Tor and circumvention: Lessons learned

Roger Dingledine
The Tor Project
https://torproject.org/

What is Tor?

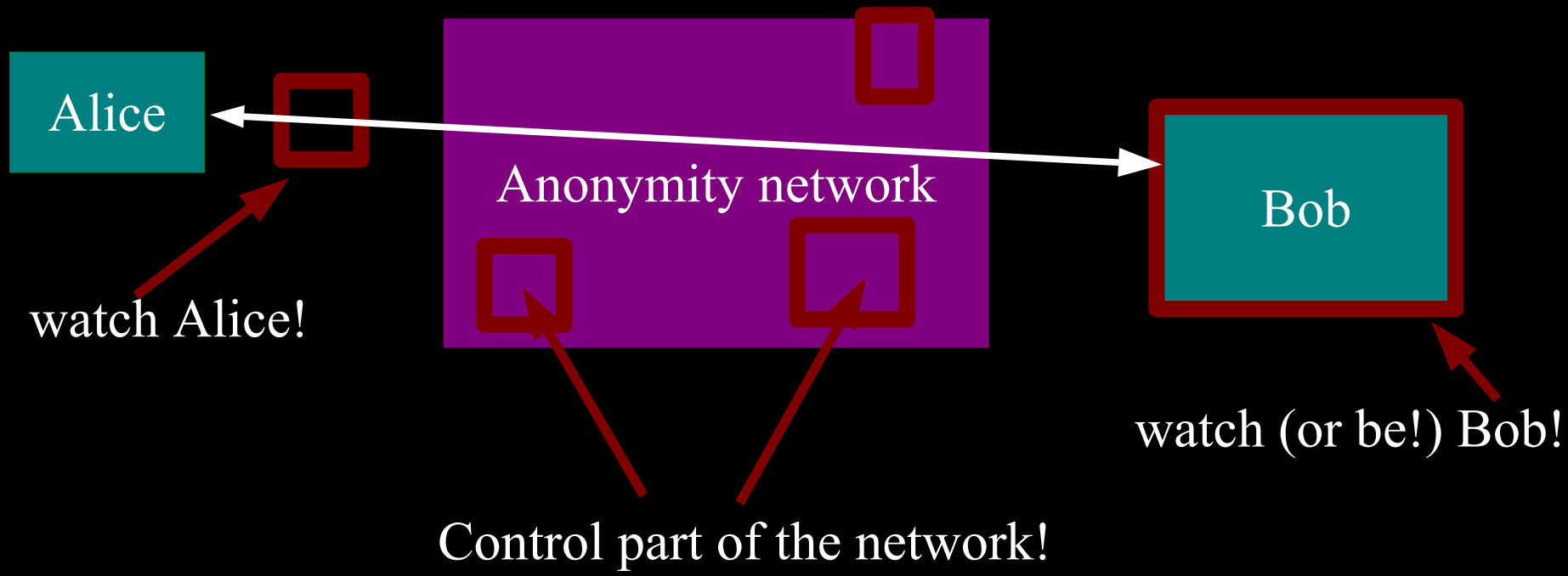
- Online anonymity 1) software, 2) network, 3) protocol
- Open source, freely available
- Community of researchers, developers, users, and relay operators
- Funding from US DoD, Electronic Frontier Foundation, Voice of America, Google, NLnet, Human Rights Watch, ...

The Tor Project, Inc.

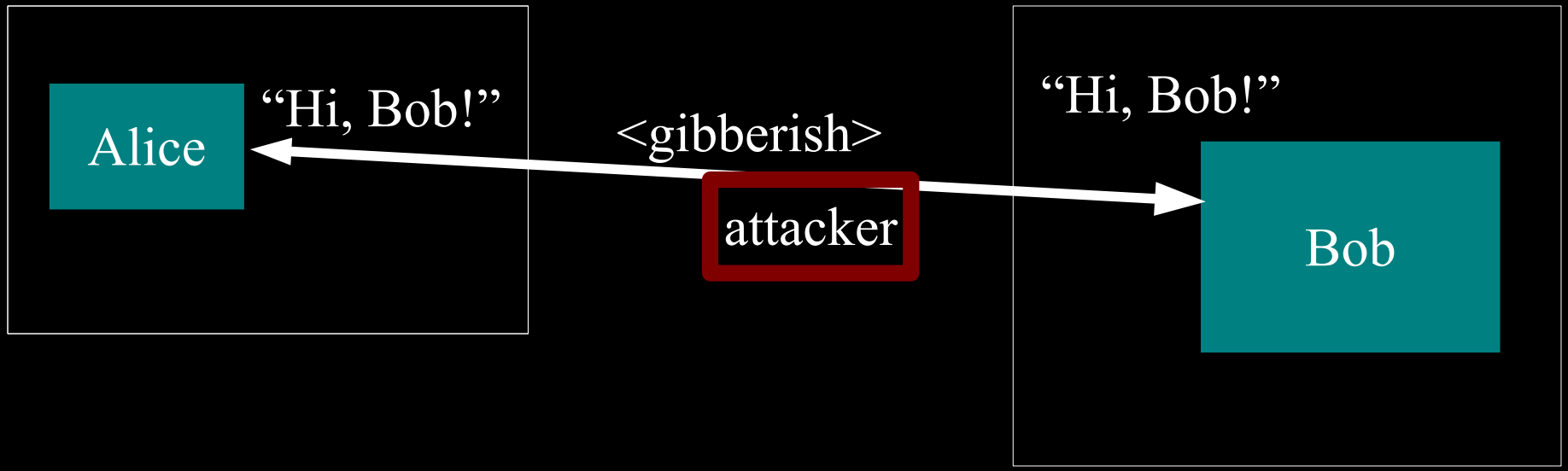
- 501(c)(3) non-profit organization dedicated to the research and development of tools for online anonymity and privacy


Estimated 500,000
daily Tor users

Threat model: what can the attacker do?



Anonymity isn't cryptography: Cryptography just protects contents.



Anonymity isn't just wishful thinking...

“You can't prove it was me!”

“Promise you won't look!”

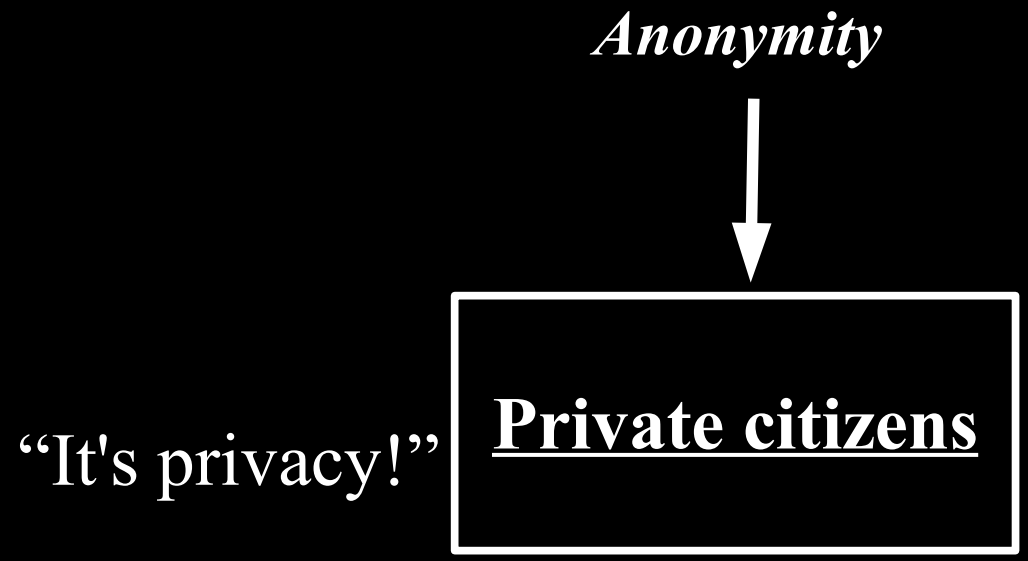
“Promise you won't remember!”

“Promise you won't tell!”

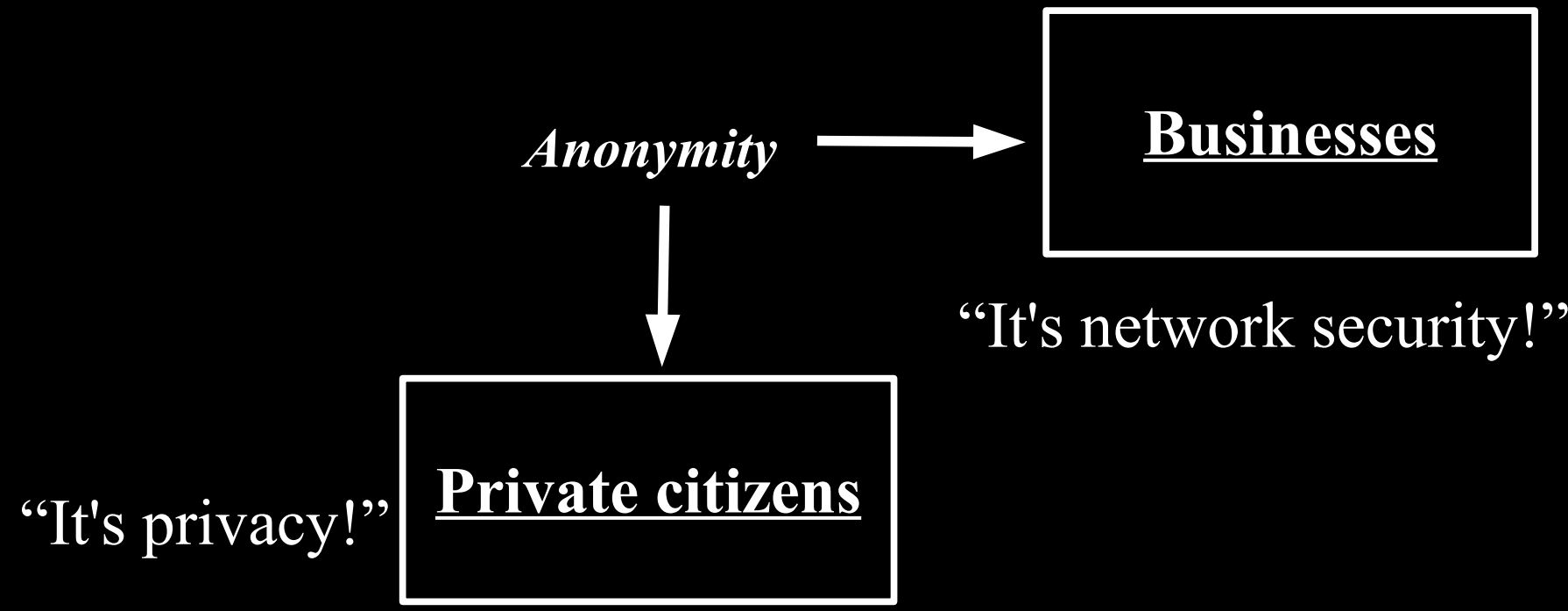
“I didn't write my name on it!”

“Isn't the Internet already anonymous?”

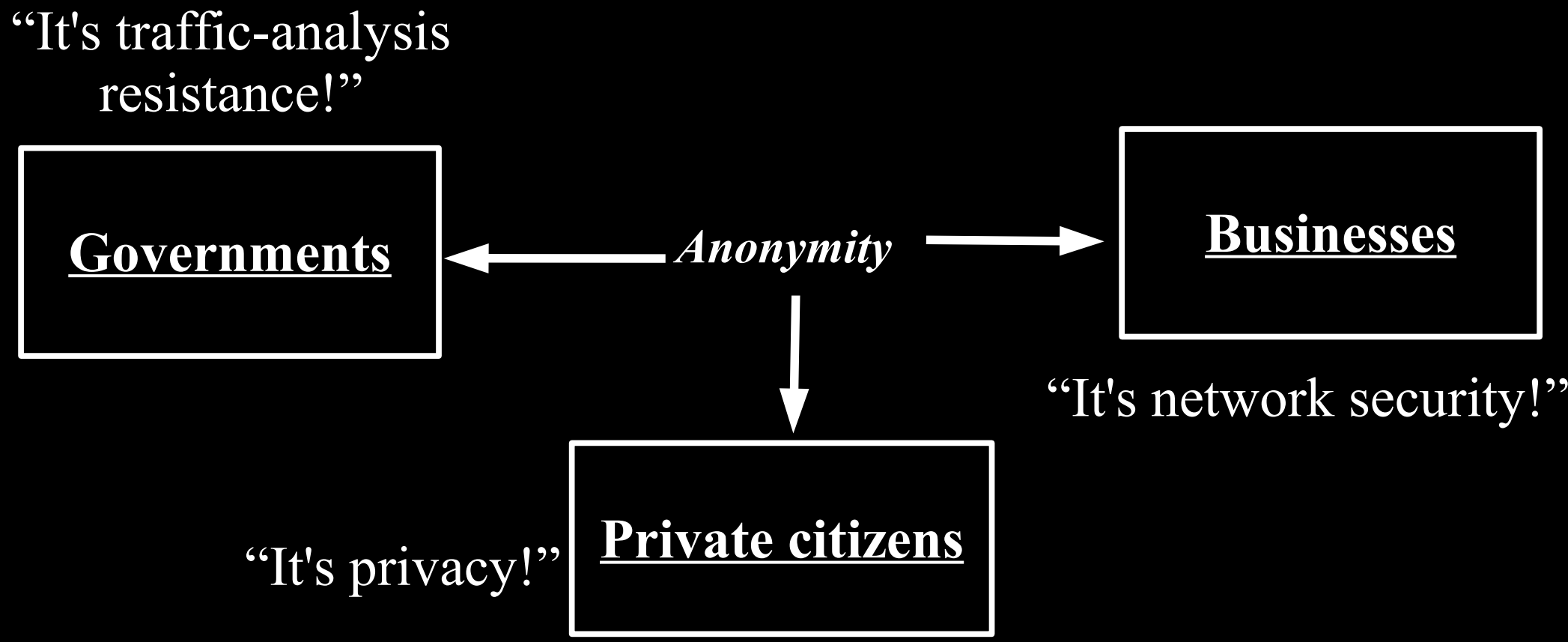
Anonymity serves different interests for different user groups.



