Open Information Access: Old Problems, Emerging Challenges

Nick Feamster
Georgia Tech

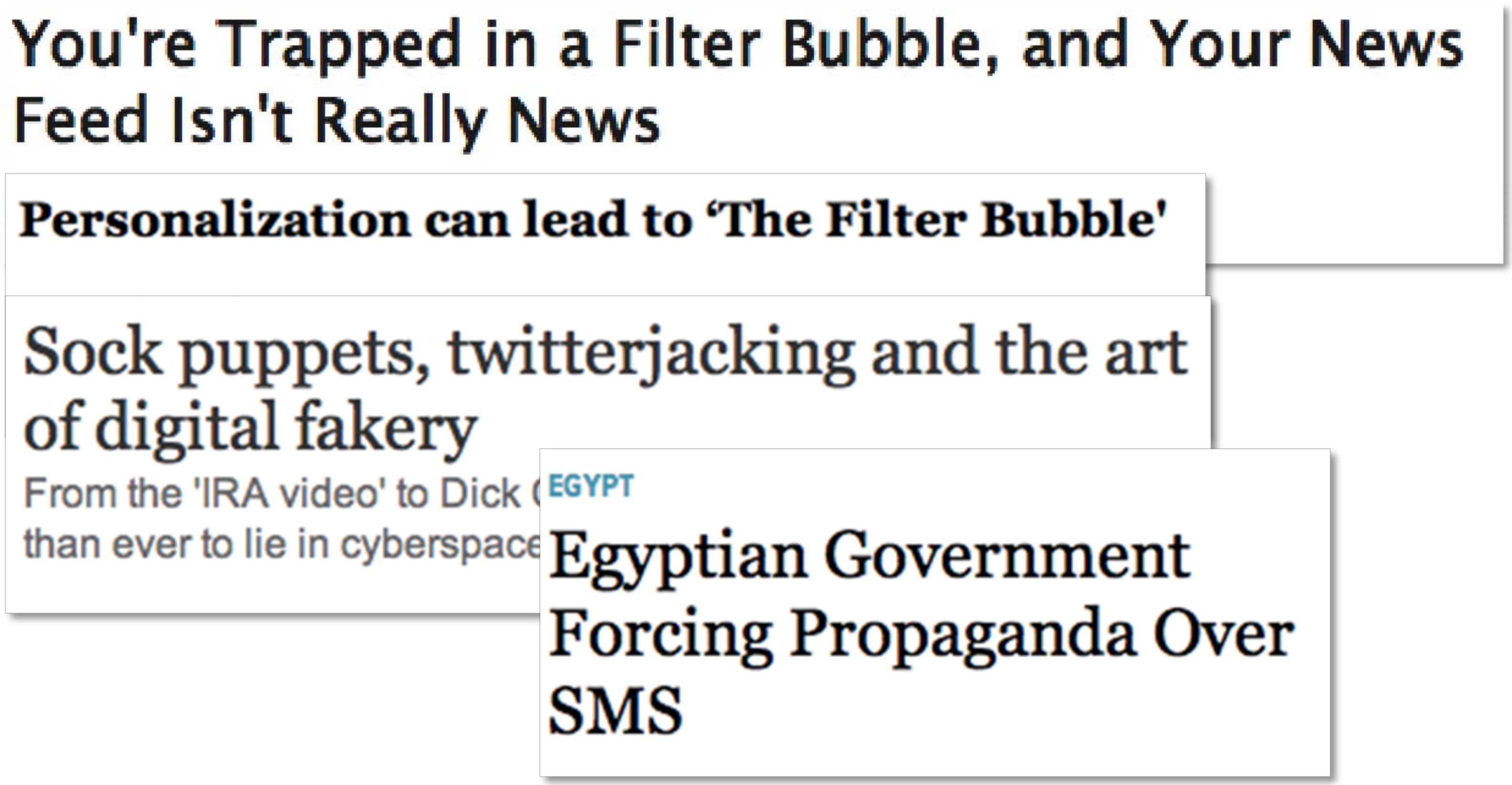
An “Old” Problem

- Many governments/companies trying to limit their citizens’ access to information
  - Censorship (prevent access)
  - Punishment (deter access)
  - Surveillance (spy on or monitor access)
  - China, Saudi Arabia, many companies

- **How can we defeat such attempts?**
  - Circumvent censorship
  - Undetectably

New Forms

- **Increasing number of attempts to "hijack" media channels to influence public opinion**



You're Trapped in a Filter Bubble, and Your News Feed Isn't Really News

Personalization can lead to 'The Filter Bubble'

Sock puppets, twitterjacking and the art of digital fakery

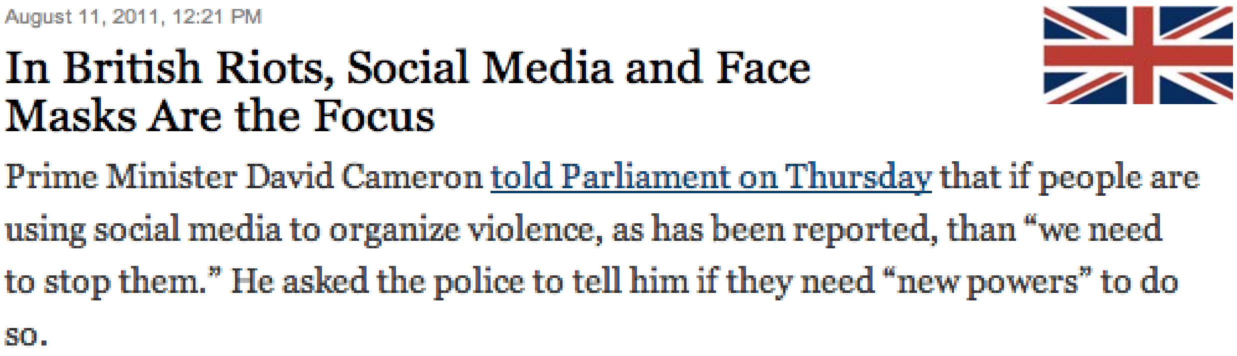
From the 'IRA video' to Dick (
than ever to lie in cyberspace

EGYPT

Egyptian Government Forcing Propaganda Over SMS

What countries censor, and why?

