CacheBrowser: Bypassing Chinese Censorship without Proxies Using Cached Content

John Holowczak, Hadi Zolfaghari and Amir Houmansadr
College of Information and Computer Sciences, University of Massachusetts, Amherst

Motivation

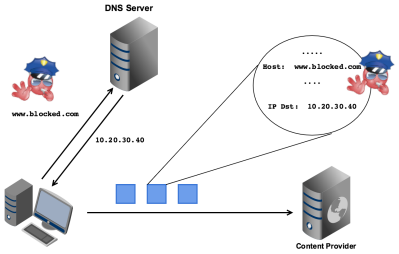
- Content caching by CDN networks poses significant technical and non-technical challenges to the censors.

In this paper:

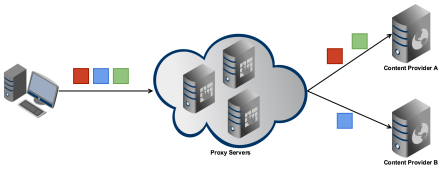
- Analyze how content cached by CDN networks can be censored.
- Design and implement a system that leverages the censorâs challenges in blocking CDN content.
- Introduce the publisher-centric approach for censorship circumvention.

Traditional Censorship Resistance

- The end-to-end communication paradigm employed in the Internet allows censors to prevent users at a *low cost* from making end-to-end connections with forbidden content publishers.
- Main censorship techniques used by censors are:
  - DNS interference
  - IP address filtering
  - Keyword/URL filtering using DPI



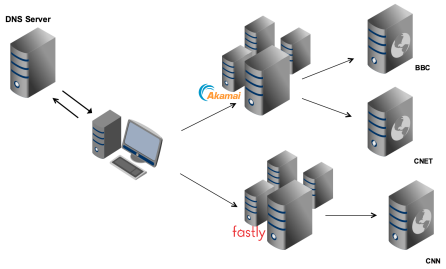
- Traditional circumvention uses a third-party **proxy** with access to the content provider.
- Client's traffic is encrypted and relayed through one or more proxies.



- Proxy-based circumventions is used by systems such as Tor and PSIPHON.
- The countermeasure used against proxy-based circumvention systems is to locate and block the proxy servers.

Our Approach

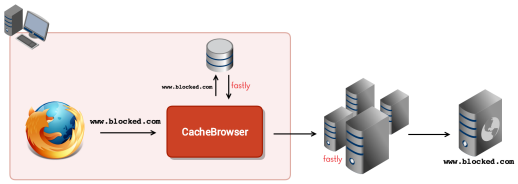
- Content Delivery Networks are becoming prevalent in the Internet.
- DNS requests for a CDN hosted domain return a CDN edge server address.
- Every CDN edge server responds to requests for many websites.



- IP Address blocking of CDN edge servers could result in high **collateral damage**.
  - The main method used to block CDN content is DNS interference.
- If a website is hosted on a CDN such as Akamai, we can request that website any edge server belonging to the CDN.

CacheBrowser:

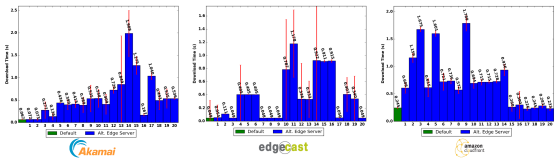
- Request blocked website from CDN edge server without DNS requests.



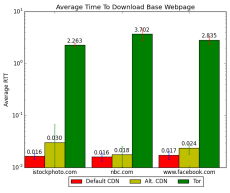
Results

Performance:

- Edge servers are not chosen by CDN mapping system.
  - CacheBrowser is slower than ordinary web browsing.

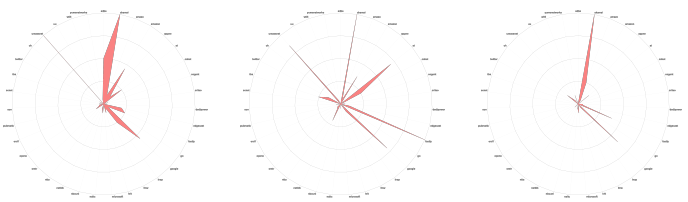


- No third-party proxy is used.
  - CacheBrowser is faster than proxy-based circumvention systems.



Future Work

- CacheBrowser reveals connections to different servers for loading resources/ads.
  - Censor could apply website fingerprinting based on server connections.
  - CacheBrowser does not provide privacy from the censor.
- Countermeasures against website fingerprinting:
  - Increasing the use of cached content by the browser.
  - Loading or not loading resources to make it look like a non-blocked website.
  - Server-side shaping of the website to make it look like a non-blocked website.
- Is website fingerprinting based on server connections feasible?



(a) www.nbc.com (b) www.cnn.com (c) www.bbc.com

Figure: Website fingerprints based on connections made

The work is generously support by the NSF grant CNS-1553301

jholowcz@umass.edu, hadi@cs.umass.edu, amir@cs.umass.edu