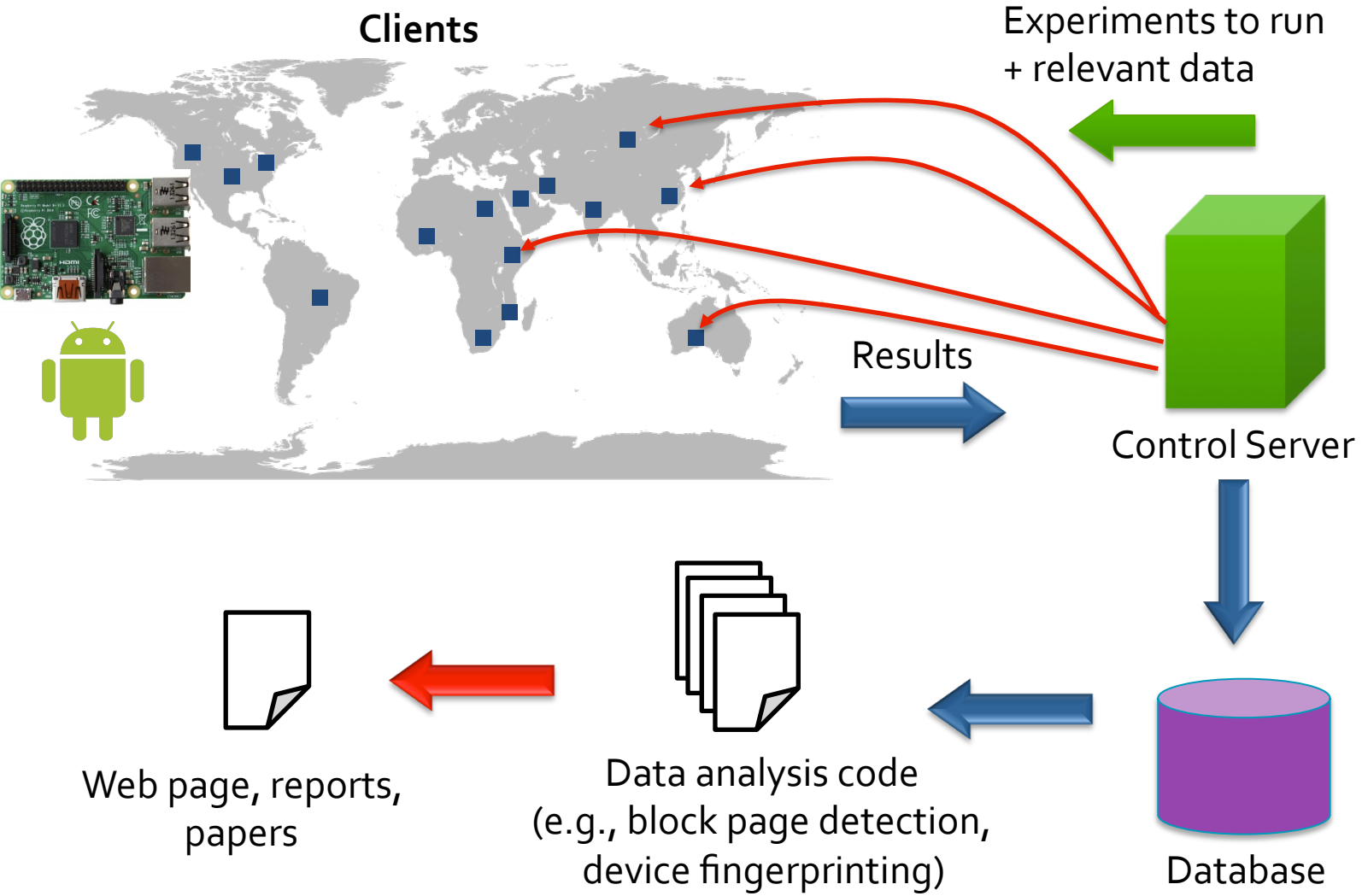
- Don't know the set of measurements to be support a priori
  - Need to keep up with new censorship technologies
- Need to be able to implement and launch new experiments
- Need to be flexible in terms of when, where, and what is run
- How to do this well?