- **Political stability**

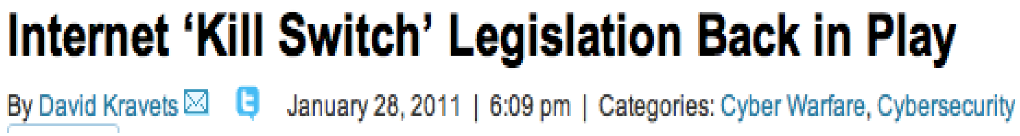


August 11, 2011, 12:21 PM

In British Riots, Social Media and Face Masks Are the Focus

Prime Minister David Cameron told Parliament on Thursday that if people are using social media to organize violence, as has been reported, than “we need to stop them.” He asked the police to tell him if they need “new powers” to do so.

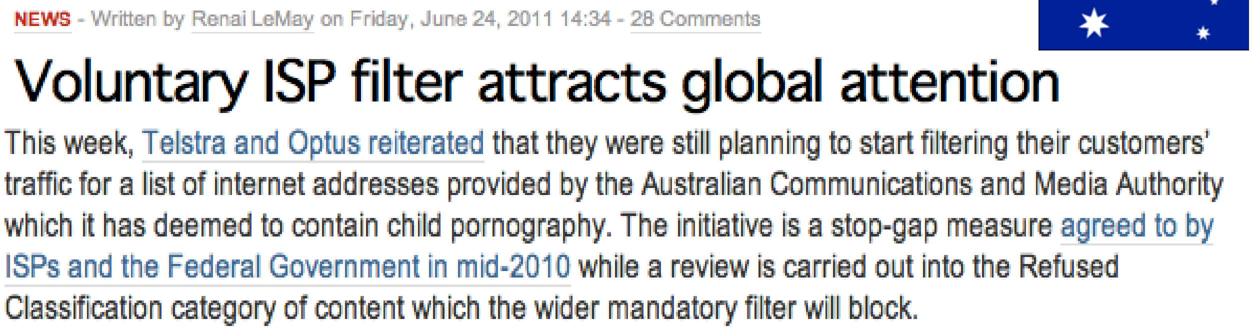
- **National security**



Internet ‘Kill Switch’ Legislation Back in Play

By David Kravets | January 28, 2011 | 6:09 pm | Categories: Cyber Warfare, Cybersecurity