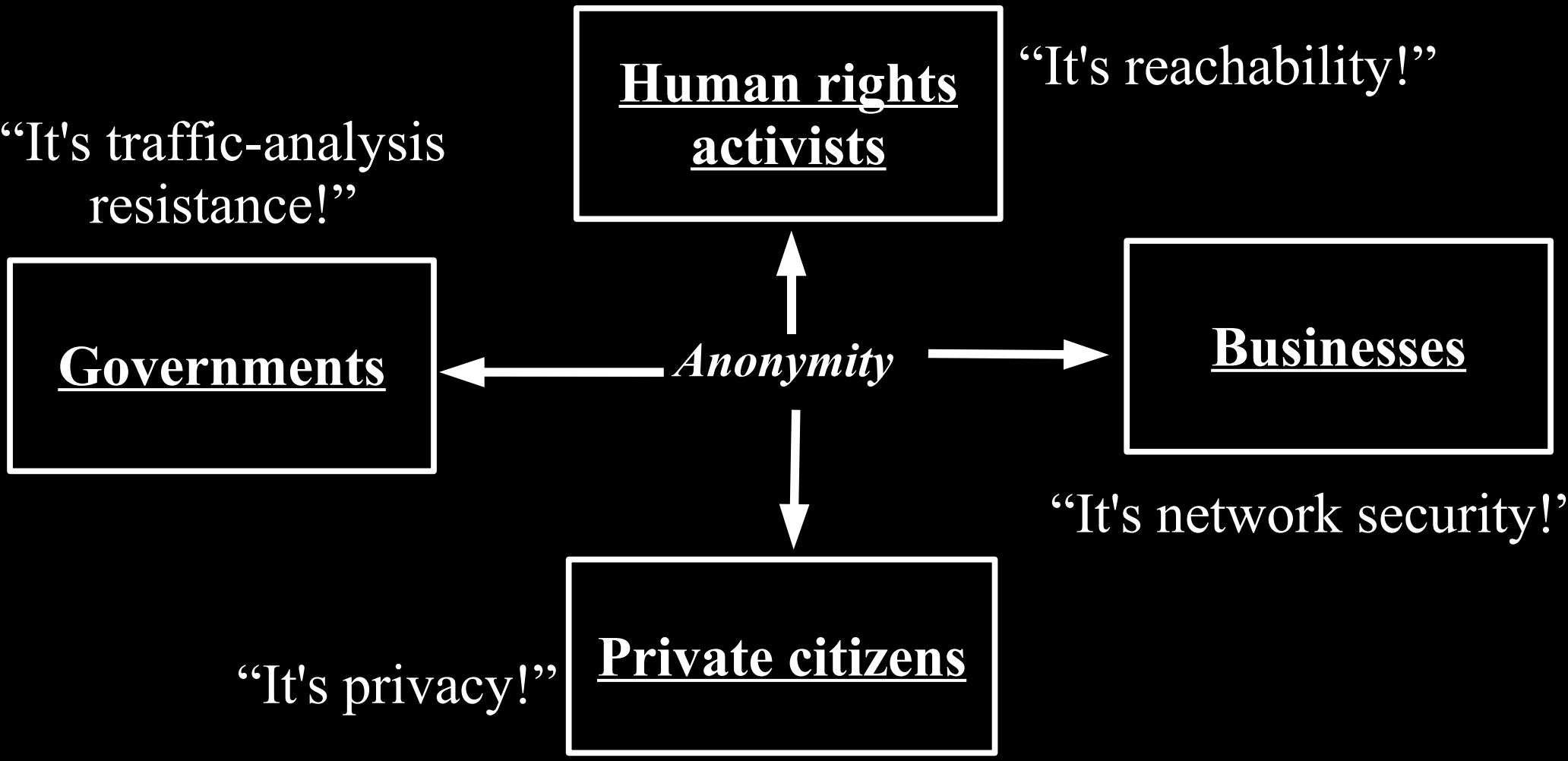
Anonymity serves different interests for different user groups.



Anonymity serves different interests for different user groups.



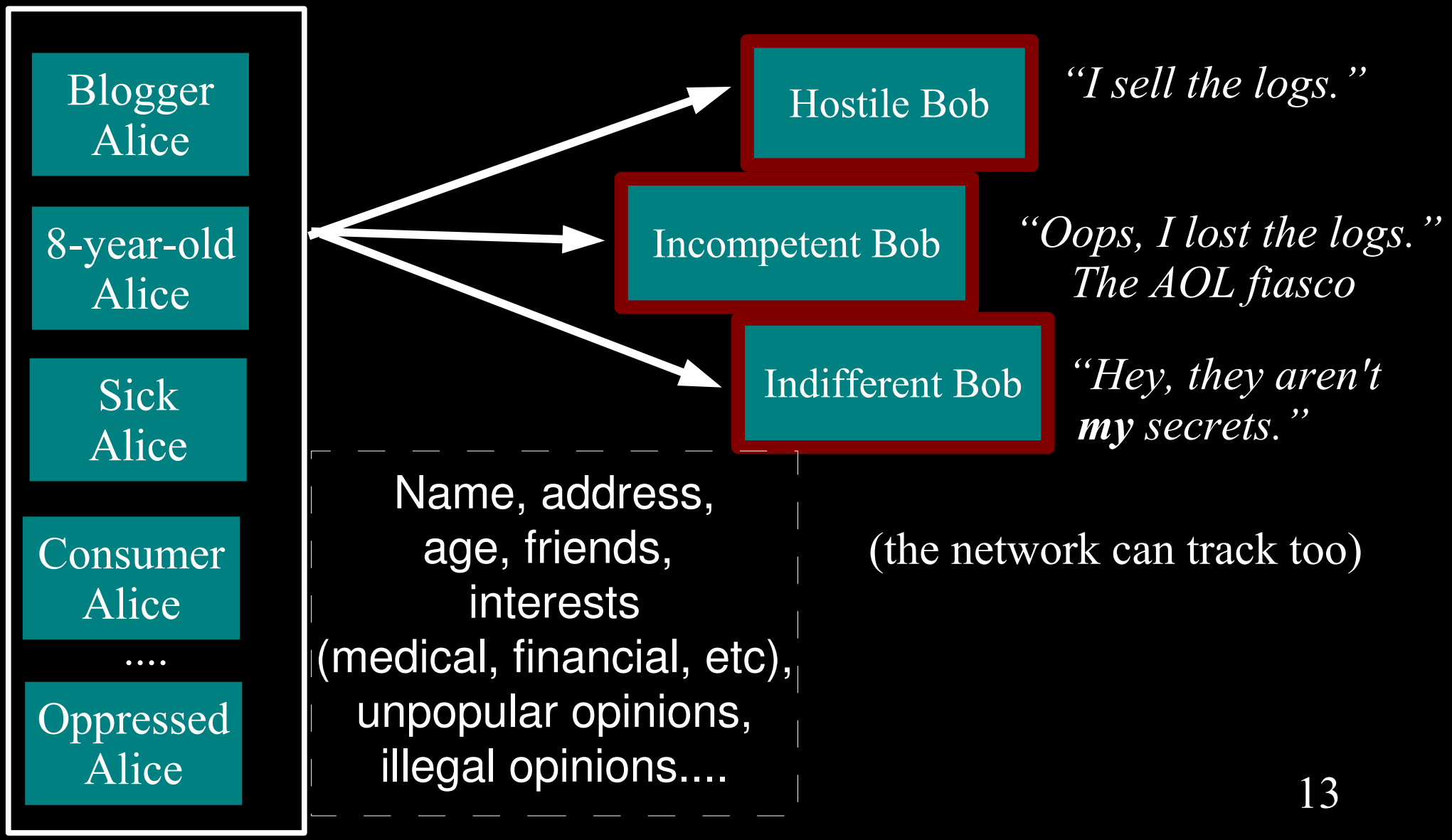
Anonymity serves different interests for different user groups.



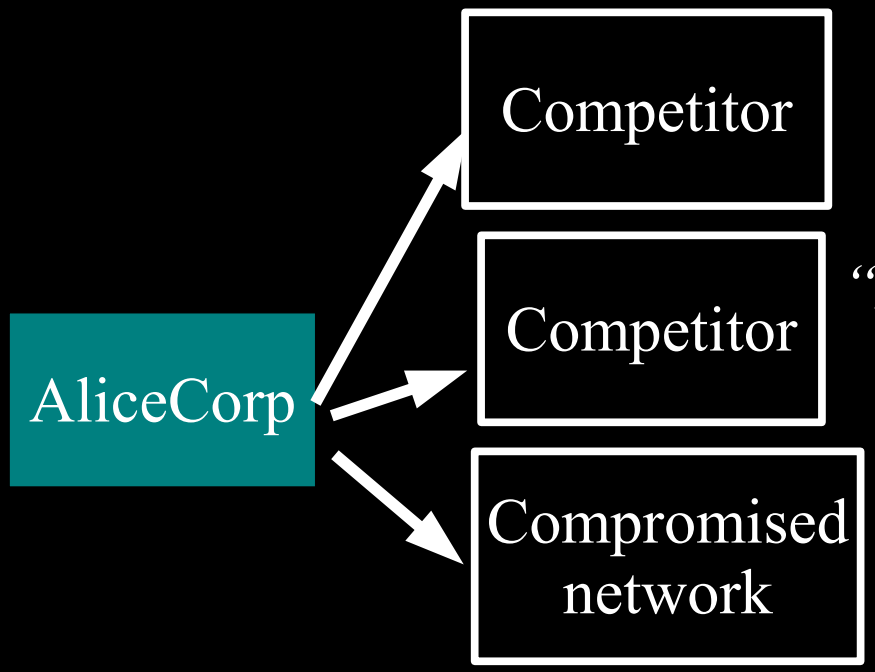
Four pieces to today's talk

- 1) Who uses Tor and why?
- 2) The Tor design in a nutshell
- 3) Tor and censorship
- 4) What NSF is funding us for

Regular citizens don't want to be watched and tracked.



Businesses need to keep trade secrets.

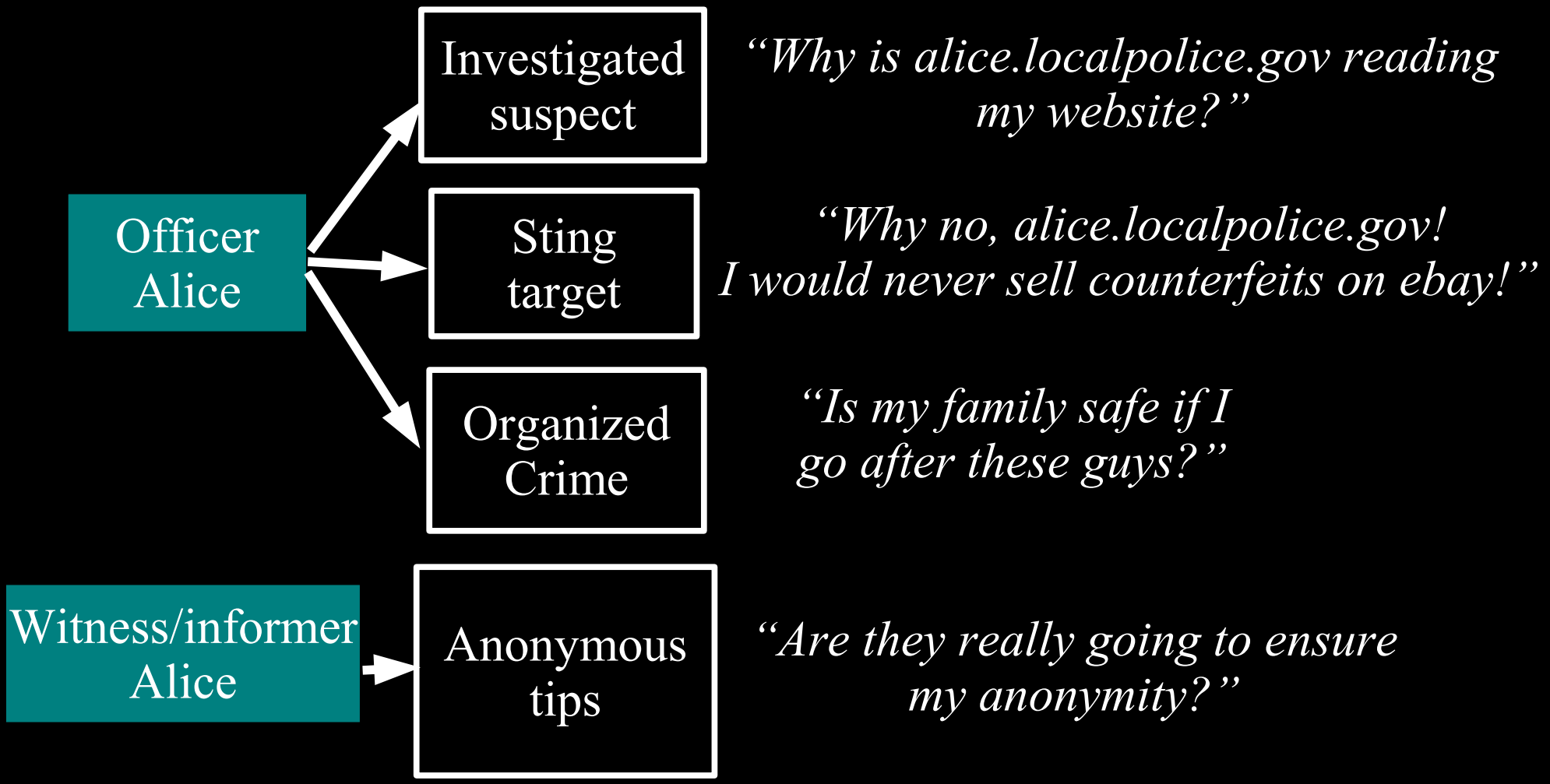


"Oh, your employees are reading our patents/jobs page/product sheets?"