Overview of ICLab



Overview of ICLab



ICLab use case 1: Yemen conflict

UNIVERSITY OF TORONTO

MUNK SCHOOL OF GLOBAL AFFAIRS

Information Controls during Military Operations

THE CASE OF YEMEN DURING THE 2015 POLITICAL AND ARMED CONFLICT

21 October 2015

ICLab use case 1: Yemen conflict

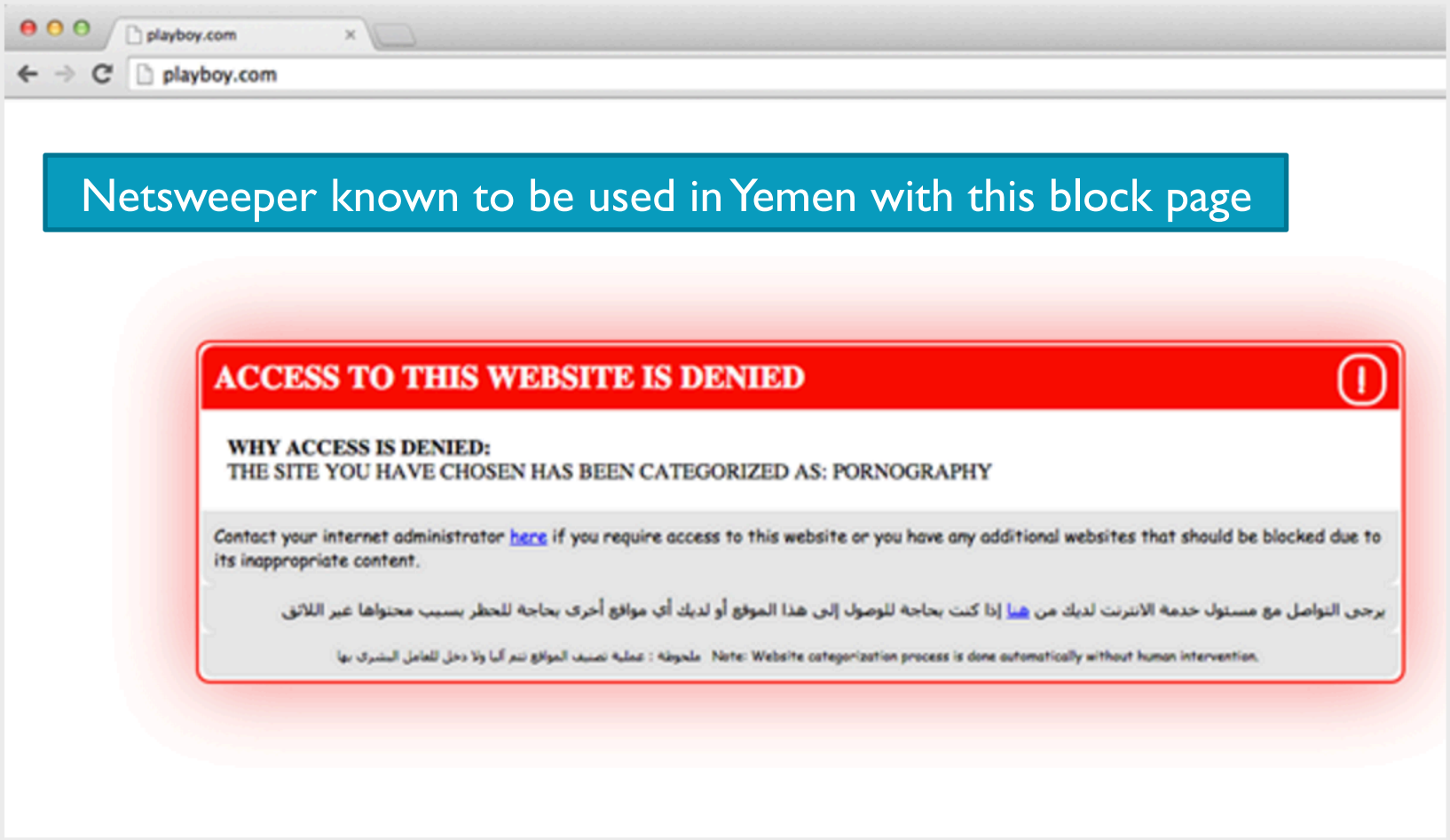