- **Social values**

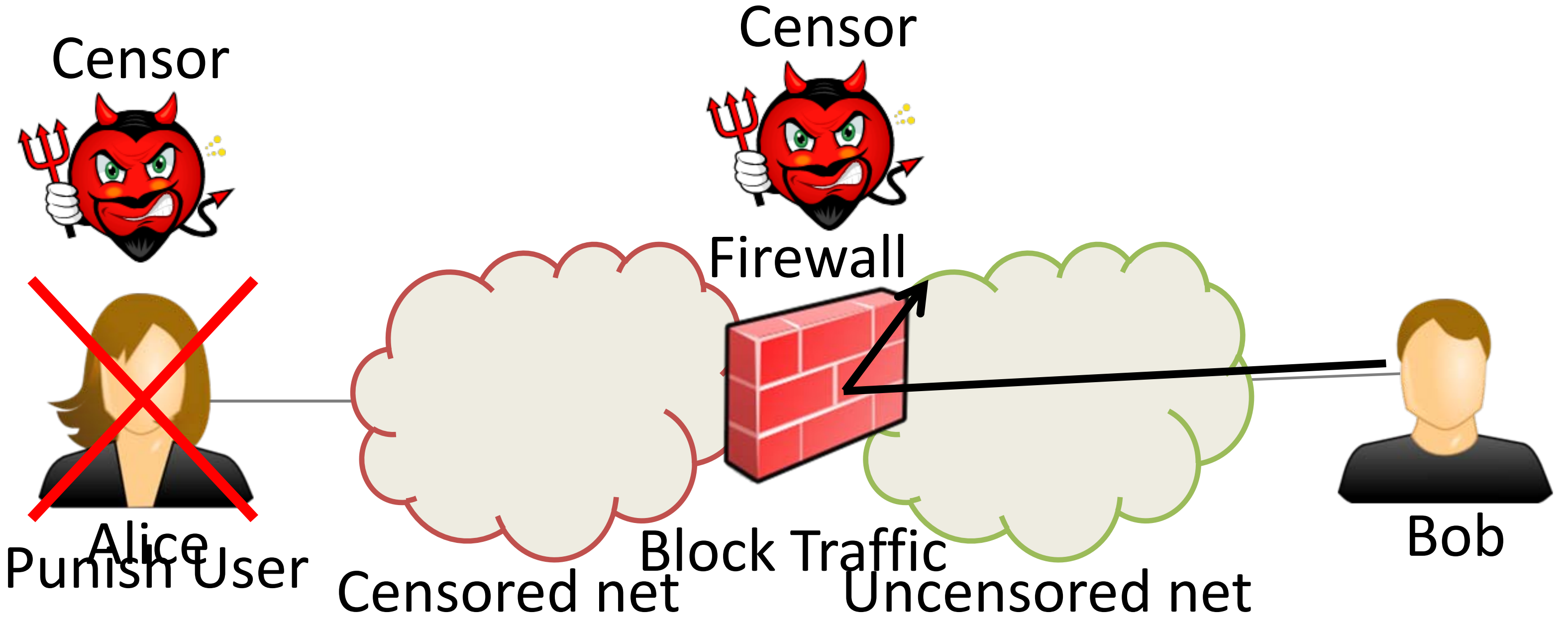


NEWS - Written by Renai LeMay on Friday, June 24, 2011 14:34 - 28 Comments

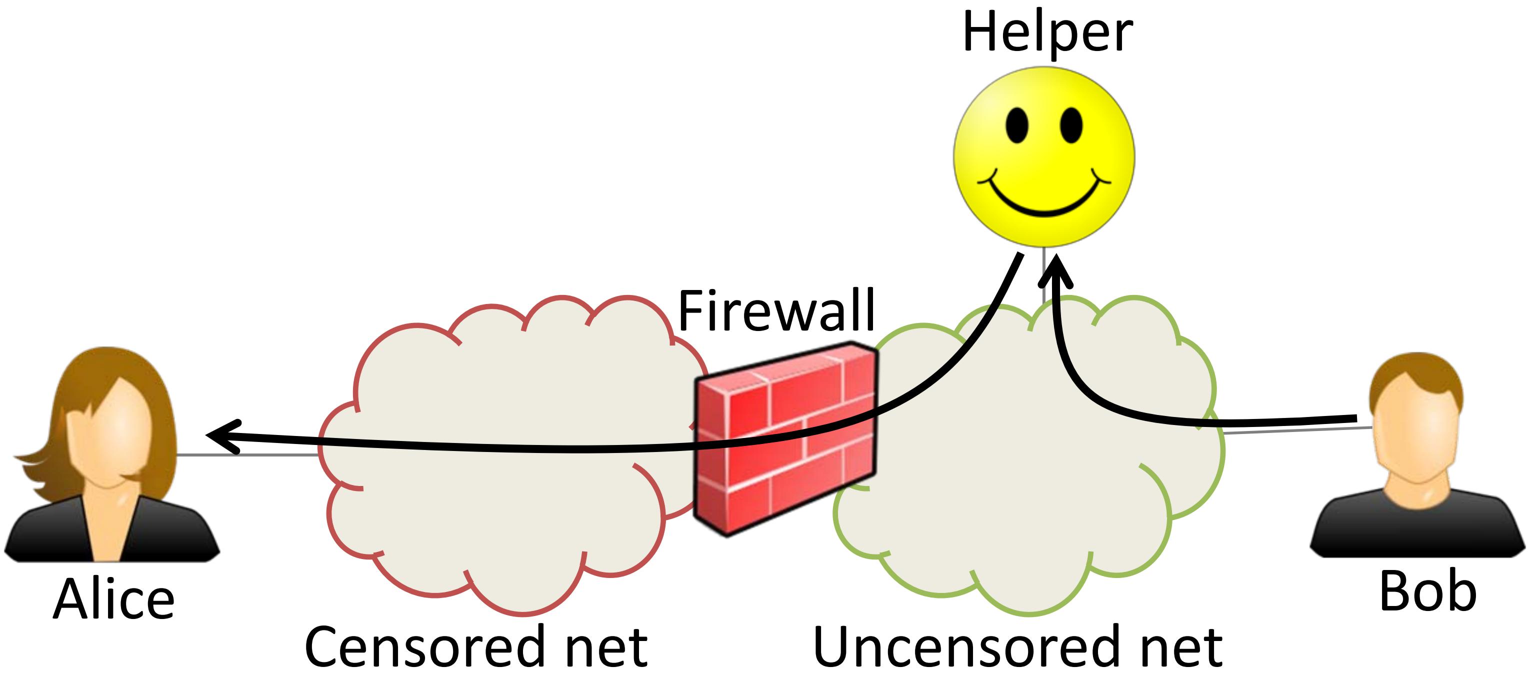
Voluntary ISP filter attracts global attention

This week, Telstra and Optus reiterated that they were still planning to start filtering their customers' traffic for a list of internet addresses provided by the Australian Communications and Media Authority which it has deemed to contain child pornography. The initiative is a stop-gap measure agreed to by ISPs and the Federal Government in mid-2010 while a review is carried out into the Refused Classification category of content which the wider mandatory filter will block.

Conventional Internet Censorship



Solution: Use a Helper



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

Requirements

Confidentiality…*and covertness.*

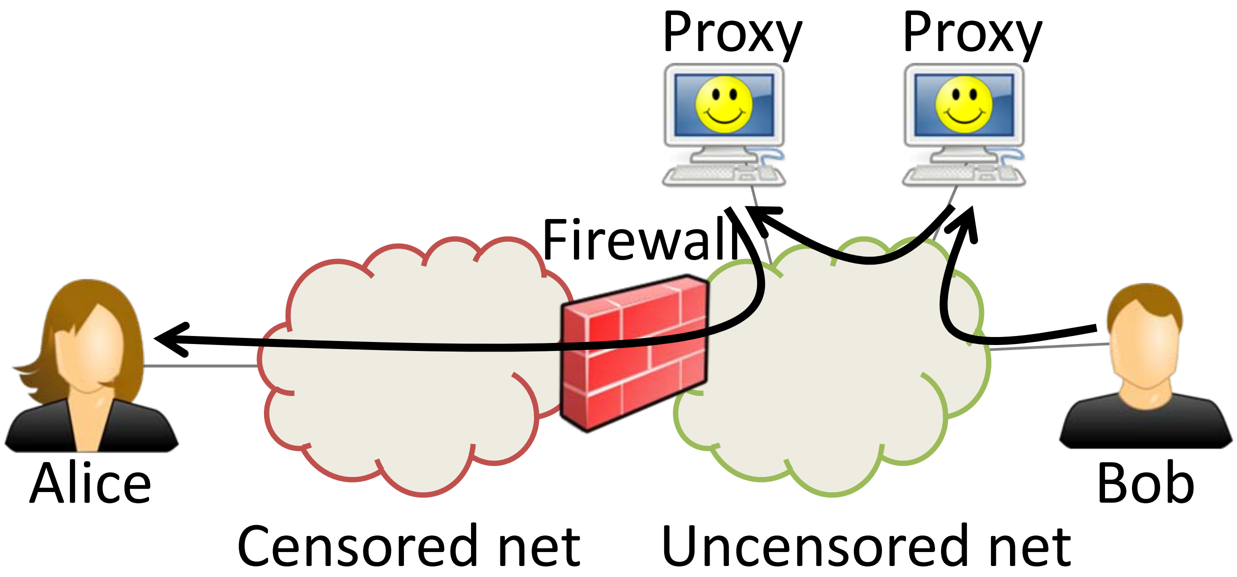
- Communication robustness
  - Even without detecting, censor could scramble covert channel
- Covertness/deniability
  - Detection could be embarrassing or worse
  - Even suspicion could be a problem
  - If server detected, can be blocked
- Minimal dedicated infrastructure
- Performance (bandwidth, latency)

Deniability is Hard

- Easy to hide what you are getting
  - E.g., just use SSL or some other confidential channel
- And easy to “get through” censors
  - Reflection (e.g., Safeweb)
- But hard to hide that you are doing it!
- To be practical, all these problems must be solved
- Want *both* confidentiality and covertness