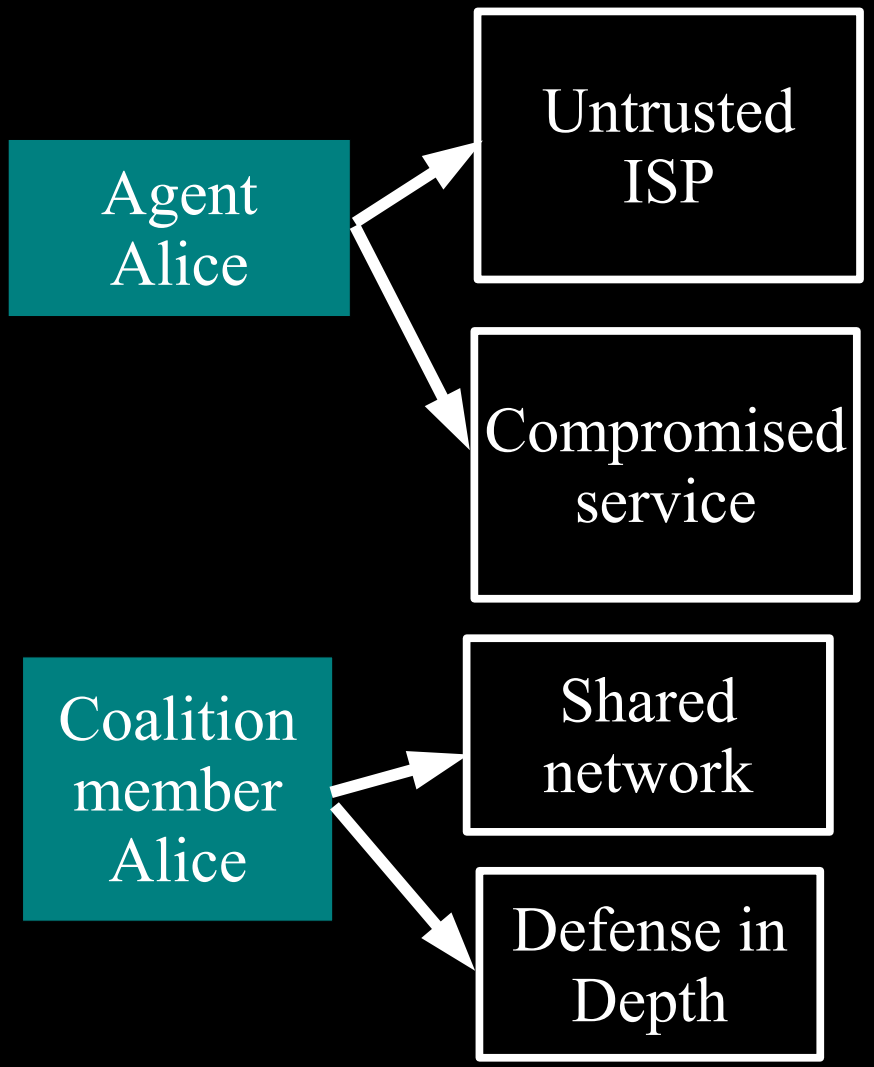
"Hey, it's Alice! Give her the 'Alice' version!"

"Wanna buy a list of Alice's suppliers? What about her customers? What about her engineering department's favorite search terms?"

Law enforcement needs anonymity to get the job done.



Governments need anonymity for their security



"What will you bid for a list of Baghdad IP addresses that get email from .gov?"

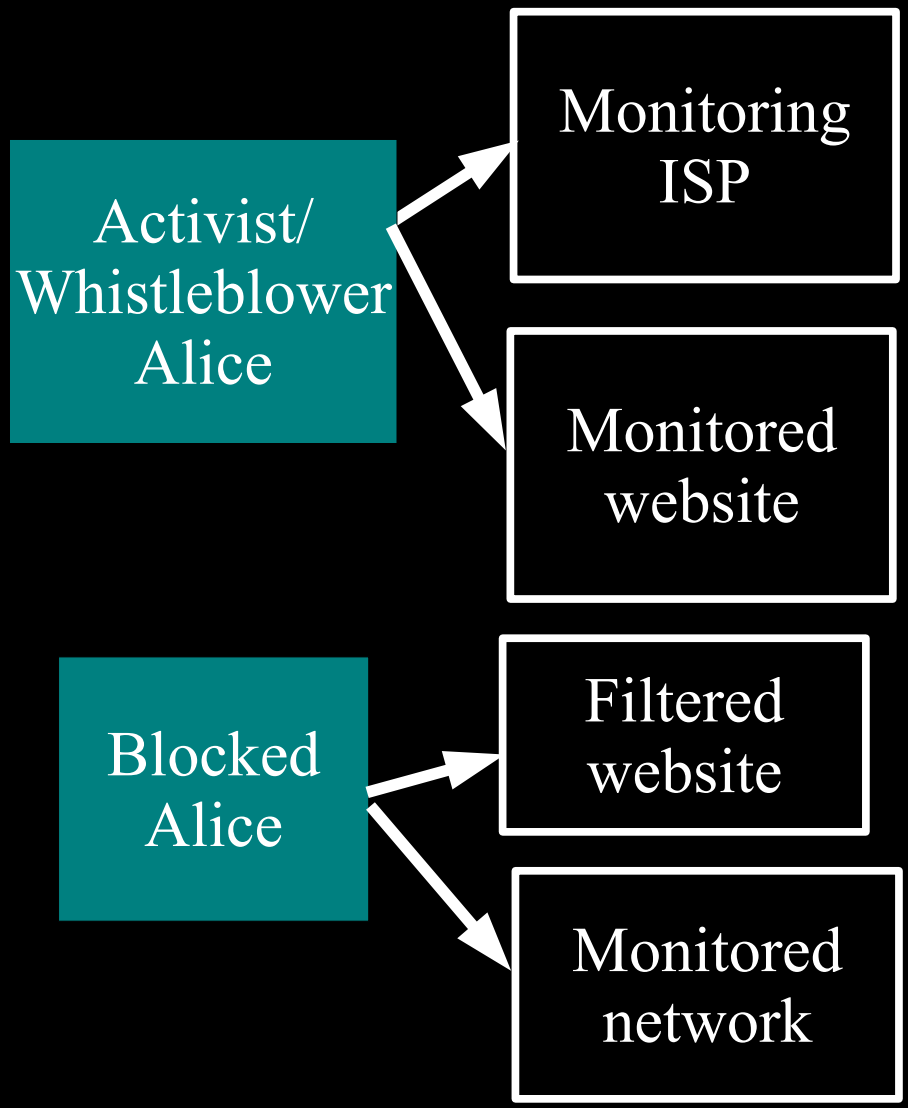
"Somebody in that hotel room just checked his Navy.mil mail!"

"*What* does *FBI Google for?*"

"Do I really want to reveal my internal network topology?"

"What about insiders?"

Journalists and activists need Tor for their personal safety



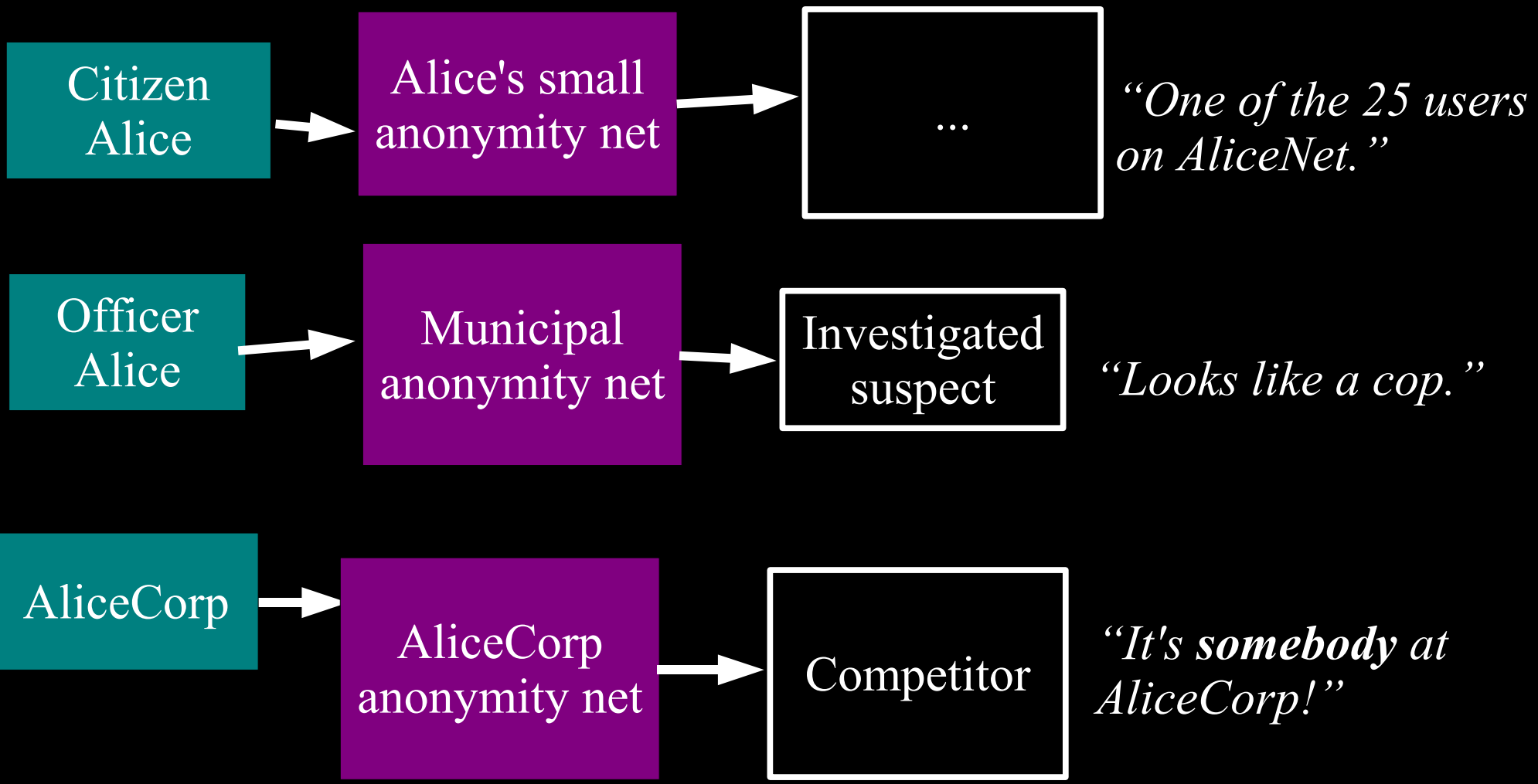
"Did you just post to that website?"

"Where are the bloggers connecting from?"
"I run livejournal and track my users"
"Of course I tell China about my users"

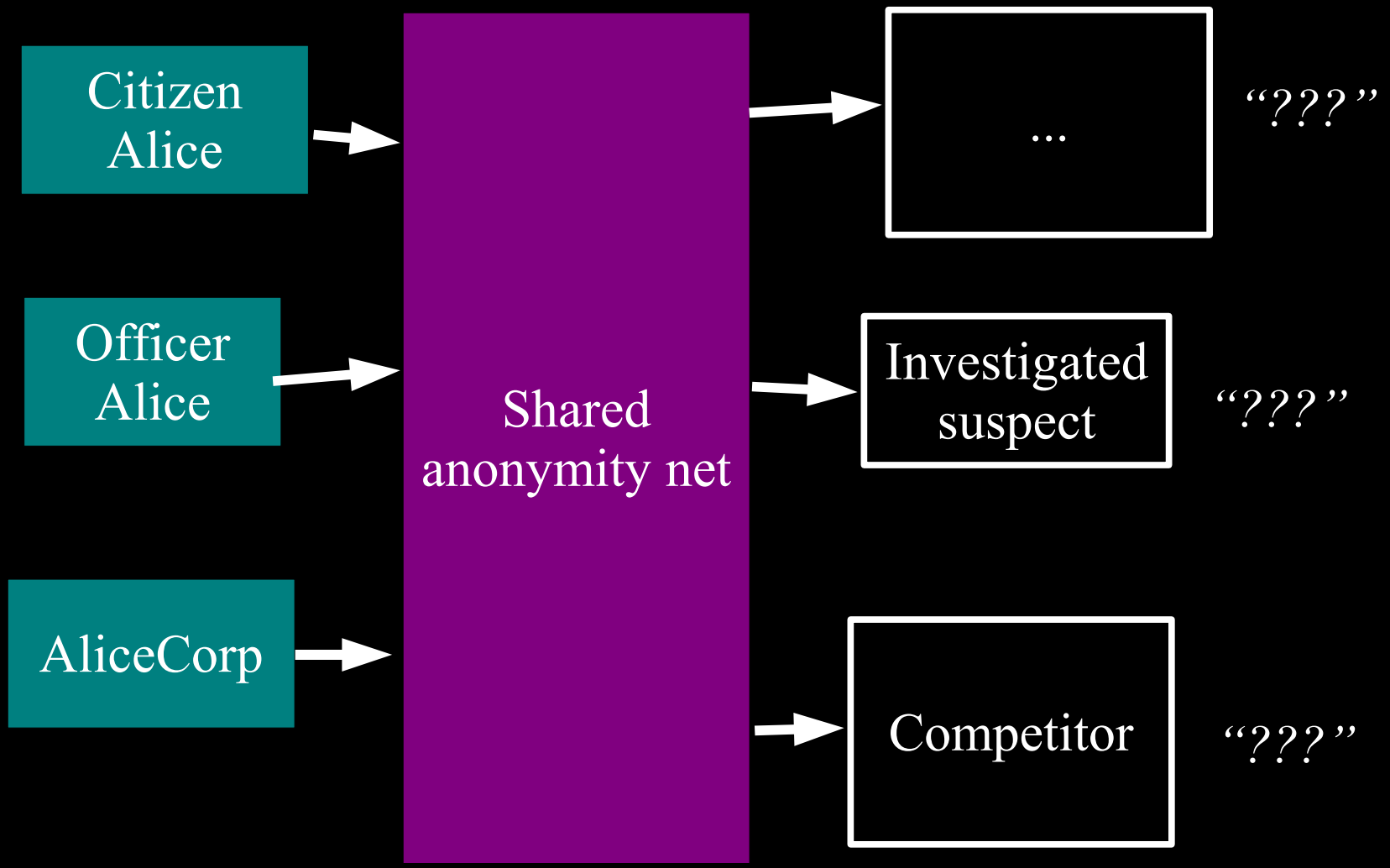
"What does the Global Voices website say today?"
"I want to tell people what's going on in my country"

"I think they're watching. I'm not even going to try."

