Do You See What I See?
Differential Treatment of Anonymous Users

Sheharbano Khattak (University of Cambridge)
David Fifield (UC Berkeley)
Sadia Afroz (ICSI)
Mobin Javed (UC Berkeley)
Srikanth Sundaresan (ICSI)
Vern Paxson (UC Berkeley, ICSI)
Steven J. Murdoch (University College London)
Damon McCoy (ICSI)
How Regular Users See the Web
How Tor Users See the Web

Access Denied
You don't have permission to access "http://www.foxnews.com/" on this server.
Reference #18.4d6e5668.1445538492.20b9c5ab

Error 403
Access denied. Your IP address [171.25.193.132] is blacklisted. If you feel this is in error please contact your hosting providers abuse department.

We're sorry, but we could not fulfill your request for / on this server.
You do not have permission to access this server. Before trying again, run anti-virus and anti-spyware software and remove any viruses and spyware from your computer.
Your technical support key is: 591f-3905-2b02-1b1f
You can use this key to fix this problem yourself.
If you are unable to fix the problem yourself, please contact webmaster at monticello.org and be sure to provide the technical support key shown above.
Difference w/ Traditional Censorship

User-side Censorship
Difference w/ Traditional Censorship

User-side Censorship

Publisher-side Censorship
How Do Websites Block Tor?

Entry  Middle  Exit

abc.com
How Do Websites Block Tor?

[Diagram showing Tor network with labeled nodes: Entry, Middle, Exit, and a server.

Publicly known]
Measuring Tor Blocking by the Web

- Network layer blocking
- Application layer blocking
Network-layer Discrimination
Does An IP Address Block Tor?

**SYN (port 80)**

**SYN-ACK**

**SYN (port 80)**

**RESET / NO RESPONSE**
Measuring Tor Blocking at Scale

- IPv4 ~ over 3 billion addrs
- 4 Tor Exit Nodes (USA, Romania, Netherlands)
- 3 Control Nodes (Michigan, Cambridge, Berkeley)
..But What is The Web?

- Web Footprint—a set of IP addresses that respond successfully to our control scans on port 80
Challenges in Defining The Web

• What if a probe or response is lost?
  ✤ Redundant probing

• Temporal and spatial churn in the Web Footprint:
  ✤ Lax Web Footprint: IP addresses for which all control nodes see a response at least once (~96% of Web Footprint)
  ✤ Strict Web Footprint: IP addresses for which all control nodes received a successful response on all days (~50% of Web Footprint)
Challenges in Defining The Web

• What if a probe or response is lost?
  ✤ Redundant probing

• Temporal and spatial churn in the Web Footprint:
  ✤ **Lax Web Footprint**: IP addresses for which all control nodes see a response at least once (~96% of Web Footprint)
  ✤ **Strict Web Footprint**: IP addresses for which all control nodes receive a successful response on all days (~50% of Web Footprint)
At least 1.2% of the Web blocks Tor
AS distribution of Top 5 Tor Blockers (Lax Footprint)
AS distribution of Top 5 Tor Blockers (Strict Footprint)

ASN distribution (top 5) of IP addresses that block Tor across exit nodes.

Exit Node (% of Strict Footprint that blocks the exit)

- Axigy1 (1.91%)
- Axigy2 (1.23%)
- NForce2 (2.55%)
- Voxility1 (1.82%)

ASN distribution (%) of IP addresses that block the exit

- MCCI-AS,IR
- RMH-14-Rackspace,US
- RACKSPACE-Rackspace,US
- DREAMHOS...
- Rackspace Lt...
- KUNET-AS,KR
- REDSTATION...
- SINGLEHOP...
- BBIL-AP BH...
- OCN NTT,JP
- GO-DADDY-
- FBDC FreeBi...
Geo Distribution of Top 5 ASes that do wholesale Tor blocking
Application-layer Discrimination
Does a Website Block Tor?

HTTP GET

HTTP GET

HTTP GET
Does a Website Block Tor?
Does a Website Block Tor?

Berkeley

HTTP GET

200 OK

All Tor Exits (~900)

HTTP GET

Not 200

Alexa Top 1000
3.67% of Alexa Top 1k block Tor

Sorry, you're not allowed to access this page.