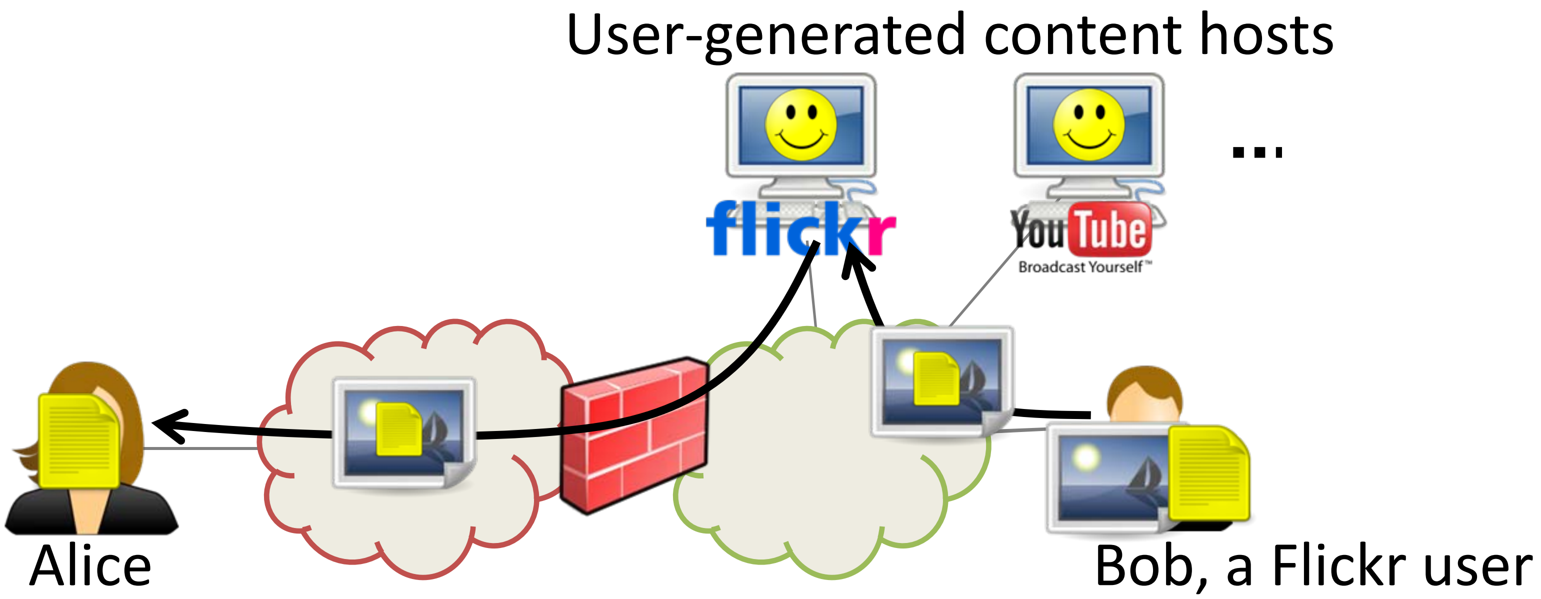
What about Proxies and Mixnets?

(*e.g.*, Tor)



- Censors can **block proxies** if the proxy list is public
- **Not deniable** if encryption is incriminating
- **Requires dedicated infrastructure** (network of proxies)

Collage: Let User-Generated Content Help Defeat Censorship

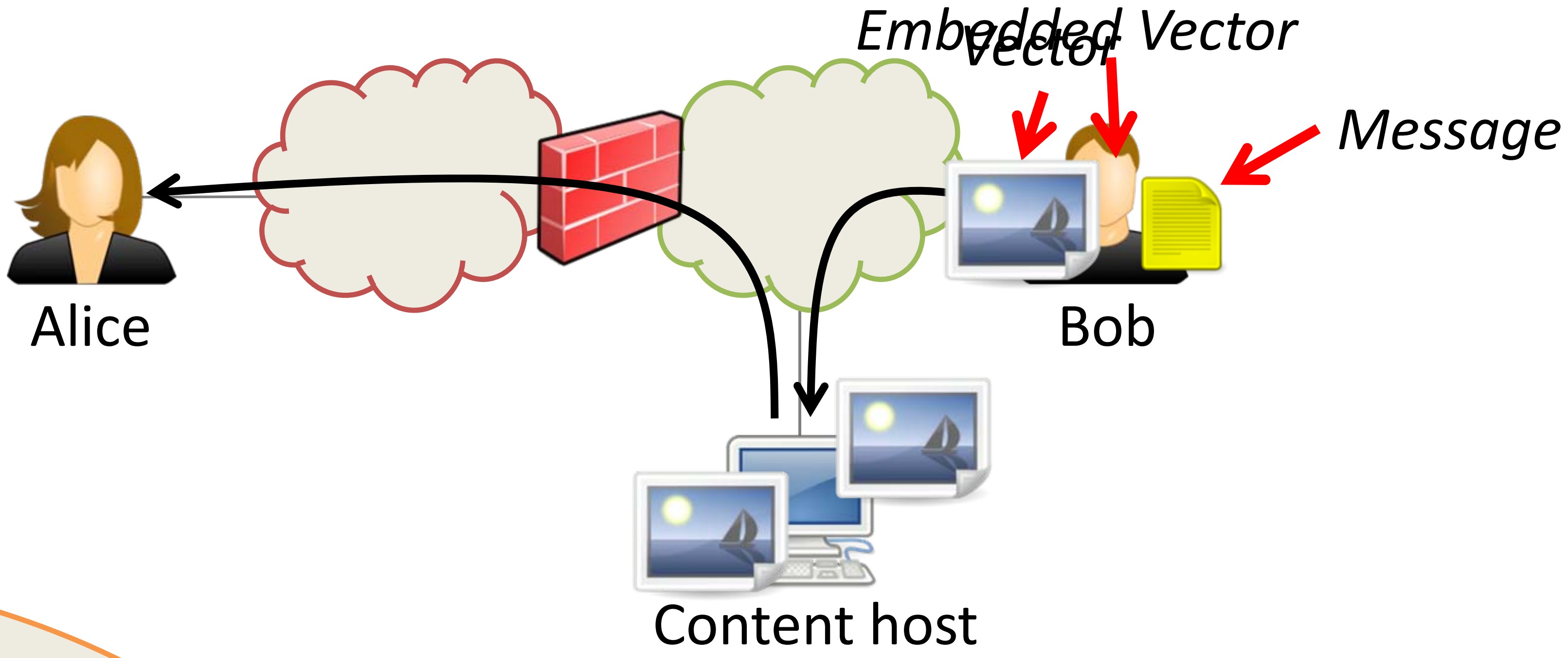


- **Robust** by using redundancy
- Users generate **innocuous-looking traffic**
- **No dedicated infrastructure** required

Why Might Collage Work?

- Lots of User-Generated Content (**UGC**)
  - More than 4 billion Flickr images
  - A day of video uploaded to YouTube every minute
- **Many** sites host UGC
- We have tools to **store censored data** in UGC
  - Steganography, watermarking

Collage, Step-by-Step



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 [illegible] host

- [illegible]
- [illegible] recipient should know it

Embedding Messages in Vectors

- **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

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**

Steganography: hard to detect

Watermarking: hard to remove

Do the reverse to decode

Where Are Bob's Vectors?