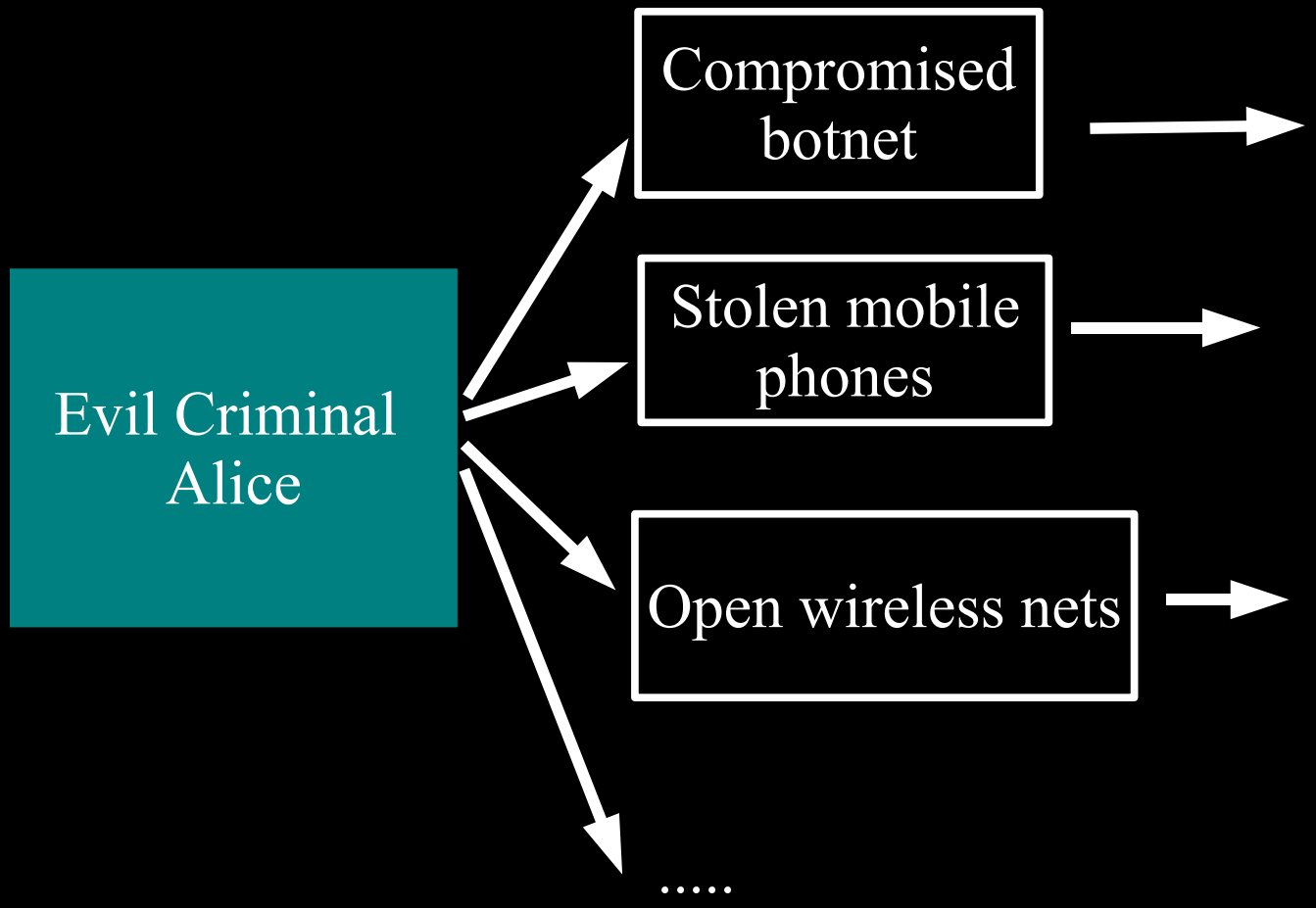
You can't get anonymity on your own: private solutions are ineffective...



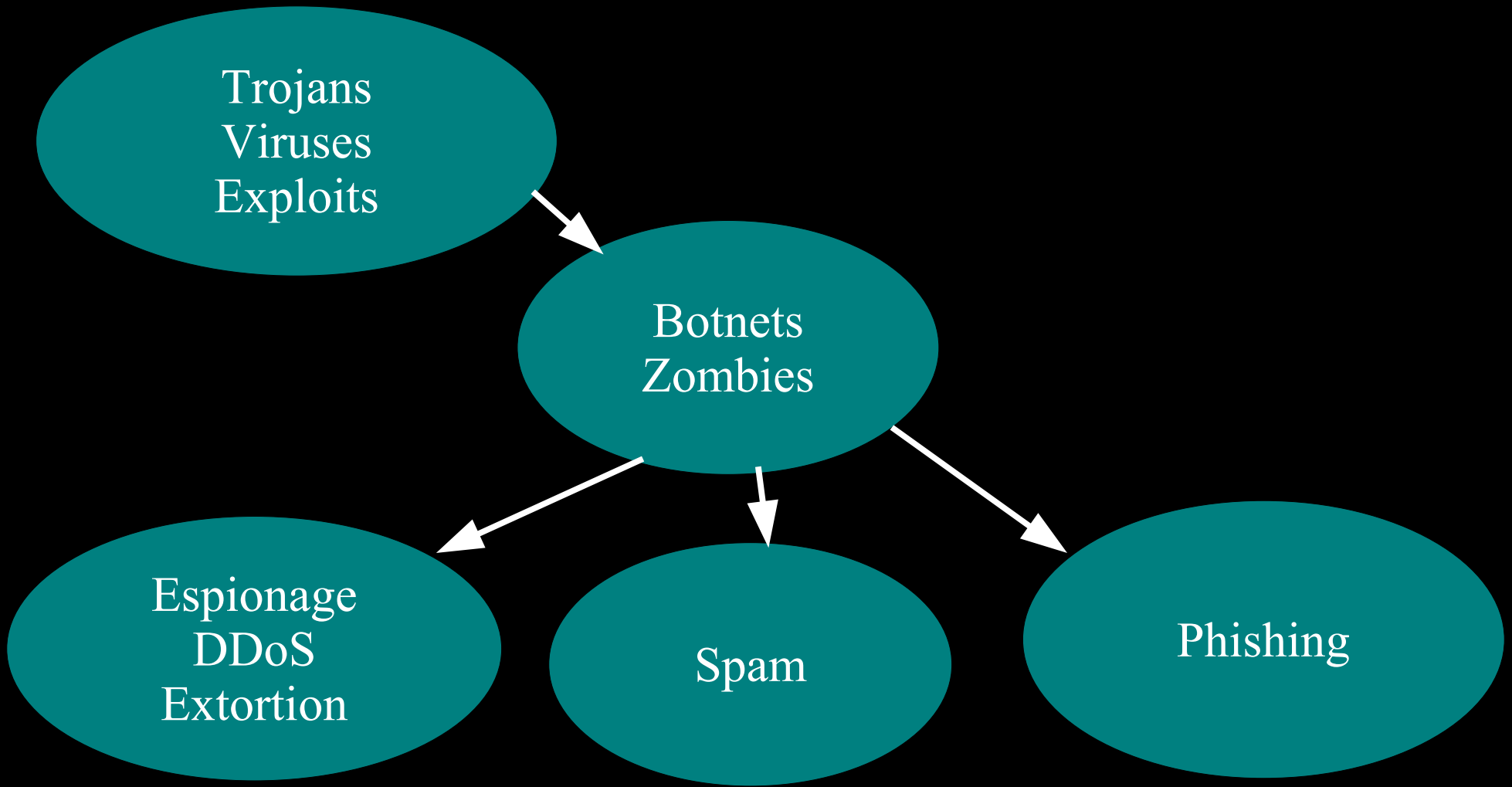
... so, anonymity loves company!



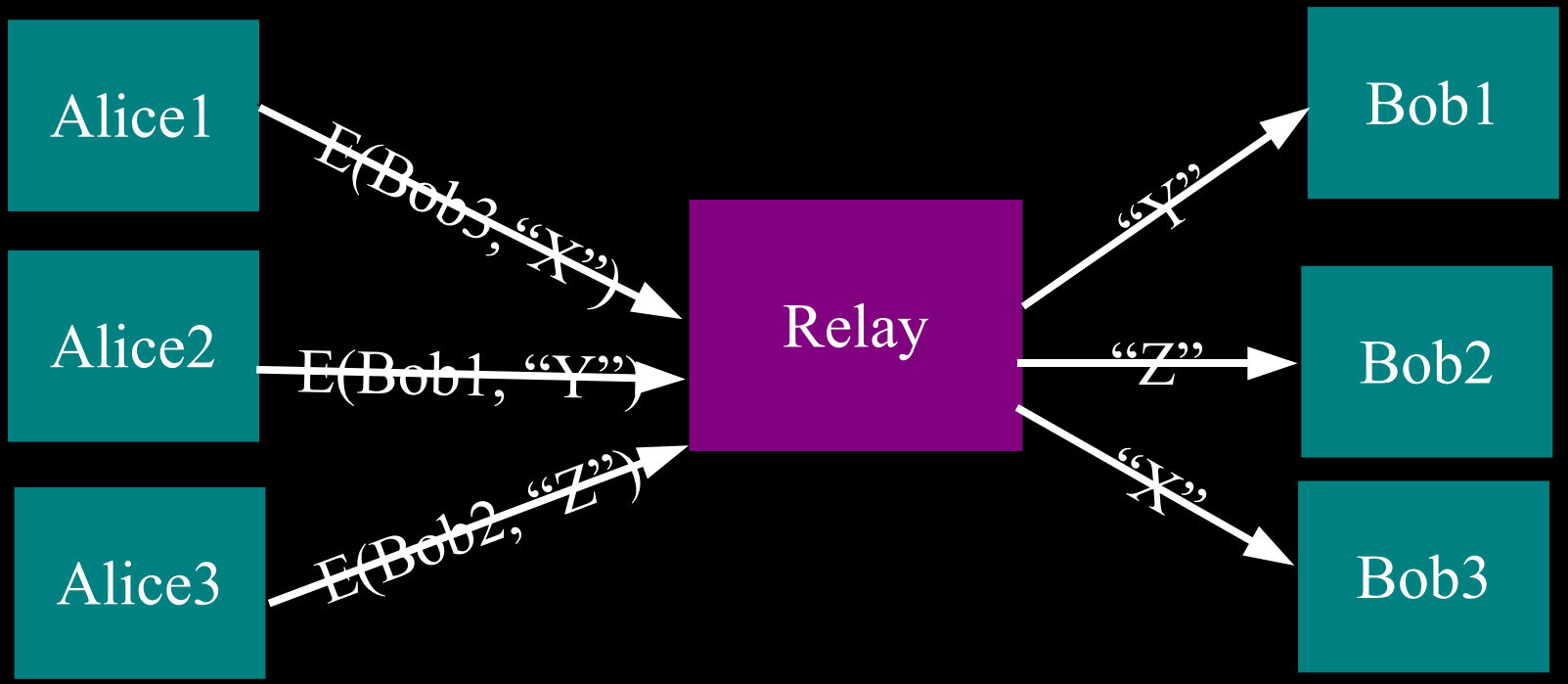
Yes, bad people need anonymity too. But they are *already* doing well.



Current situation: Bad people on the Internet are doing fine

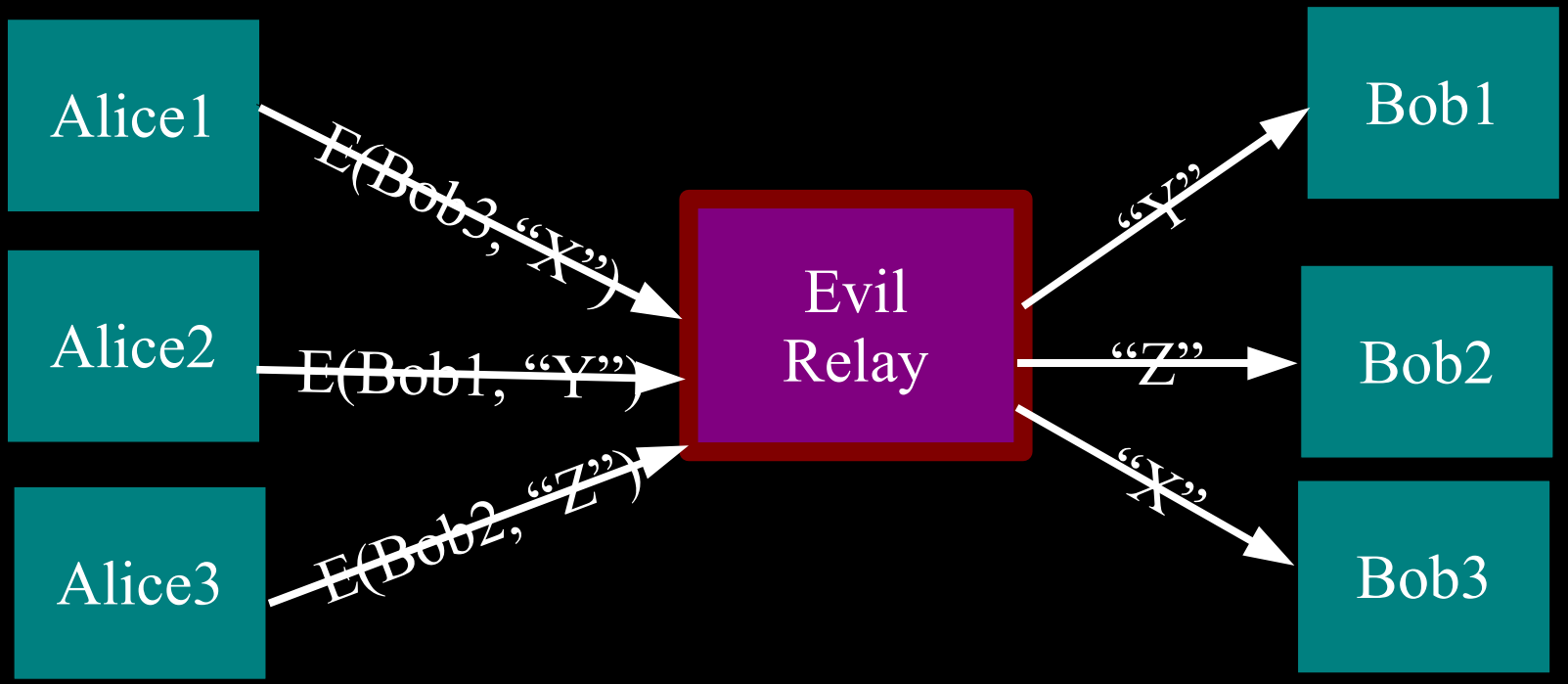


The simplest designs use a single relay to hide connections.

