Communication Without Boundaries

Breaching the Great Firewall of China

George Louthan John Hale

iSec™

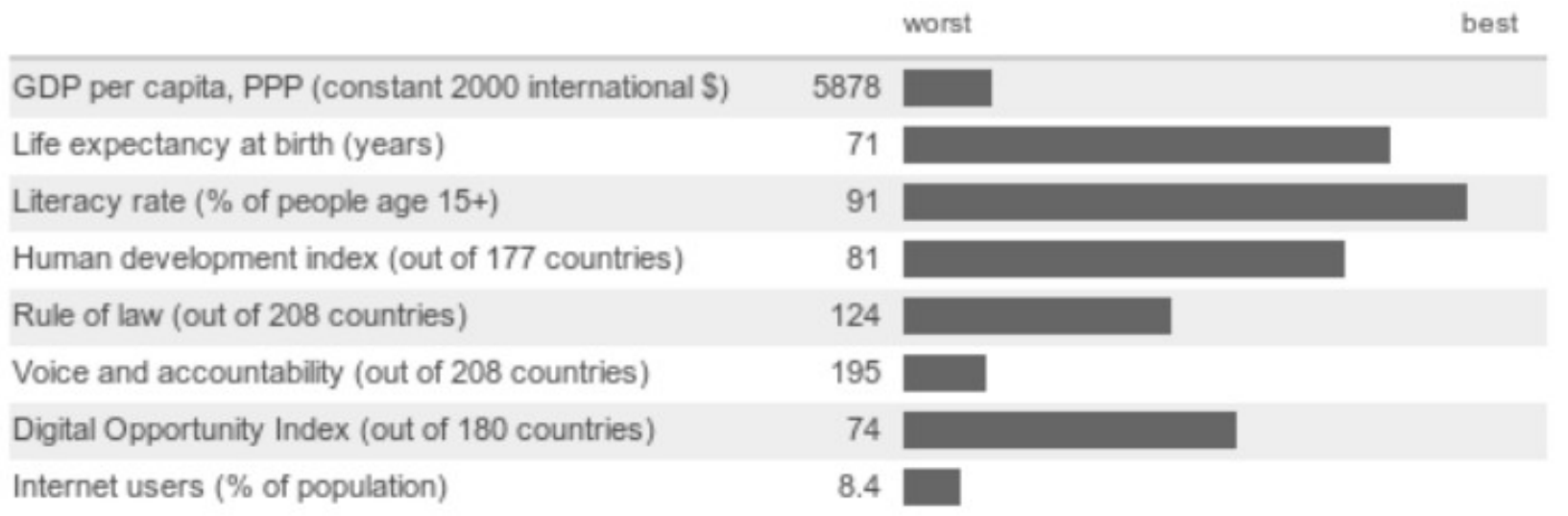
The University of Tulsa
Institute for Information Security

Overview

- About Internet in the People's Republic of China
- Overview of the Great FW of China
  - What it does
  - How it works
- Previous work in defeating it
  - Shortcomings
- Proposal for new methods
- Call to action

About the PRC

- 8.4% of the Chinese population is online
  - That's 137 million people.
- OpenNet Initiative's statistics:



Source: OpenNet Initiative, http://opennet.net/research/profiles/china

The Great Firewall of China

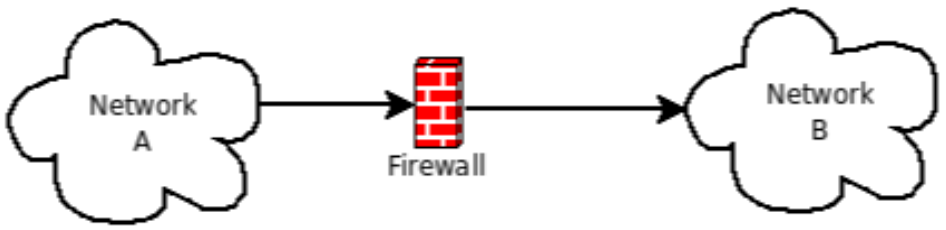
- Officially *Golden Shield Project*
- Blocks "harmful" content
  - Pornography
  - Political sites
  - Criminal activities
  - Full list of prohibited content available at http://tinyurl.com/cwudp6
- Not really a firewall
  - More aptly, a set of technical controls working together to censor Chinese Internet traffic
  - Includes some conventional firewall-like aspects