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

Solution: A predictable way of mapping message identifiers to subsets of content hosts

Collage steps:

1. Obtain message
2. Pick message identifier
3. Obtain cover media
4. Embed message in cover
5. **Upload UGC to content host**

---

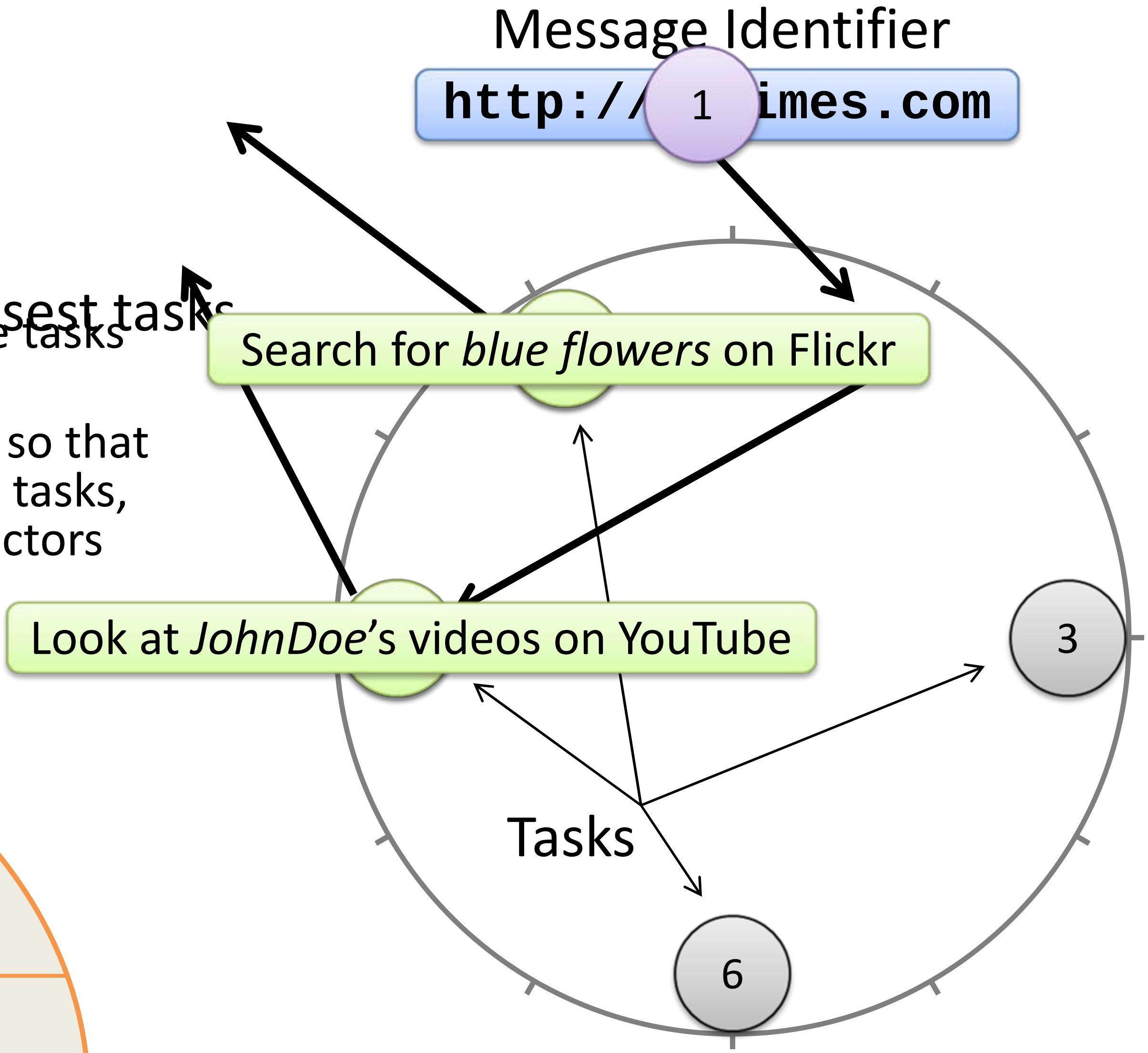
6. **Find and download UGC**
7. Decode message from UGC

Solution: Task Mapping

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

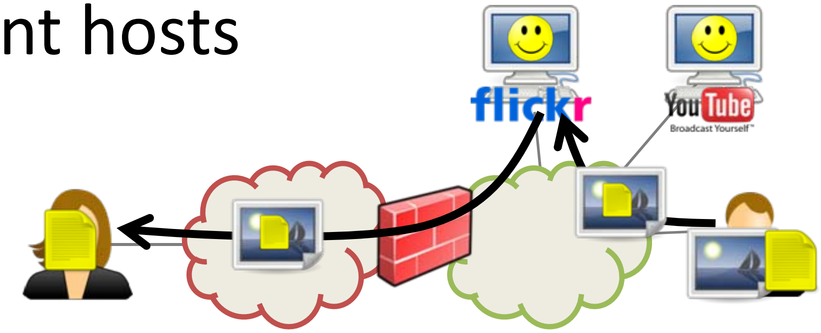

Message Identifier
http:// 1 imes.com
Search for *blue flowers* on Flickr
Look at *JohnDoe*’s videos on YouTube
3
Tasks
6

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

How Does Collage Meet the Design Goals?

- **Robust** against blocking
  - Erasure coding
  - Many content hosts
- **Deniable** against user identification
  - Traffic only to/from content hosts
  - Depends upon task construction
- Require **no dedicated infrastructure**
  - Messages stored on content hosts



Performance Metrics

- Sender and receiver **traffic overhead**
- Sender and receiver **transfer time**
- **Storage** required on content hosts

But these metrics can vary a lot:

- Different content hosts
- Different tasks
- Different applications

Case Study

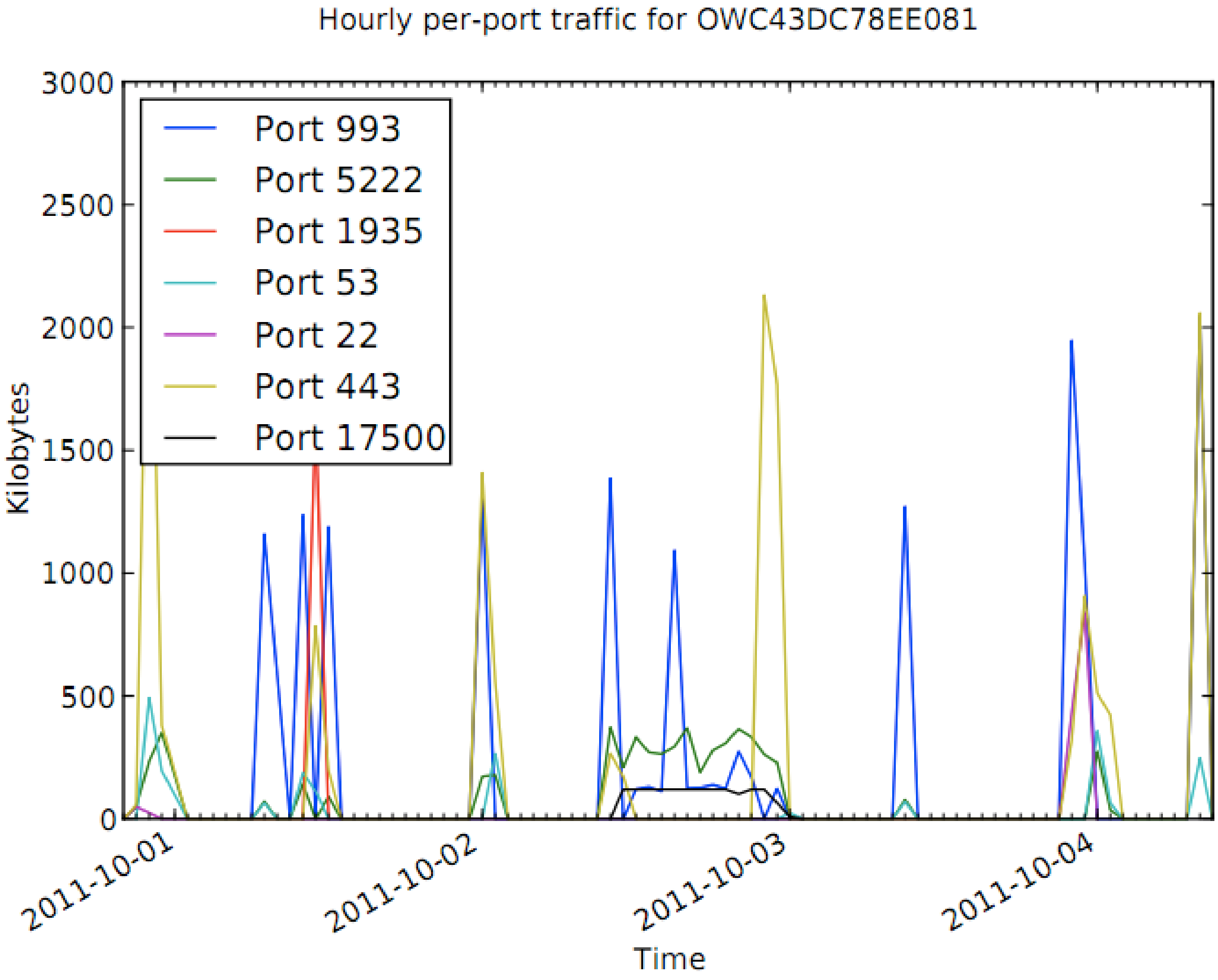
	News Articles	**Covert Tweets**
Content host	Flickr	Twitter
Message size	30 KB	140 Bytes
Vectors needed	5	30
Storage needed	600 KB	4 KB
Sending traffic	1,200 KB	1,100 KB
Sending time	5 minutes	60 minutes
Receiving traffic	6,000 KB	600 KB
Receiving time	2 minutes	½ minute

Experiments performed on a 768/128 Kbps DSL connection

Challenge: Evaluating Deniability

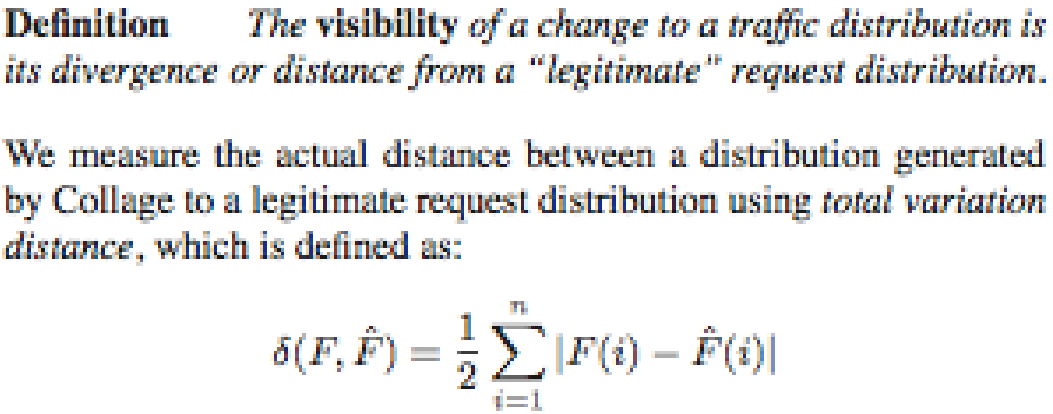
- Defining metrics
  - Formalize **statistical deniability** and develop a metric for measuring the deniability of user traffic.
  - Define distance metrics to establish a distance from normal (“visibility”)
- Establish fingerprints for normal user behavior
  - Characterize traffic mixes in different settings
    - **Home: Measurement of user behavior from home routers**
    - Enterprise/campus
    - Transit

Traffic Volumes by Port Over Time



Statistical Deniability

- **Definition:** Properties of covert traffic "match" those of legitimate traffic



Definition *The* **visibility** *of a change to a traffic distribution is its divergence or distance from a "legitimate" request distribution.*

We measure the actual distance between a distribution generated by Collage to a legitimate request distribution using *total variation distance*, which is defined as:

$$\delta(F,\hat{F})=\frac{1}{2}\sum_{i=1}^{n}|F(i)-\hat{F}(i)|$$

- Can be done by measuring distances between distributions
  - L1 Norm
  - L2 Norm
  - K-L Distance
- Can apply to different properties
- Number of distributions depends on capabilities of censor

New Challenges

- **Monitoring** censorship and surveillance
- **Manipulation** of online content (e.g., propaganda)
- **Personalization** as censorship (“filter bubbles”)

Challenges with Monitoring Censorship

- Censorship is ill-defined
- Measurements may be blocked
- Reports may be blocked
- Operating the measurement tool may be incriminating
- Reports may be falsified
- Personalization may be confused with censorship