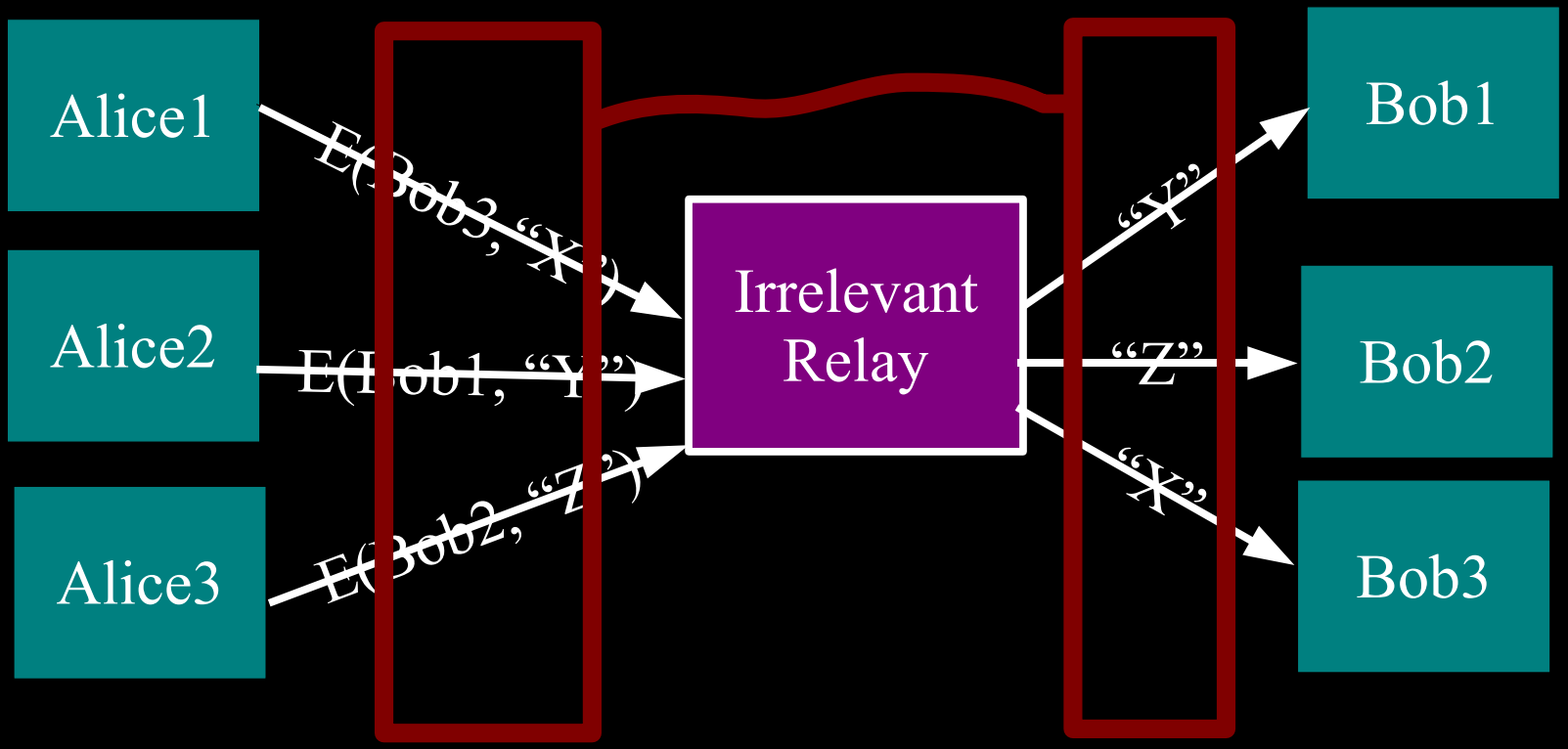


(example: some commercial proxy providers)

But a single relay (or eavesdropper!) is a single point of failure.

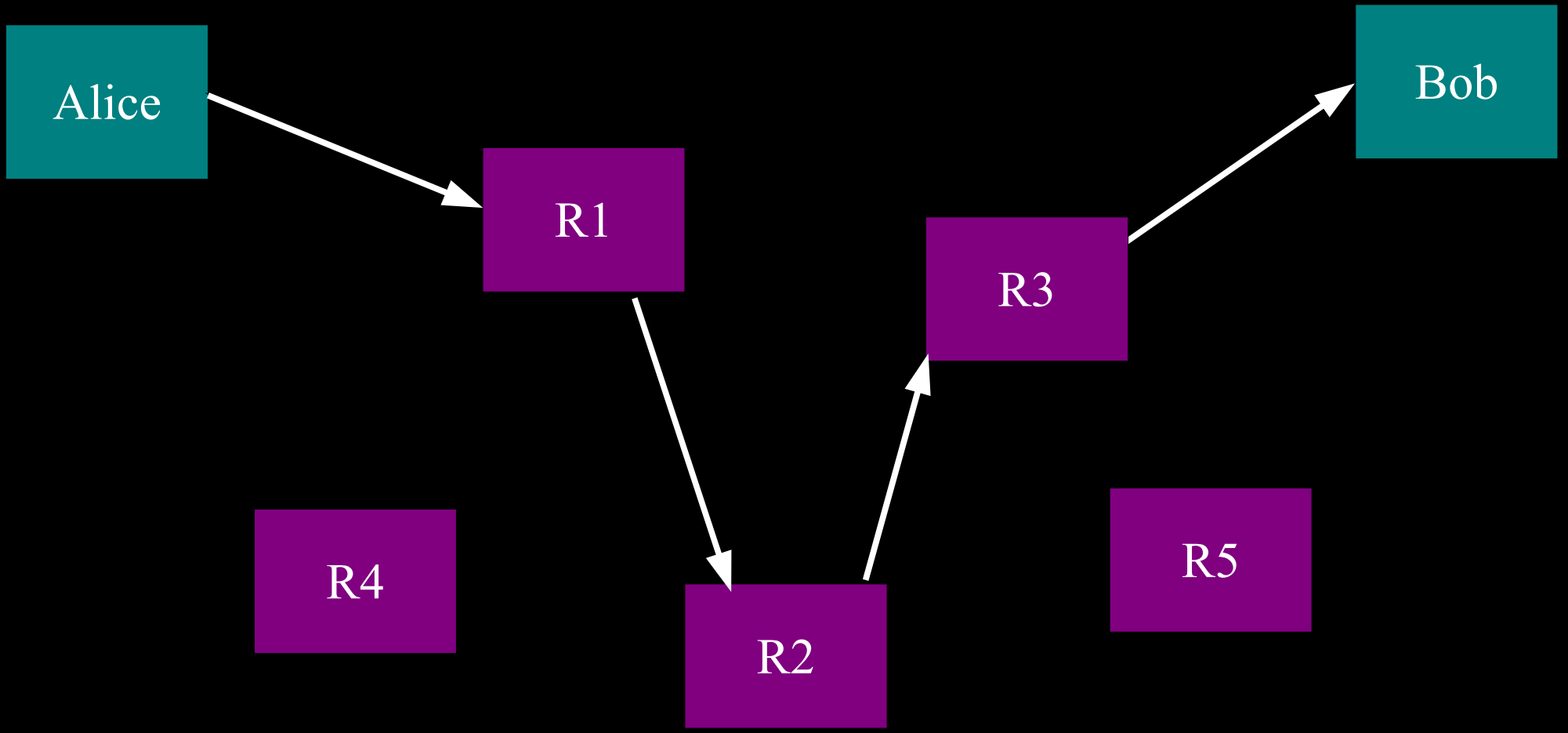


... or a single point of bypass.

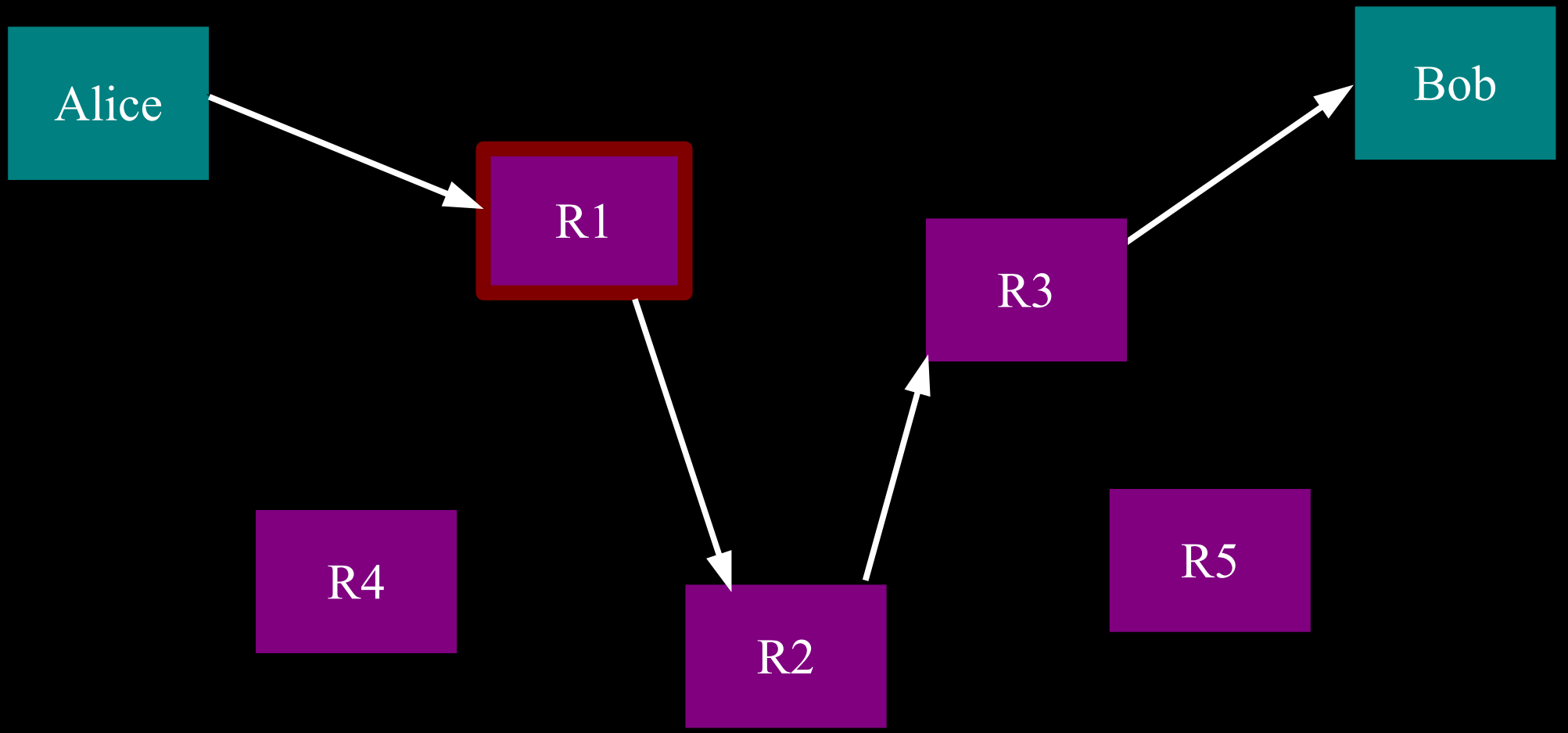


Timing analysis bridges all connections through relay ⇒ An attractive fat target

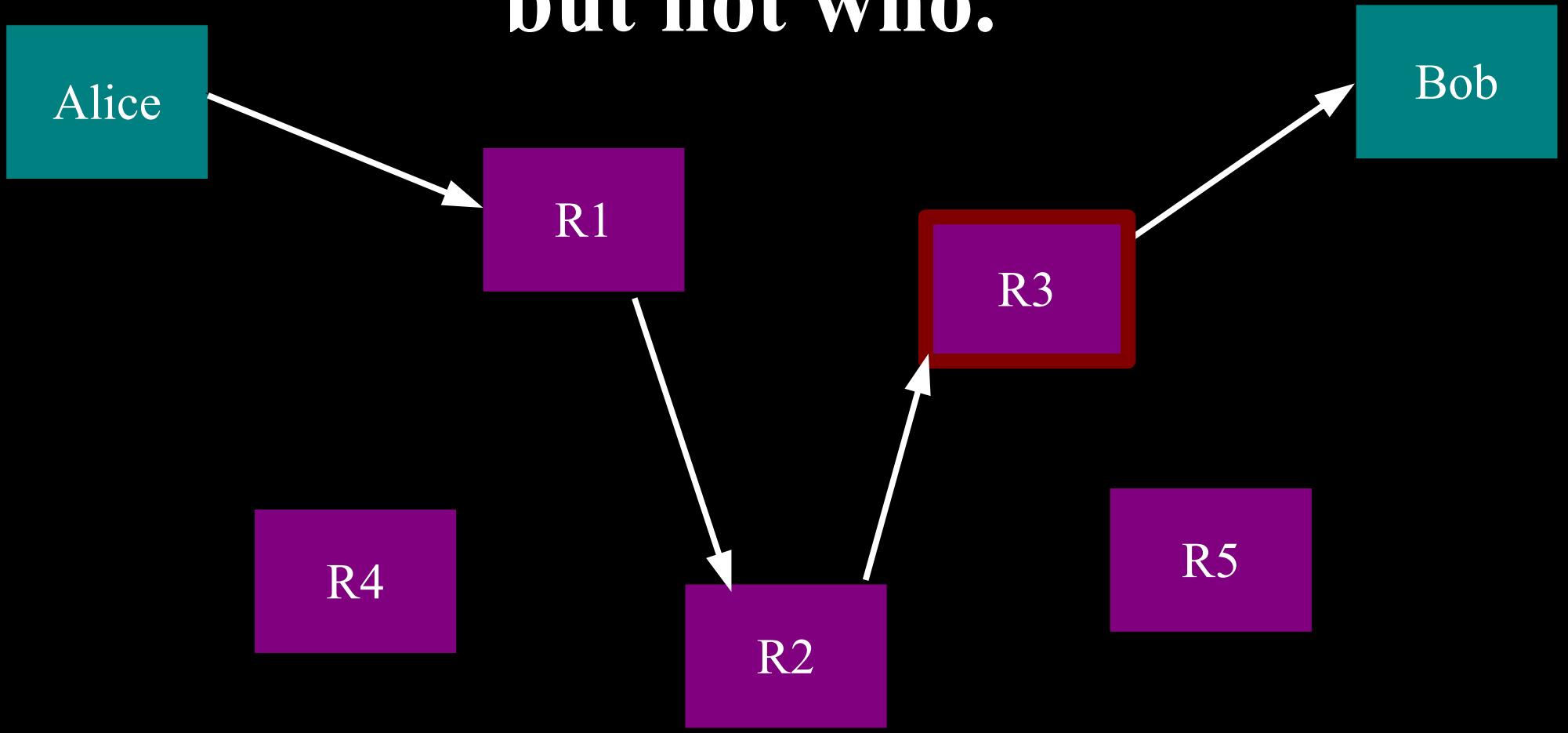
So, add multiple relays so that no single one can betray Alice.



