Combating web censorship through social links

Jinyang Li, Nikolaos Michalakis, Nguyen Tran, Alana Libonati
web censorship

- Large problem in many countries
  - China, Vietnam, Saudi Arabia, Syria, USA, Cuba...
- Big sites: Google, Wikipedia, MySpace, blogs...
- Reasons: political, religious, social
- Means: Telecomms block traffic, cyber police monitoring, admission control (permits)
censorship map

Countries putting obstacles to the free flow of information
Internet “Blackholes”

Source: www.rsf.org
Possible solutions

- Proxies
  - strip out your identity (anonymizer.com)
  - better proxies like Infranet also use covert channels to hide the information.
  - rely on volunteers, not too many -> easy to shut down
Possible solutions

- Anonymization Proxy Networks
- TOR and Crowds
  - few nodes, central components -> shut down
- P5, Herbivore p2p architectures
  - suffer from sybil attack-> block traffic
Possible solutions

- Censorship resistant publishing
  - Freenet, Tangler: content-oriented solution, bad for dynamic content
  - not good for general purpose browsing
Four requirements

1. Need proxies (connection-based solution) and lots of them.

2. Nothing is centralized (bootstrap, proxy discovery, software download).

3. Resilient to adversarial nodes that try to discover and block every proxy.

4. Resilient to adversary monitoring and dropping network traffic.
Requirement One

- Use p2p which has potentially many nodes, therefore many proxies.
- Central system tends to have small # of proxies.
- But why would a peer want to run a proxy?
- Idea: package with a more interesting application.
- We are building a cooperative backup application [Nguyen].
- Scenario: everyone runs the app, and some also run a proxy.
Requirement Two

- Everything must be decentralized
  - governments are good at shutting down centralized system components.
- No fixed set of bootstrapping nodes.
- Proxy discovery using p2p layer.
- Not every node is a proxy.
Requirement Three

- If the government can create many nodes it can discover many proxies.
- If one node knows many other proxies, then few bad guys can discover a lot of proxies.
  - No too many colluding bad nodes can enter.
  - No single node can know too many proxies.
Use social overlay

- Social overlay: nodes establish links to other nodes if their owners know each other in the physical world.
- Hard for a malicious entity (unfriendly government) to create many links in the network.
- Also relatively stable.
- Organic growth, no central mediator.
Setup

Diagram showing nodes P1, P2, P3, N1, N2 connected in a network.
The random route

- Approach: use random routes [SybilGuard]
Properties

- Every node learns as many proxies as friends.
- Routes are deterministic so nodes learn the same over time.
- Is this discovery load balanced?
Skewed Discovery

Increase capacity of each link.
the social network

- Using a crawler to explore links inside the Orkut social network and see if a skewed graph among countries really appears [Alana].
- Plan is also to use social network models and simulation for larger numbers of users.
Outstanding issues

- Nodes temporarily go offline.
- Changes in the social network.
- Changes in the set of proxies.
Requirement Four

- Need to hide network traffic...
  - Sender: “Hi, I’m looking for a proxy”.
  - Chinese firewall: “Your request timed out”
- Has been addressed already: Infranet uses covert channels (but slow)
- Sufficient to encrypt traffic with shared keys among friends.
Other Issues

- Strong assumption that # of links an adversary has in social network is small.
- Other compelling applications except backup?
- Friend’s data sharing.