Manipulation of Content

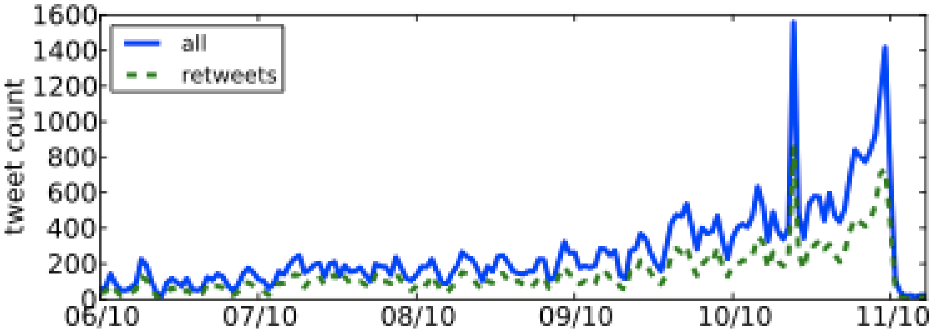
- **Sock-puppeting:** False appearance of independent speakers

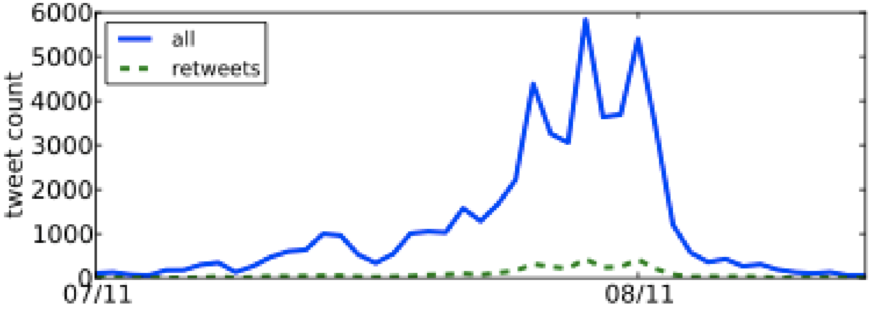
- **Astroturfing:** False appearance of a grassroots movement

Research: Detecting Propaganda

- Twitter: Medium for spreading information
- Can it be used for influencing public opinion?
- Can we detect when that is the case?

Nevada Senate Race

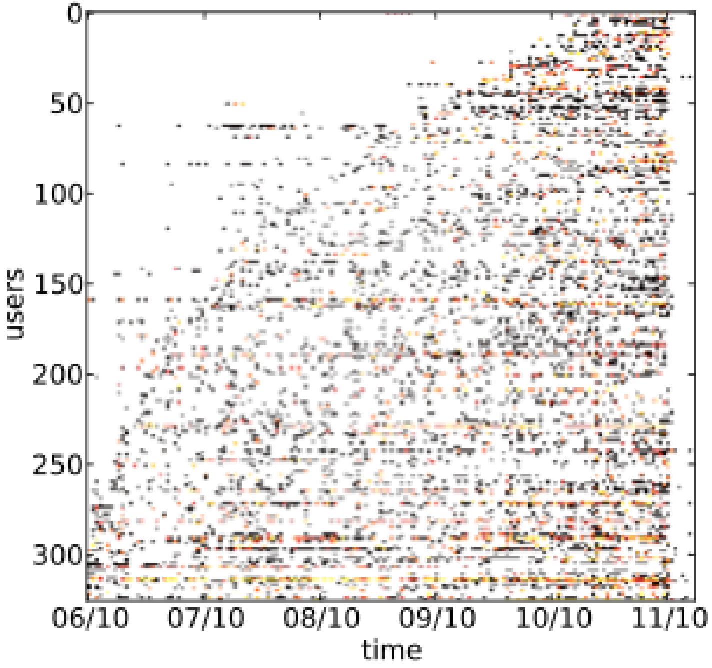


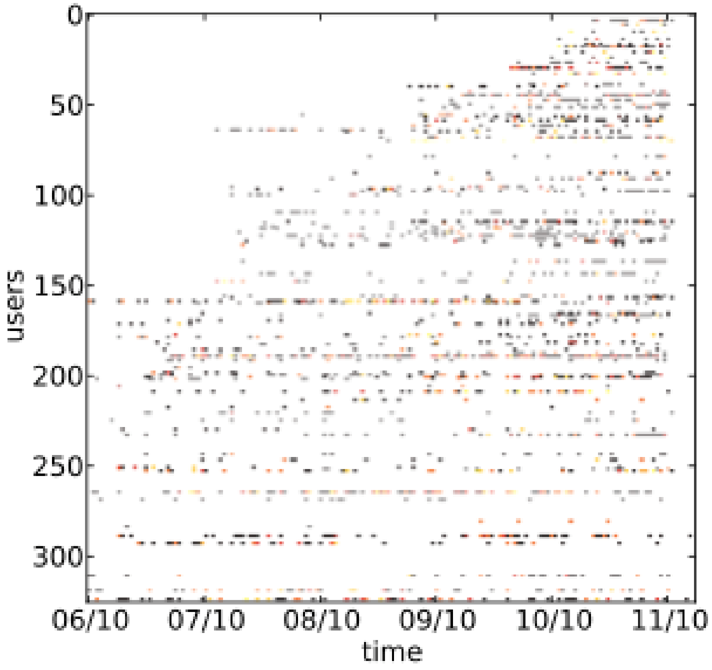


Properties of Propagandists

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

(w/Cristian Lumezanu and Hans Klein)