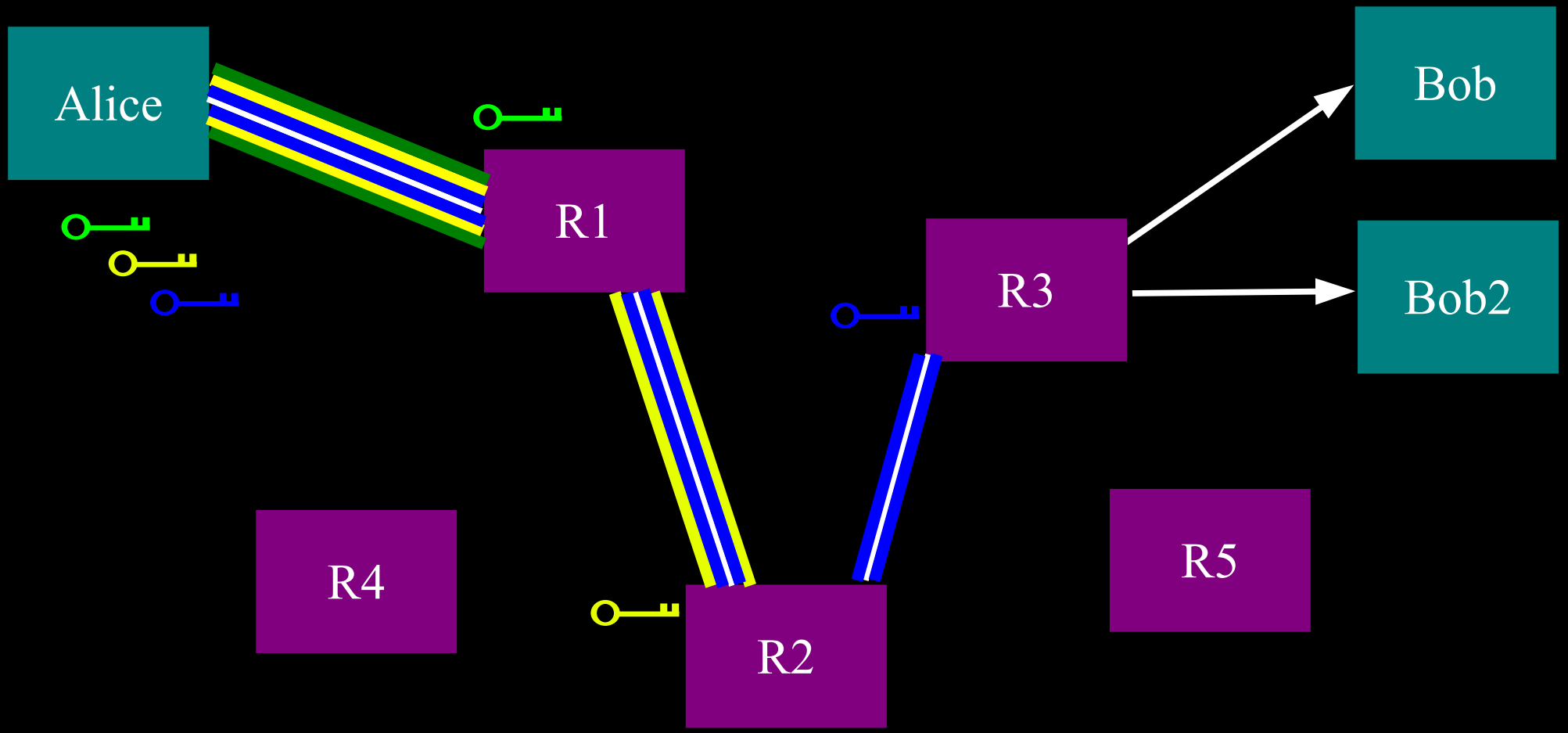
A corrupt first hop can tell that Alice is talking, but not to whom.



A corrupt final hop can tell that somebody is talking to Bob, but not who.



Alice makes a session key with R1 ...And then tunnels to R2...and to R3



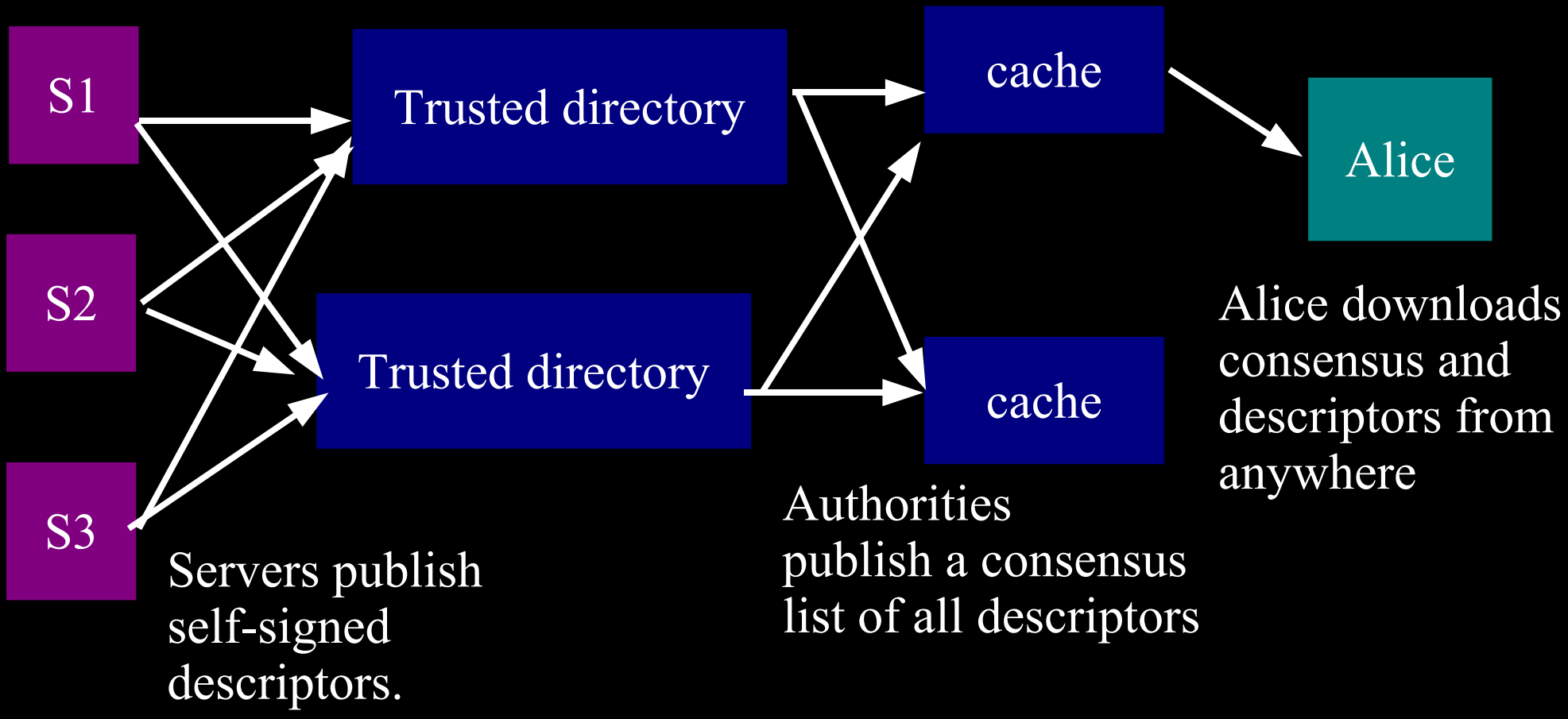
What we spend our time on

- Performance and scalability
- Maintaining the whole software ecosystem
- Blocking-resistance (circumvention)
- Basic research on anonymity
- Reusability and modularity
- Advocacy, education, and trainings around the world
- Metrics, data, and analysis

Relay versus Discovery

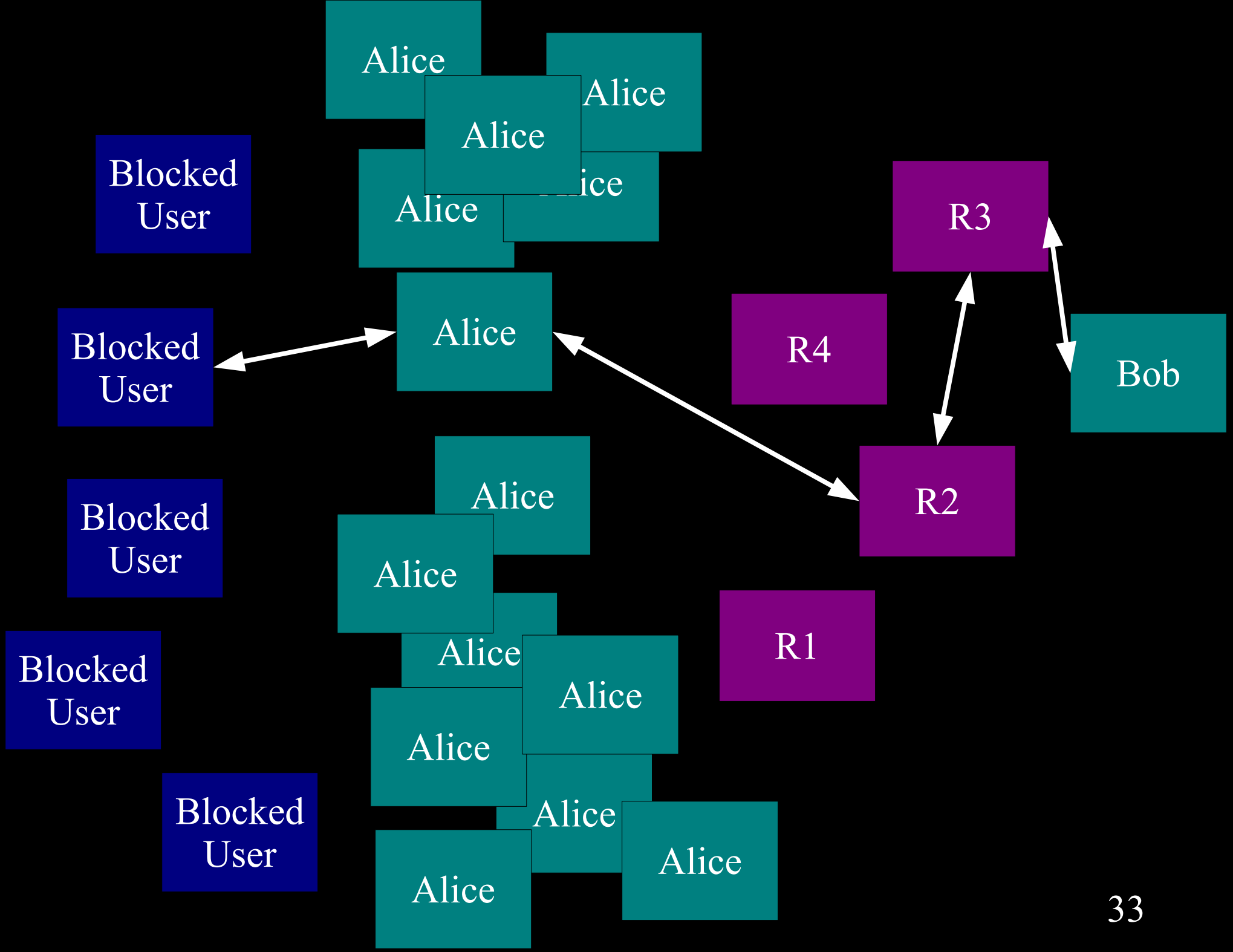
- There are two pieces to all these “proxying” schemes:
- a **relay** component: building circuits, sending traffic over them, getting the crypto right
- a **discovery** component: learning what relays are available

The basic Tor design uses a simple centralized directory protocol.