Chinese Methods for Censoring Traffic

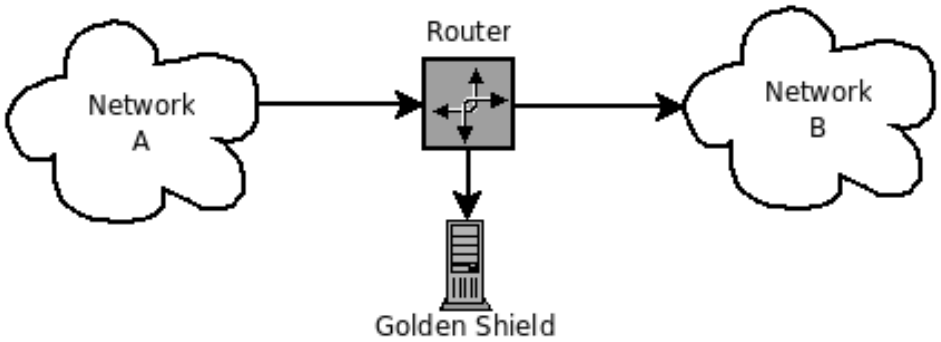
- DNS hijacking
  - Chinese DNS servers convert blacklisted URLs into incorrect IP addresses
- IP Address Blocking
  - Access to some IP addresses is blocked outright
- Content Filtering (Primary topic of discussion)
  - URL keywords
    - Connections may be broken when URLs containing certain keywords are detected
  - Response keywords
    - Same, but for actual content of web pages

Topology – Not a Real Firewall

- Traditional Firewall
  - Sits between networks, actually filtering traffic

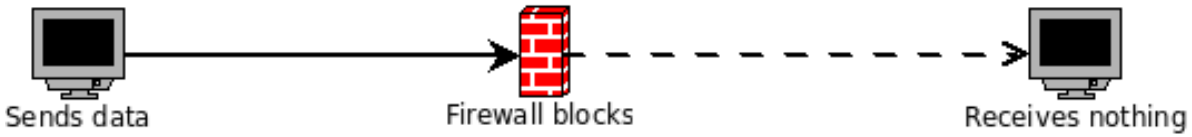


- Golden Shield
  - Examines and interacts, but does not filter, *per se*

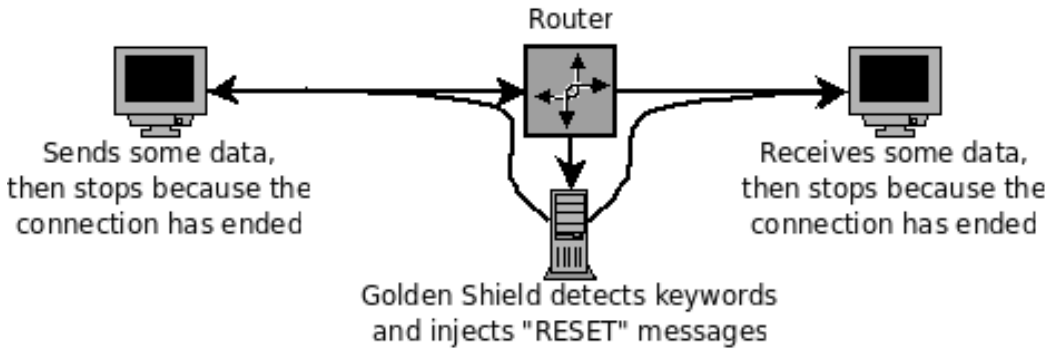


Stopping traffic – Not a Real Firewall

- Traditional Firewall
  - Blocked traffic is not allowed to pass through



- Golden Shield
  - Violators receive spoofed “connection reset” messages



Getting through: Existing methods

- Defeating DNS hijacking
  - Use non-Chinese DNS servers
- Defeating IP Address Blocking
  - Go through a proxy
  - Onion routing (e.g. TOR)
- Defeating Content Filtering
  - Adjust MTU (maximum transmission unit)
  - Escape characters
  - Ignore “reset” packets

Problems with existing methods

- Require tech savvy on the inside
  - Alternate DNS servers, proxies, MTU adjustment, onion routing, ignore “reset” packets
  - That's all but one of the methods listed on previous slide
- Easily defeated
  - Alternate DNS servers, proxies, escape characters
- Require local configuration that's detectable
  - Alternate DNS servers, proxies, MTU adjustment, onion routing, ignore “reset” packets
  - Potentially incriminating
  - Not always possible (Internet cafes, for example)

Proposal for new methods: Requirements

- Robust
  - Should take a lot of work to defeat
  - Need rough parity between effort to implement and effort to defeat
- Transparent to users inside China
  - Sites should be accessible as normally as possible
  - Nothing requiring technical know-how
- Undetectable locally
  - No unusual configuration or installation of software on computers inside China
- Prevent flow of prohibited keywords BOTH into and out of China