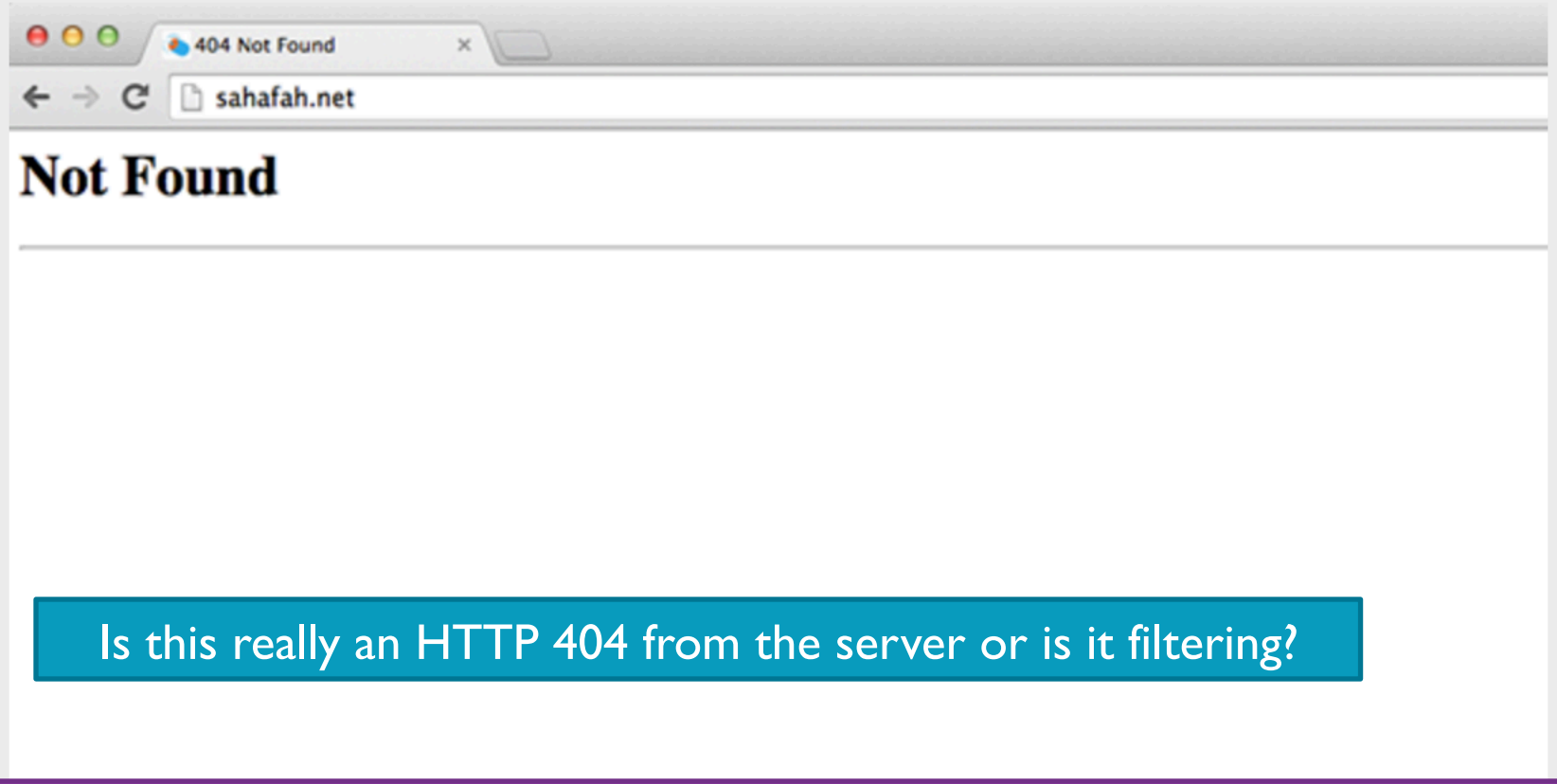


Figure 13: Descriptive Netsweeper blockpage for an attempt to access playboy.com

**Tests run by a pseudonymous researcher who was already in the country.

ICLab use case 1: Yemen conflict



Is this really an HTTP 404 from the server or is it filtering?