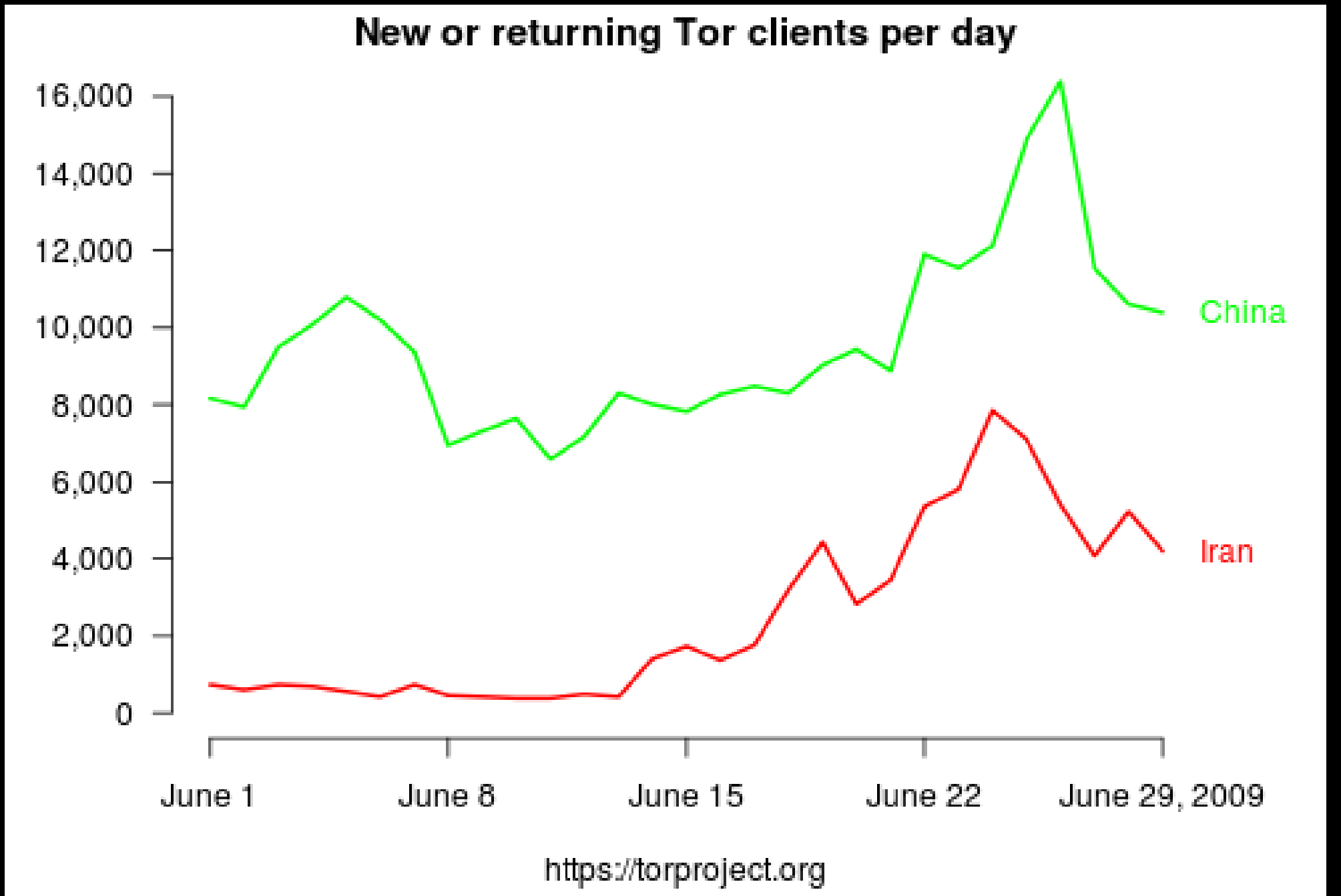
Attackers can block users from connecting to the Tor network

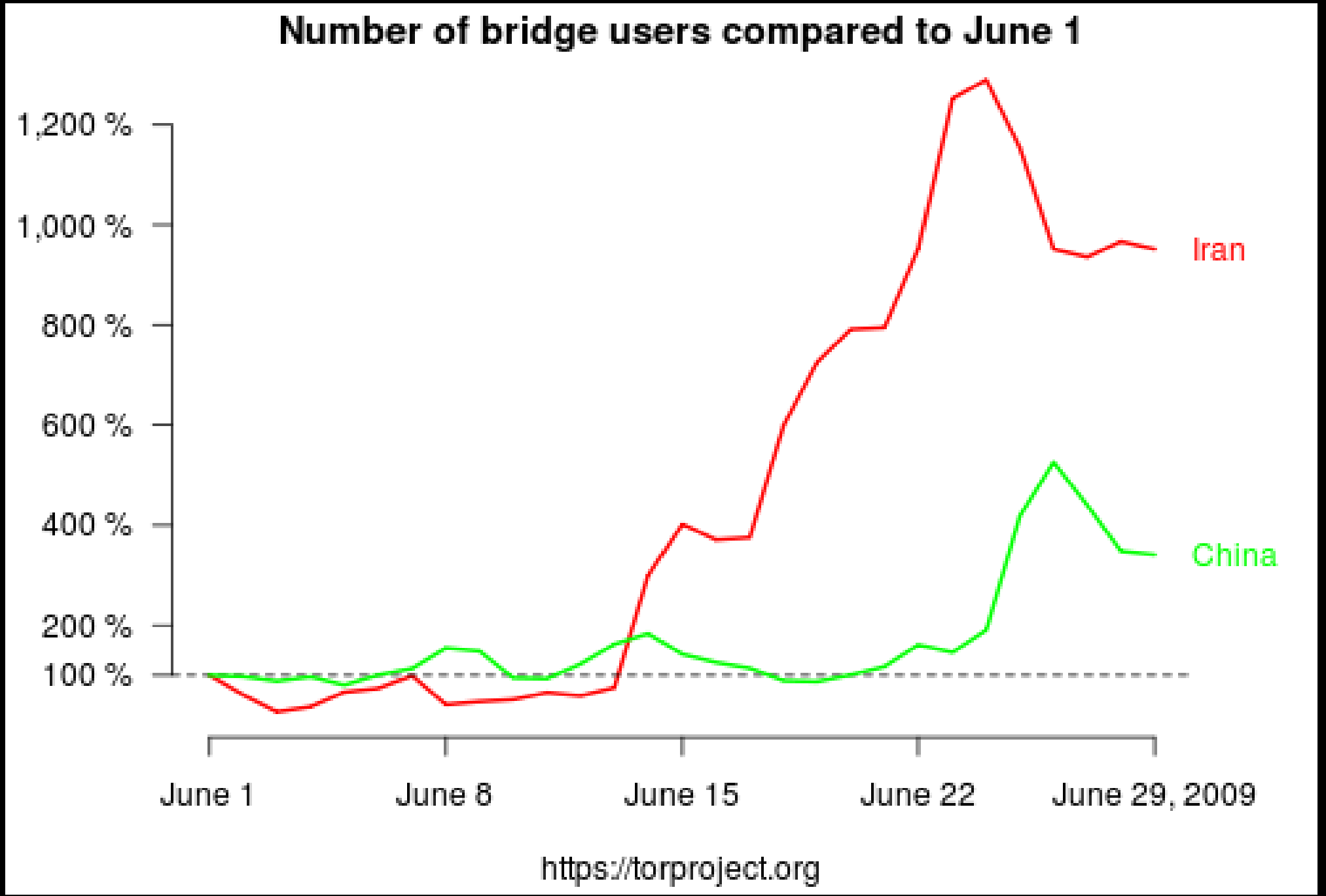
- By blocking the directory authorities
- By blocking all the relay IP addresses in the directory
- By filtering based on Tor's network fingerprint
- By preventing users from finding the Tor software

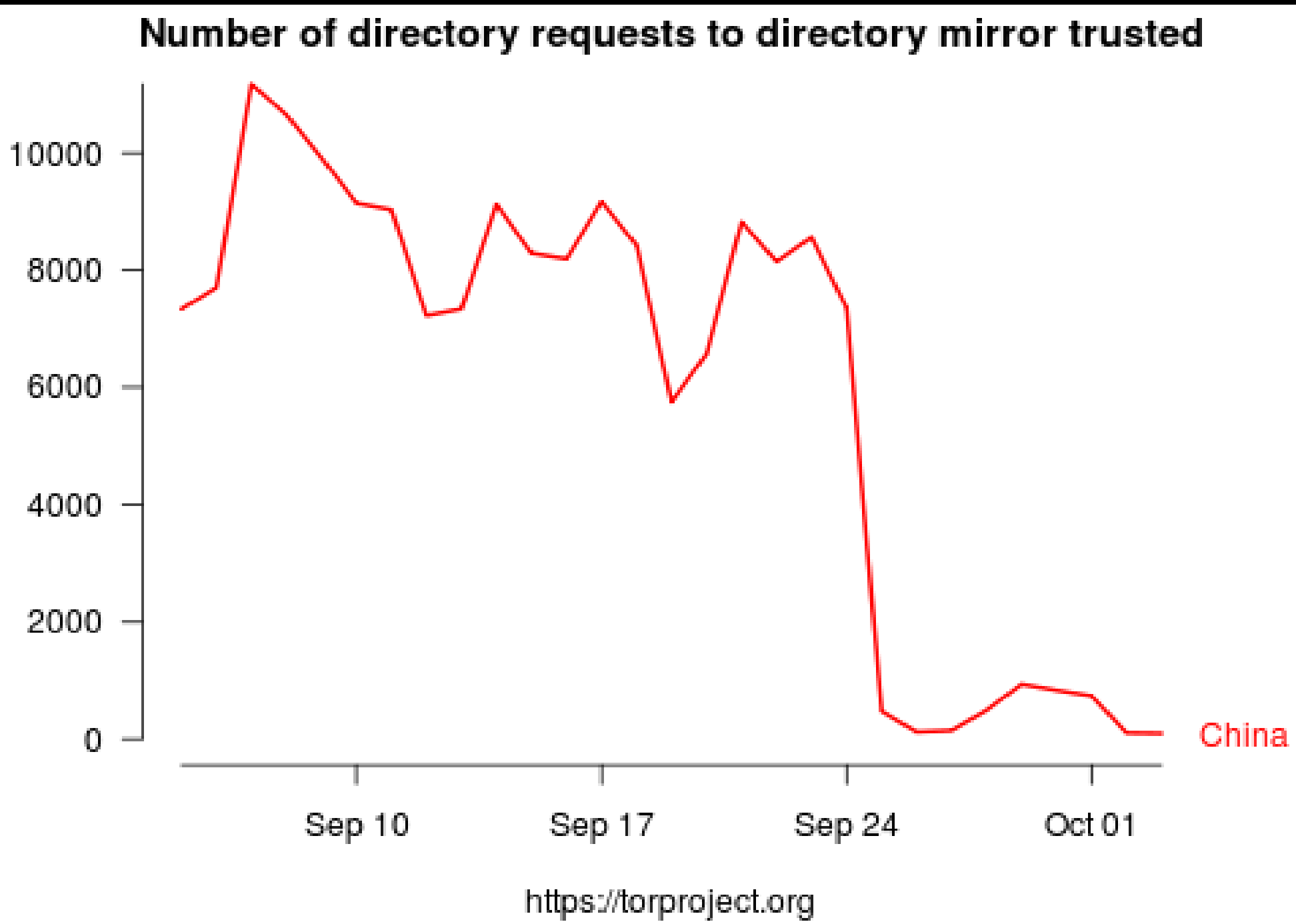

Alice
Alice
Alice
Alice
Alice
Blocked User
Blocked User
Blocked User
Blocked User
Blocked User
Alice
Alice
Alice
Alice
Alice
Alice
Alice
R3
R4
R2
R1
Bob