Proposal: 4 research topics

- Prohibited keyword discovery
  - Figure out precisely what is filtered
- Site migration
  - Development of a surreptitious, as-transparent-as-possible means for migrating a web location's IP and/or URL
  - Hard!
- Automated site auditing
  - Ensure navigation of a site does not require transmission of keywords (provides “inside-to-outside” data flow safety)
- Selective TCP segmentation
  - Ensure that packet breaks fall in the middle of keywords
  - Construct a device that “sanitizes” web data headed for China (provides “outside-to-inside” data flow safety)

Automated site auditing

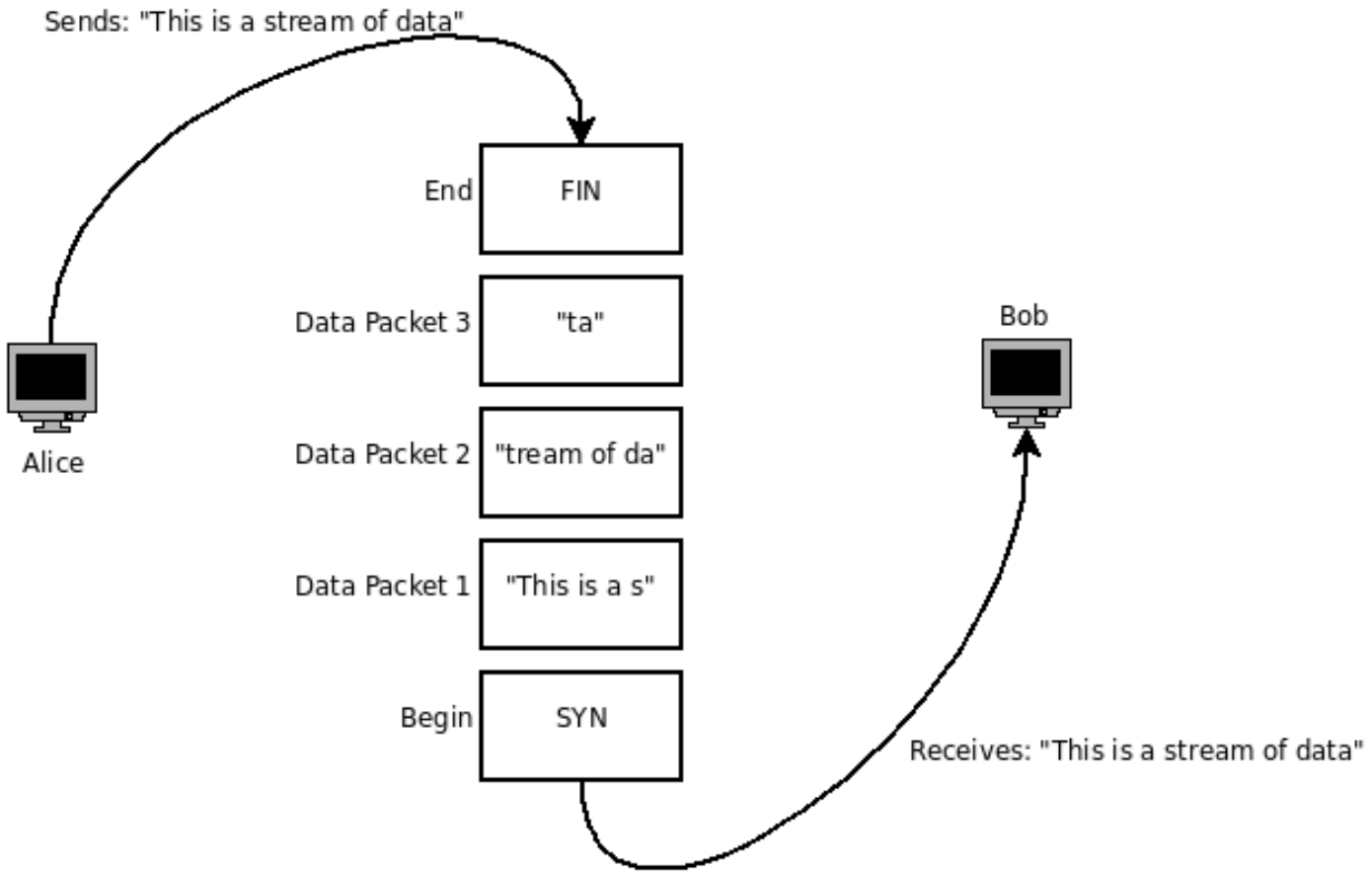
- The problem:
  - Search Google for “Tienanmen Square Massacre”