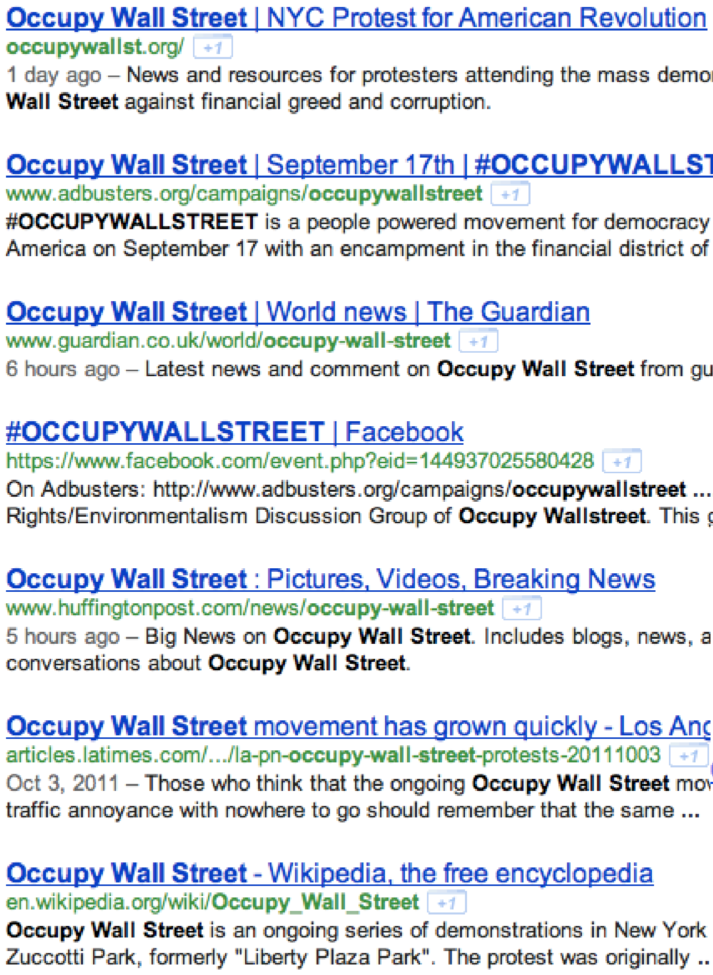


Filter Bubbles

> "A squirrel dying in front of your house may be more relevant to your interests right now than people dying in Africa" – Mark Zuckerberg

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

Bursting the Bubble: Approach



[Occupy Wall Street | NYC Protest for American Revolution](#)
occupywallst.org/
1 day ago – News and resources for protesters attending the mass demonstration on **Wall Street** against financial greed and corruption.

[Occupy Wall Street | September 17th | #OCCUPYWALLSTREET ...](#)
www.adbusters.org/campaigns/**occupywallstreet**
#OCCUPYWALLSTREET is a people powered movement for democracy that began in America on September 17 with an encampment in the financial district of ...

[Occupy Wall Street | World news | The Guardian](#)
www.guardian.co.uk/world/**occupy-wall-street**
6 hours ago – Latest news and comment on **Occupy Wall Street** from guardian.co.uk.

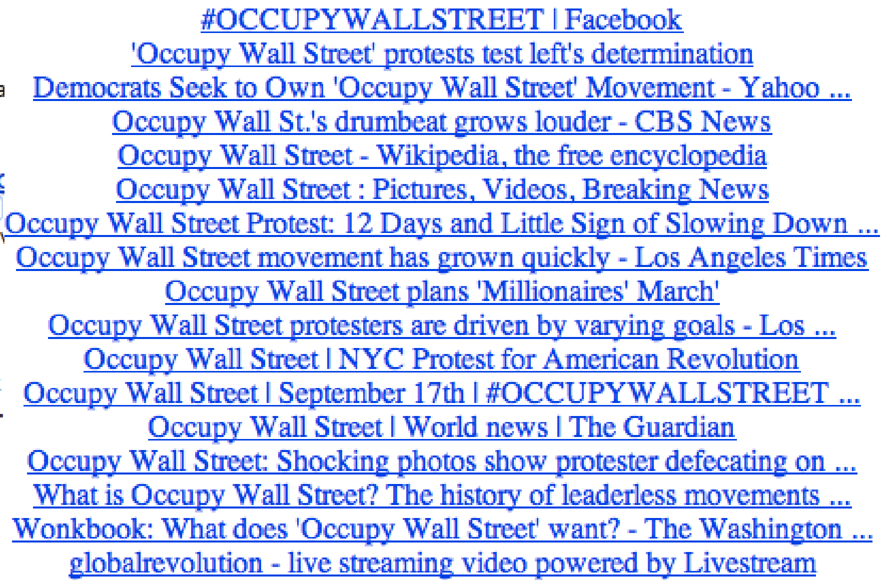
[#OCCUPYWALLSTREET | Facebook](#)
https://www.facebook.com/event.php?eid=144937025580428
On Adbusters: http://www.adbusters.org/campaigns/**occupywallstreet** .... The Animal Rights/Environmentalism Discussion Group of **Occupy Wallstreet**. This group ...

[Occupy Wall Street : Pictures, Videos, Breaking News](#)
www.huffingtonpost.com/news/**occupy-wall-street**
5 hours ago – Big News on **Occupy Wall Street**. Includes blogs, news, a
conversations about **Occupy Wall Street**.

[Occupy Wall Street movement has grown quickly - Los Anc](#)
articles.latimes.com/.../la-pn-**occupy-wall-street**-protests-20111003
Oct 3, 2011 – Those who think that the ongoing **Occupy Wall Street** mo
traffic annoyance with nowhere to go should remember that the same ...

[Occupy Wall Street - Wikipedia, the free encyclopedia](#)
en.wikipedia.org/wiki/**Occupy_Wall_Street**
Occupy Wall Street is an ongoing series of demonstrations in New York
Zuccotti Park, formerly "Liberty Plaza Park". The protest was originally ..

- When user issues search query, launch the same query from multiple locations
  - Use stateless accounts to ensure that search history does not influence results
- Compare user's results to "normalized" results



[#OCCUPYWALLSTREET | Facebook](#)
['Occupy Wall Street' protests test left's determination](#)
[Democrats Seek to Own 'Occupy Wall Street' Movement - Yahoo ...](#)
[Occupy Wall St.'s drumbeat grows louder - CBS News](#)
[Occupy Wall Street - Wikipedia, the free encyclopedia](#)
[Occupy Wall Street : Pictures, Videos, Breaking News](#)
[Occupy Wall Street Protest: 12 Days and Little Sign of Slowing Down ...](#)
[Occupy Wall Street movement has grown quickly - Los Angeles Times](#)
[Occupy Wall Street plans 'Millionaires' March'](#)
[Occupy Wall Street protesters are driven by varying goals - Los ...](#)
[Occupy Wall Street | NYC Protest for American Revolution](#)
[Occupy Wall Street | September 17th | #OCCUPYWALLSTREET ...](#)
[Occupy Wall Street | World news | The Guardian](#)
[Occupy Wall Street: Shocking photos show protester defecating on ...](#)
[What is Occupy Wall Street? The history of leaderless movements ...](#)
[Wonkbook: What does 'Occupy Wall Street' want? - The Washington ...](#)
[globalrevolution - live streaming video powered by Livestream](#)

Conclusion

- Censorship is pervasive, even in Western countries
- Deniability is an important property
  - We don’t have good ways of monitoring/quantifying it
- Monitoring censorship is important
- New forms for censorship are emerging
  - Propaganda
  - Filter bubbles