Your IP address is: 46.28.110.136

Please retry your request and contact Yelp if you continue experiencing issues.

This IP has been automatically blocked.
If you have questions, please email: blocks-b1451523930

Request denied

Sorry, we are unable to serve your request at this time due to unusual traffic from your network connection.

Please visit our help page and provide the information below for further assistance.

Reason codes:
3.67% of Alexa Top 1k block Tor

- "You don’t have permission to access this website"
- Shows CAPTCHA
LOOK AT ALL THE CAPTCHAS I'VE TO TYPE

WHEN USING TOR
#dontblocktor

**LOOK AT ALL THE CAPTCHAS I'VE TO TYPE**

On February 9th, 2016 Anonymous said:
Tor is DEAD!
And Cloudflare KILLED it!
RIP Tor!

reply
#dontblocktor

LOOK AT ALL THE CAPTCHAS I'VE TO TYPE

On February 9th, 2016 Anonymous said:
Tor is DEAD!
And Cloudflare KILLED it!
RIP Tor

On January 24th, 2016 Anonymous said:
Every time I use Tor I want to KILL everyone who works for Cloudflare.

WHEN USING TOR
On February 9th, 2016 Anonymous said:
Tor is DEAD!
And Cloudflare KILLED it!
RIP Tor

On January 24th, 2016 Anonymous said:
Every time I use Tor I want to KILL everyone who works for Cloudflare.

On February 9th, 2016 Anonymous said:
What is the point of Tor any more when it seems like the whole web is on Cloudflare!?
last night I had a dream that there were two new ways to solve cloudflare captchas: writing sentences w/ emoji & proving \( P = NP \)
How many of the ~900 Tor exits are blocked?
~20 of Alexa top 1k websites block > 50% of the exits

~60 of Alexa top 1k websites block < 25% of the exits
Why do exits get blocked?

- Two flavours:
  - Web services use Tor specific blacklist
    - Block all the Tor exits
  - Web services use abuse-based blocking
    - Block only exits with high abuse rate
Which exits are likely to have high abuse rate?

- Our hypothesis: high bandwidth and old age
Which exits are likely to have high abuse rate?

- Our hypothesis: high bandwidth and old age
- No statistically significant effect!
  - Except for few …
Which exits are blocked?
Old and high bandwidth

Tor blocked by 4chan.org
Which exits are blocked?
Old and high bandwidth

Tor blocked by change.org
Akamai blocks most exits

Tor blocked by bestbuy.com

Fraction of webpages blocked

Exit probability
Homepage unblocked but blocked activity

- Google homepage was never blocked but searching was blocked from 23-40% of the ~900 exits.

Response to https://www.google.com/#q=hello
Exits that were never blocked

- 42 exits were never blocked
Exits that were never blocked

- 42 exits were never blocked

Uptime of one of the 42 exits
Historical Tor Blocking

- **Open Observatory Network Interference (OONI)**
  - Studies censorship in different countries
  - Visits website through Tor and without Tor
  - Over 2300 websites visited (Sep’14-Aug’15)

explorer.ooni.io
6.8% of 2300 websites blocked Tor

Tor blocking rate over time

- timeout
- CloudFlare
- all others

Fraction of blocked requests

Sites that explicitly block Tor

- Convio: Not Implemented Tor IP not allowed
- ezinearticles.com

It appears that you are using Tor anonymizing software
No Problem! We just need you to enter a Captcha so we can confirm that you are a person and not a bot.
Meanwhile at CloudFlare..

@tab2space @xxdesmus soon there will be a customer option to whitelist (or blacklist) all Tor nodes. Same as how we allow with countries.
Solution?

- Contextual awareness
- Redesigning anonymity networks
Solution?

- Contextual awareness
- Redesigning anonymity networks

- Anonymous blacklisting
- Redesigning automated abuse-based blocking
Summary

• At least 1.2% of the Web block Tor (n/w)

• At least 3.67% of Alexa top 1k sites block Tor (app)

• Fine-grained discrimination?

• Who else is subject to this kind of discrimination?
Thanks

Q&A

Sheharbano.Khattak@cl.cam.ac.uk

sadia.afroz@berkeley.edu