- The solution
  - Create an automatic tool to audit a site for whether keywords can appear in URLs
  - Harder problem than it sounds to check for this and offer alternative design decisions
  - Even a “best practices” document for minimizing inside-China to outside-China transmission of keywords would be valuable

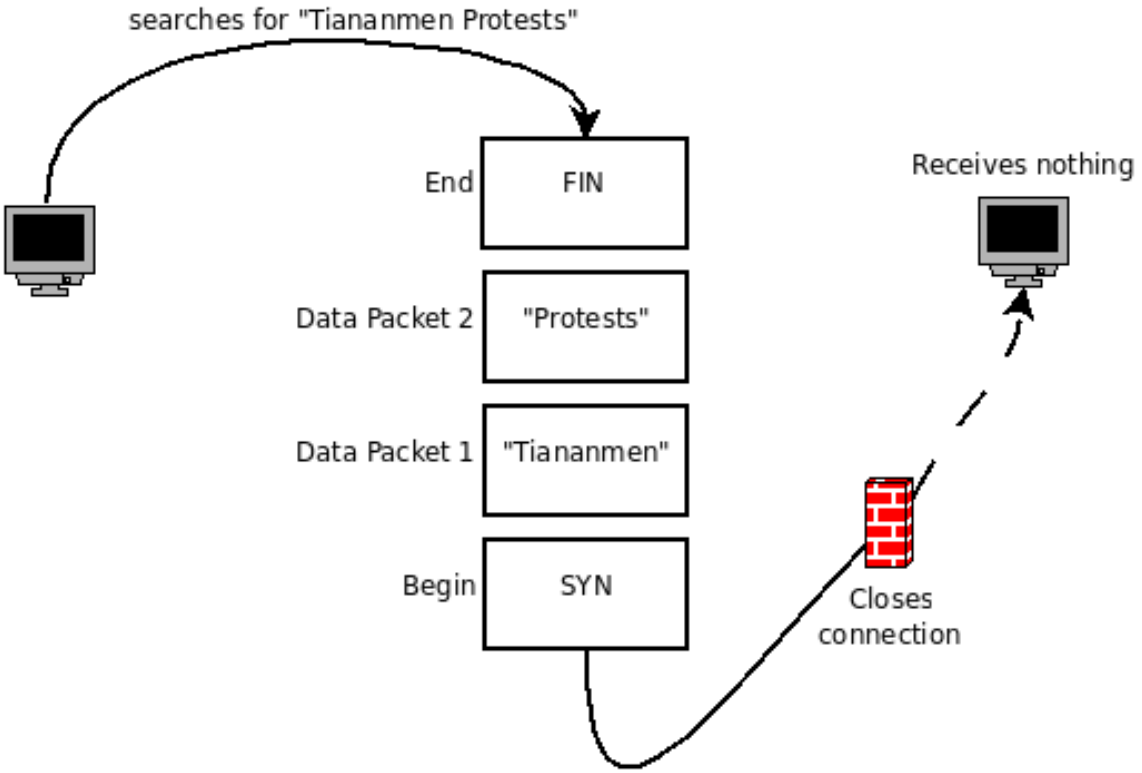
Selective Segmentation

- WWW uses the Transport Control Protocol (TCP)
  - TCP packs “stream” data into discrete “packets”
  - Drawing the dividing lines is called “segmentation”
  - No matter where the lines are drawn, the recipient reassembles it to the same stream – easy on a small scale.
- But large-scale reassembly into streams is hard
  - Easier to search packets individually than the stream
  - Packets are pretty big, so this usually works fine
  - But if a packet break falls in the middle of a keyword, then this method misses it
  - Can we automate this?