We can look at network packet headers to find out!

Figure 20: HTTP 404 error displayed during an attempt to access sahafah.net on YemenNet

ICLab use case 2: Iranian filtering

Filterwatch // Episode 5 – Measuring Internet Censorship

On this month's show, we'll talk about a recent report published by Small Media and ICLab which aimed to test whether and by what method certain websites were blocked in Iran. We'll speak to PhD candidate Abbas Razaghpanah, who led the technical research on the report, about some of the findings and what they mean.

LISTEN

Monthly blog/podcast with SmallMedia UK.

Tests carried out using a virtual private server.

ICLab use case 2: Iranian filtering

Key question: How do tech sanctions influence Internet censorship?

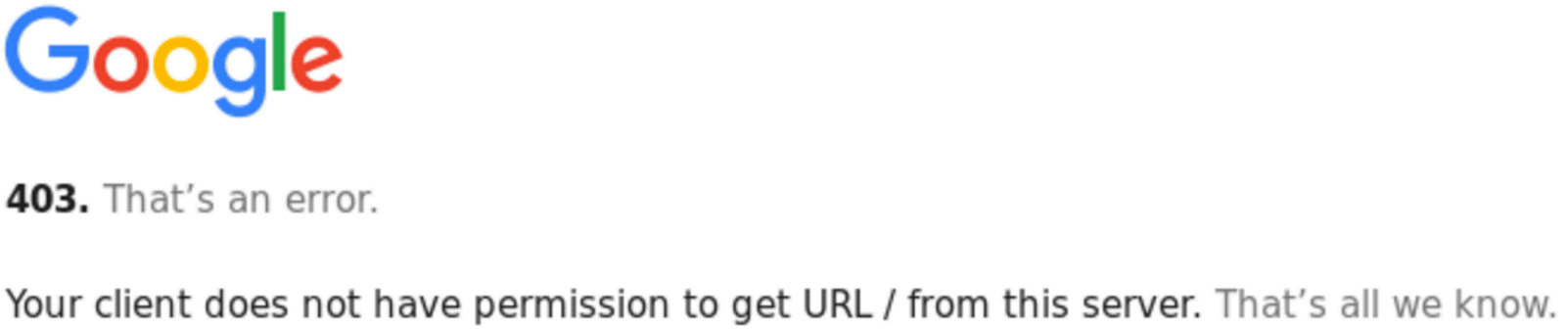
URL	VPS-1	VPS-2	VPS-3	VPS-4
https://developers.google.com/	Serverside	Serverside	Serverside	Serverside
https://get.adobe.com/flashplayer/	Local (RST)	OK	OK	Local (RST)
https://developer.android.com/studio/index.html	Serverside	Serverside	Serverside	Serverside
https://analytics.google.com/	Serverside	Serverside	Serverside	Serverside
https://code.google.com/codejam	OK	OK	OK	OK
https://github.com/	OK	OK	OK	OK
https://azure.microsoft.com	OK	OK	OK	OK
https://aws.amazon.com/ec2/	OK	OK	OK	OK
https://www.dropbox.com/	OK	OK	OK	OK
https://www.netflix.com/	OK	Local (DNS)	Local (RST)	OK
https://www.bitbucket.org	Local (RST)	Local (RST)	Local (RST)	Local (RST)

Notes:

RST indicates that a website was blocked by the Iranian authorities via packet injection.

DNS indicates that a website was blocked by the Iranian authorities via DNS tampering.

ICLab use case 2: Iranian filtering



The page we were redirected to when we tried to access Google services from inside Iran.

Looks like server-side filtering to enforce the sanctions …

What is needed?

- **Still hitting fundamental challenges in network measurement!**
  - IP geolocation
  - Measuring/predicting network paths
- **Pushing the boundaries of measurement**
  - Using TCP side channels to study censorship
  - Trade off between scale and depth of measurement
- **We need to understand how censors behave!**
  - How can we evaluate circumvention schemes without understanding the adversary?
  - Complicated by political factors

More info:
http://www.cs.umass.edu/~phillipa
@phillipa_gill