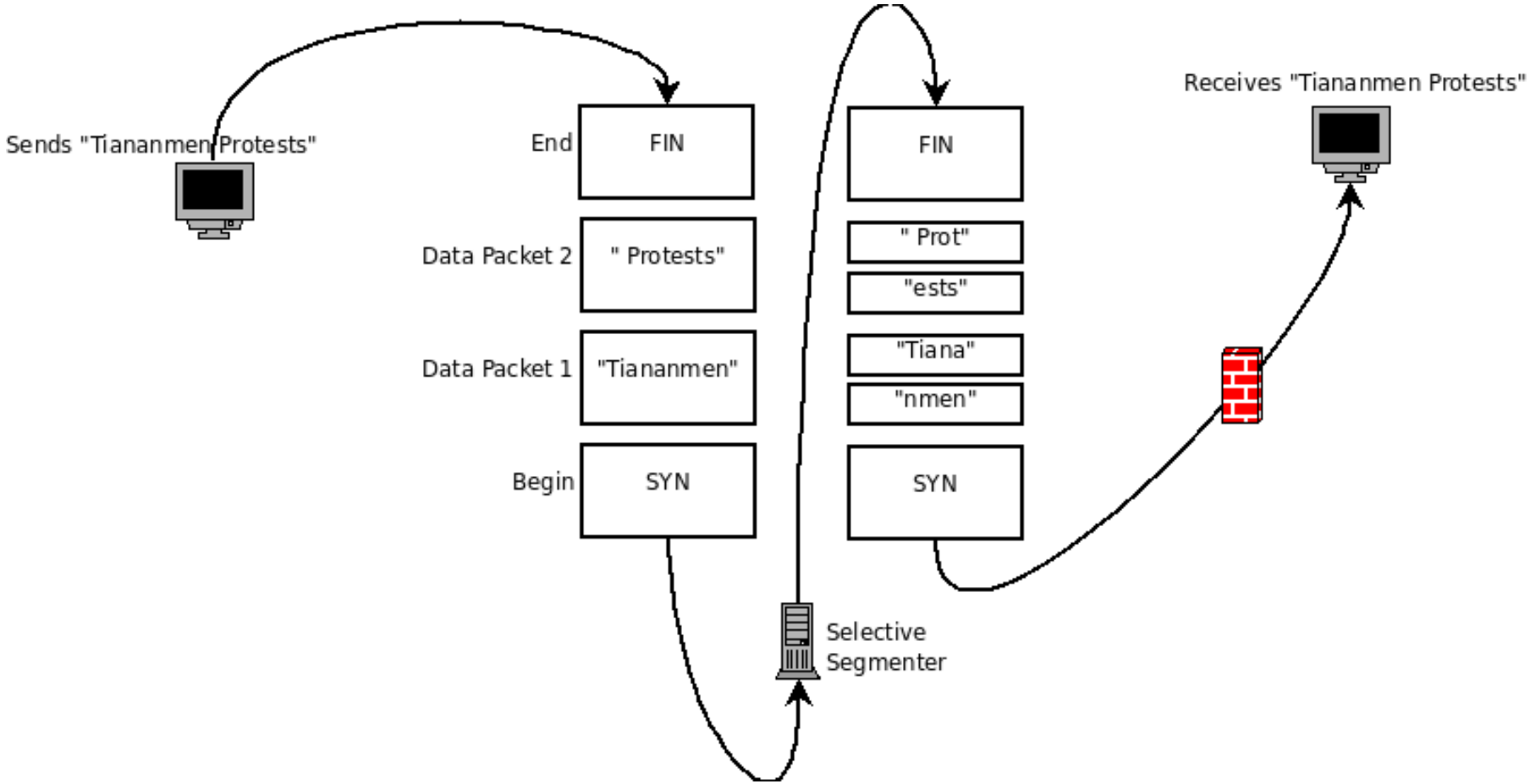
TCP traffic, normally



TCP traffic, with content filtering



TCP traffic with content filtering and selective segmentation



Take a stand.

"The Net interprets censorship as damage and routes around it"

- John Gilmore, co-founder of the Electronic Frontier Foundation

"Did you ever hear anyone say, 'That work had better be banned because I might read it and it might be very damaging to me'?"

- Joseph Henry Jackson

References and further reading

- Bennett Haselton, “List of possible weaknesses in systems to circumbent Internet censorship”. http://www.peacefire.org/circumventor/list-of-possible-weaknesses.html
- Jonathan Zittrain and Benjamin Edelman, “Empirical Analysis of Internet Filtering in China”. Available at http://cyber.law.harvard.edu/filtering/china/
- OpenNet Initiative (opennet.net)
- Congressional Executive Commission on China translation of Provisions on the Administration of Internet News Information Services. http://www.cecc.gov/pages/virtualAcad/index.phpd?showsingle=24396

References and further reading

- Rebecca MacKinnon. *YaleGlobal*. “China's Internet: Let a Thousand Filters Bloom”
- Clayton, Murdoch, and Watson. *Lecture Notes in Computer Science*. “Ignoring the Great Firewall of China.”
- Thomas Lum. “Internet Development and Information Control in the People's Republic of China”. Congressional Research Service, Library of Congress

Q & A Session

Source: http://wikileaks.org/wiki/Wikileaks_and_Internet_Censorship_-_a_comparative_study