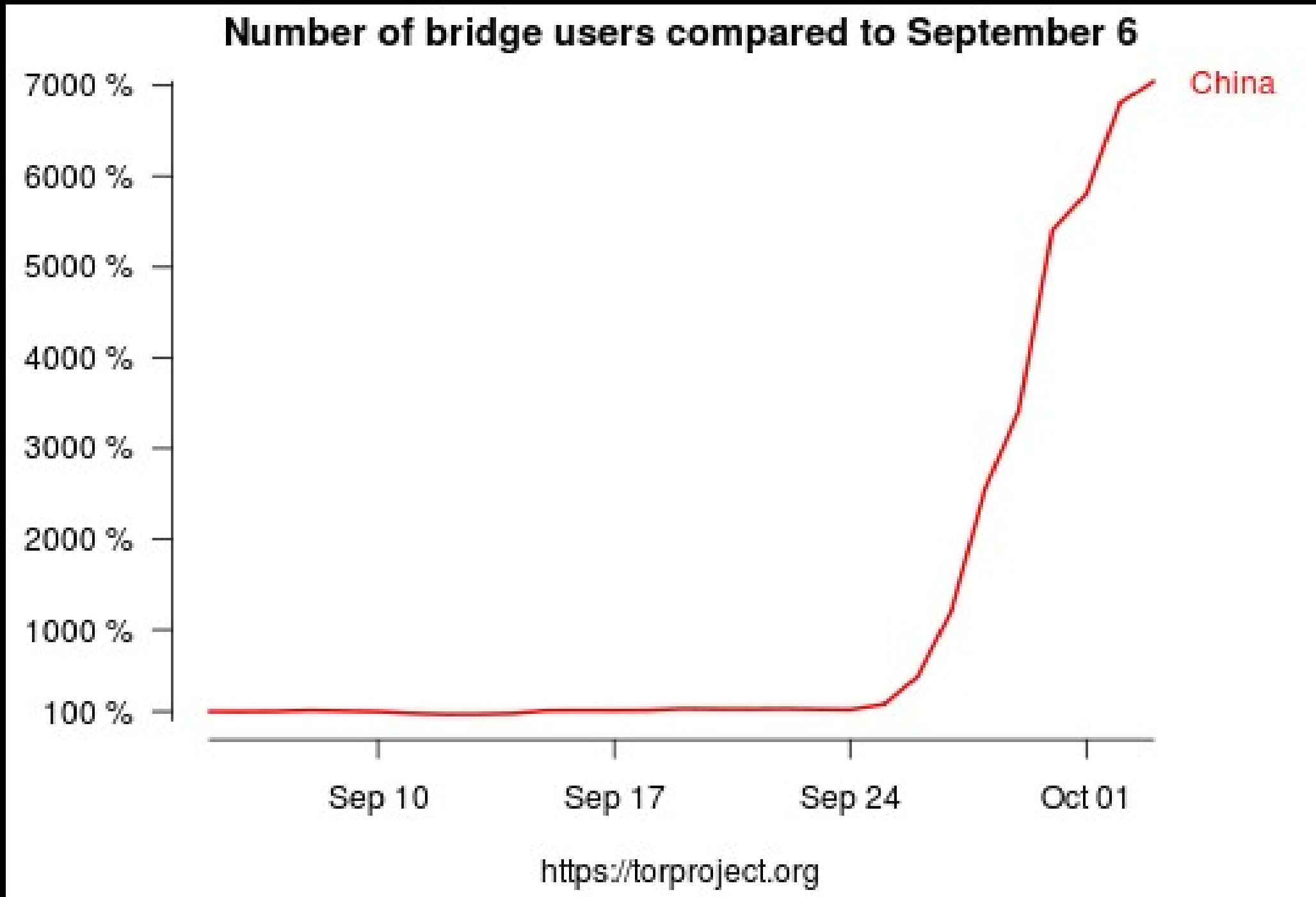
“Bridge” relays

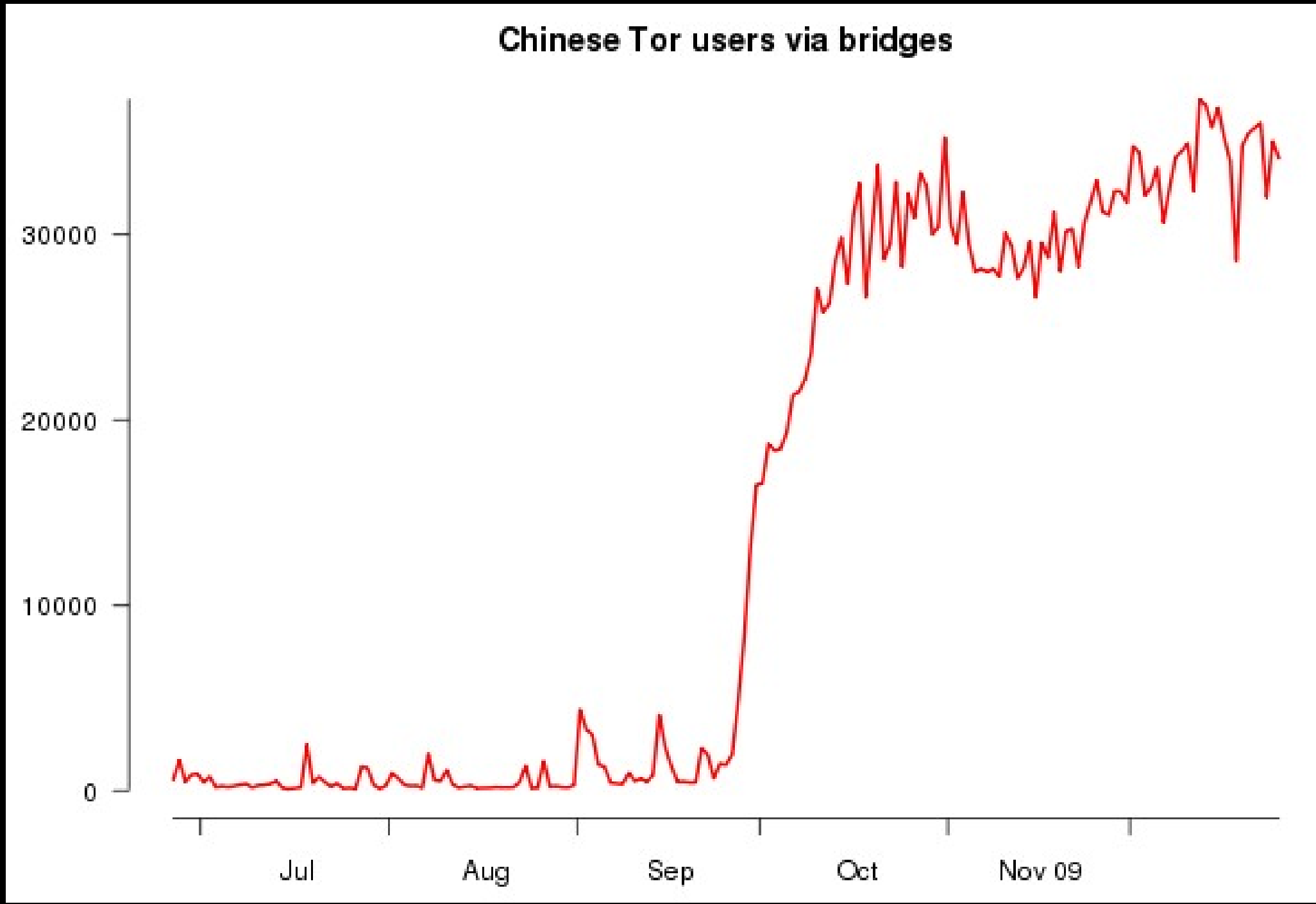
- Hundreds of thousands of Tor users, already self-selected for caring about privacy.
- Rather than signing up as a normal relay, you can sign up as a special “bridge” relay that isn't listed in any directory.
- No need to be an “exit” (so no abuse worries), and you can rate limit if needed
- Integrated into Vidalia (our GUI) so it's easy to offer a bridge or to use a bridge


New or returning Tor clients per day
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
China
Iran
June 1
June 8
June 15
June 22
June 29, 2009
https://torproject.org


Number of bridge users compared to June 1
1,200 %
1,000 %
800 %
600 %
400 %
200 %
100 %
Iran
China
June 1
June 8
June 15
June 22
June 29, 2009
https://torproject.org


Number of directory requests to directory mirror trusted
10000
8000
6000
4000
2000
0
Sep 10
Sep 17
Sep 24
Oct 01
China
https://torproject.org


Number of bridge users compared to September 6
7000 %
6000 %
5000 %
4000 %
3000 %
2000 %
1000 %
100 %
China
Sep 10
Sep 17
Sep 24
Oct 01
https://torproject.org


Chinese Tor users via bridges
30000
20000
10000
0
Jul
Aug
Sep
Oct
Nov 09

How do you find a bridge?

- If you can, go to **https://bridges.torproject.org/** and it will tell you a few based on time and your IP address
- Mail bridges@torproject.org from a gmail/yahoo address, and we'll send you a few
- From your friends and neighbors, like before

Bridge directory authorities

- Specialized dir authorities that aggregate and track bridges, but don't provide a public list:
  - You can keep up-to-date about a bridge once you know its key, but can't just grab list of all bridges.
- Identity key and address for default bridge authorities ship with Tor.
- Bridges publish via Tor, in case somebody is monitoring the authority's network.

One working bridge is enough

- Connect via that bridge to the bridge authority.
- ...and to the main Tor network.
- Remember, all of this happens in the background.
- “How to circumvent for all transactions (and trust the pages you get)”
  is now reduced to
  “How to learn about a working bridge”.

Hiding Tor's network fingerprint

- We got rid of plaintext HTTP (used by directories). Now clients tunnel their directory requests over the same TLS connection as their other Tor traffic.
- We've made Tor's TLS handshake look more like Firefox+Apache.
- When Iran kicked out Smartfilter in early 2009, Tor's old v2 dir design worked again!

Attacker's goals (1)

- Restrict the flow of certain kinds of information
  - Embarrassing (rights violations, corruption)
  - Opposing (opposition movements, sites that organize protests)
- Chill behavior by *impression* that online activities are monitored

Attacker's goals (2)

- Little reprisal against passive consumers of information.
  - Producers and distributors of information in greater danger.
- Censors (actually, govts) have economic, political, social incentives not to block the whole Internet.
  - But they don't mind collateral damage.

Main network attacks

- Block by IP address / port at firewall
- Intercept DNS requests and give bogus responses or redirects
- China: Keywords in TCP packets
- Iran: DPI to filter SSL when they want
- Russia: Don't block, just pollute

What we're up against (1)

- Govt firewalls used to be stateless. Now they're buying fancier hardware.
  - Burma vs Iran vs China
- New filtering techniques spread by commercial (American) companies :(
- How to separate “oppressing employees” vs “oppressing citizens” arms race?

What we're up against (2)

- Censorship is not uniform even within each country, often due to different ISP policies
- Attacker can influence other countries and companies to help them censor or track users. We'll see if the GNI (Global Network Initiative) changes that.

Blocking goes both ways

- If China blackholes your IP address, you can't reach Chinese websites either.
- So if exit relays are blackholed, Tor users can't read Chinese websites. :(
- And if you use dynamic IP addresses, then more and more of your neighbors can't read Chinese websites?

Javascript, cookies, history, etc

- Javascript refresh attack
- Cookies, History, browser window size, user-agent, language, http auth, ...
- Mostly problems when you toggle from Tor to non-Tor or back
- Mike Perry's Torbutton Firefox extension tackles many of these

Flash is dangerous too

- Some apps are bad at obeying their proxy settings.
- Adobe PDF plugin. Flash. Other plugins. Extensions. Especially Windows stuff: did you know that Microsoft Word is a network app?

Choose how to install it

- Tor Browser Bundle: standalone Windows exe with Tor, Vidalia, Firefox, Torbutton, Polipo, e.g. for USB stick
- Vidalia bundle: Windows/OSX installer
- Tor VM: Transparent proxy for Windows
- “Net installer” via our secure updater
- Amnesia Linux LiveCD

Only a piece of the puzzle (1)

- Assume the users aren't attacked by their hardware and software
  - No spyware installed, no cameras watching their screens, etc
- Users need to know about SSL for gmail. Cookies. End-to-end encryption.
- Many people in Iran in June were using plaintext proxies!

Only a piece of the puzzle (2)

- Users can fetch a genuine copy of Tor?
- PGP signatures are great, but nobody knows what that means, and nobody in Burma has my key.
- Gettor email autoresponder. USB key spread by hand.
- Our secure updater should help.

Sustainability

- Tor has a community of developers and volunteers.
- Commercial anonymity systems have flopped or constantly need more funding for bandwidth.
- Our sustainability is rooted in Tor's open design: clear documentation, modularity, and open source.

Responding to China blocks

- In late Sept, conflicting advice from experts:
- “Hit 'em in the nose, show that you care about your users”
- “Lie low and let it pass. You're about more than China.”
- Tor is a new approach to China bloggers: “Find new bridge” rather than “get software update”.

Publicity attracts attention

- Many circumvention tools launch with huge media splashes. (The media loves this.)
- But publicity attracts attention of the censors.
- We threaten their *appearance* of control, so they must respond.
- We can control the pace of the arms race.

Using Tor in oppressed areas

- Common assumption: risk from using Tor increases as firewall gets more restrictive.
- But as firewall gets more restrictive, more ordinary people use Tor too, for more mainstream activities.
- So the “median” use becomes more acceptable?

Other Iran user count

- Talked to chief security officer of one of the web 2.0 social networking sites:
  - 10% of their Iranian users in June 2009 were coming through Tor
  - 90% were coming from proxies in the Amazon cloud

Trust and reputation

- See Hal Roberts' January 2009 blog post about how some circumvention tools sell user data
- Many of these tools see circumvention and privacy as totally unrelated goals


I CAN HAZ
FREEDOM?
TorProject.org

Other ongoing questions

- How to detect if bridges are blocked (and what to do once we know)
- Better strategies for giving bridges out (Twitter, better use of social networks; Kaist design project)

Bridge communities

- Volunteers run several bridges at once, or coordinate with other volunteers.
- The goal is that some bridges will be available at any given time.
- Each community has a bridge authority, to add new bridges to the pool, expire abandoned or blocked bridges, etc.
- All automated by the Tor client.

How to scale the network?

- The clients need to learn info about the relays they can use. Eventually this means partial network knowledge, and non-clique topology.
- Everybody-a-relay, and the anonymity questions that come with that.

Advocacy and education

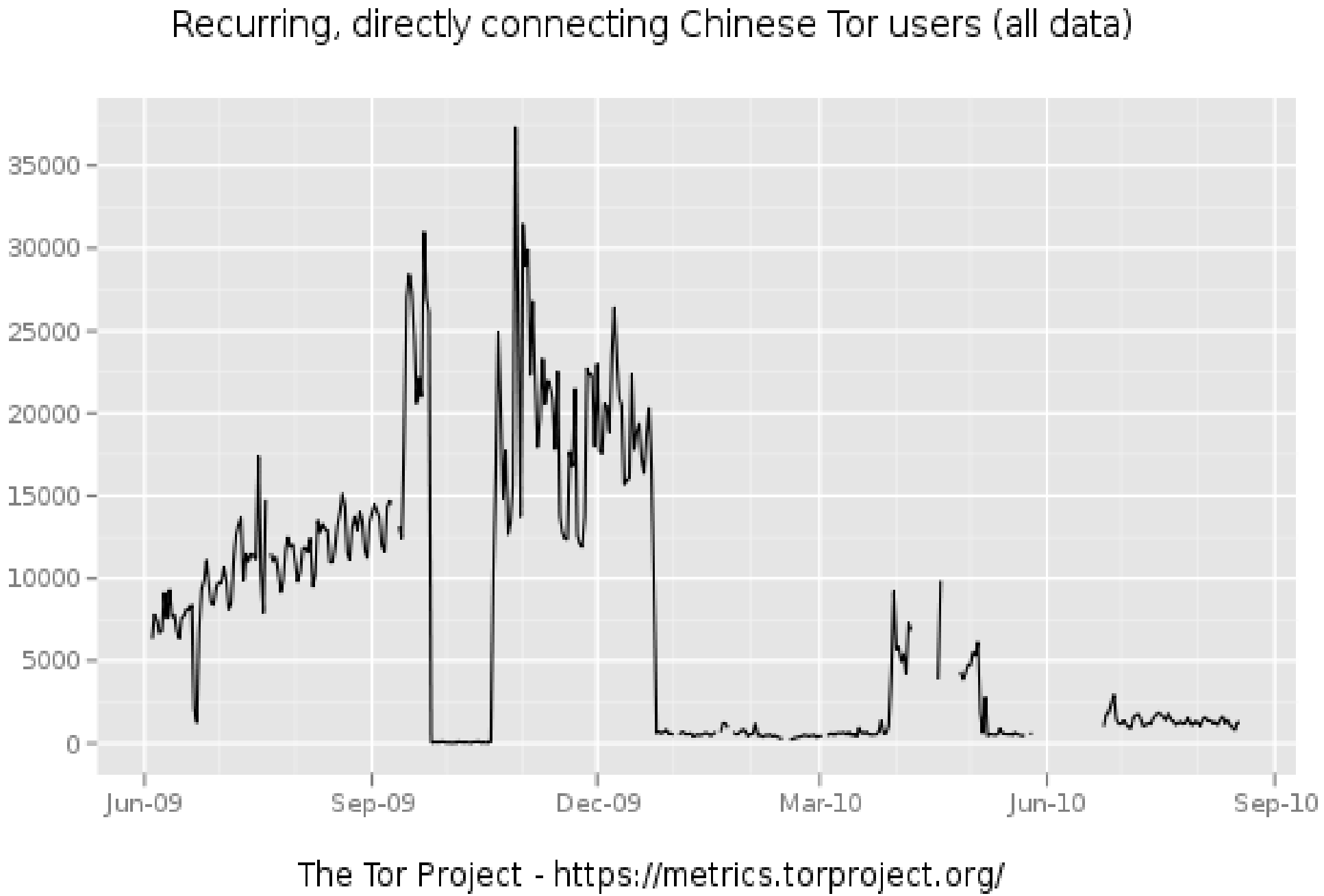
- Unending stream of people (e.g. in DC) who make critical policy decisions without much technical background
- Worse, there's a high churn rate
- Need to teach policy-makers, business leaders, law enforcement, journalists, ...
- Data retention? Internet driver's license?

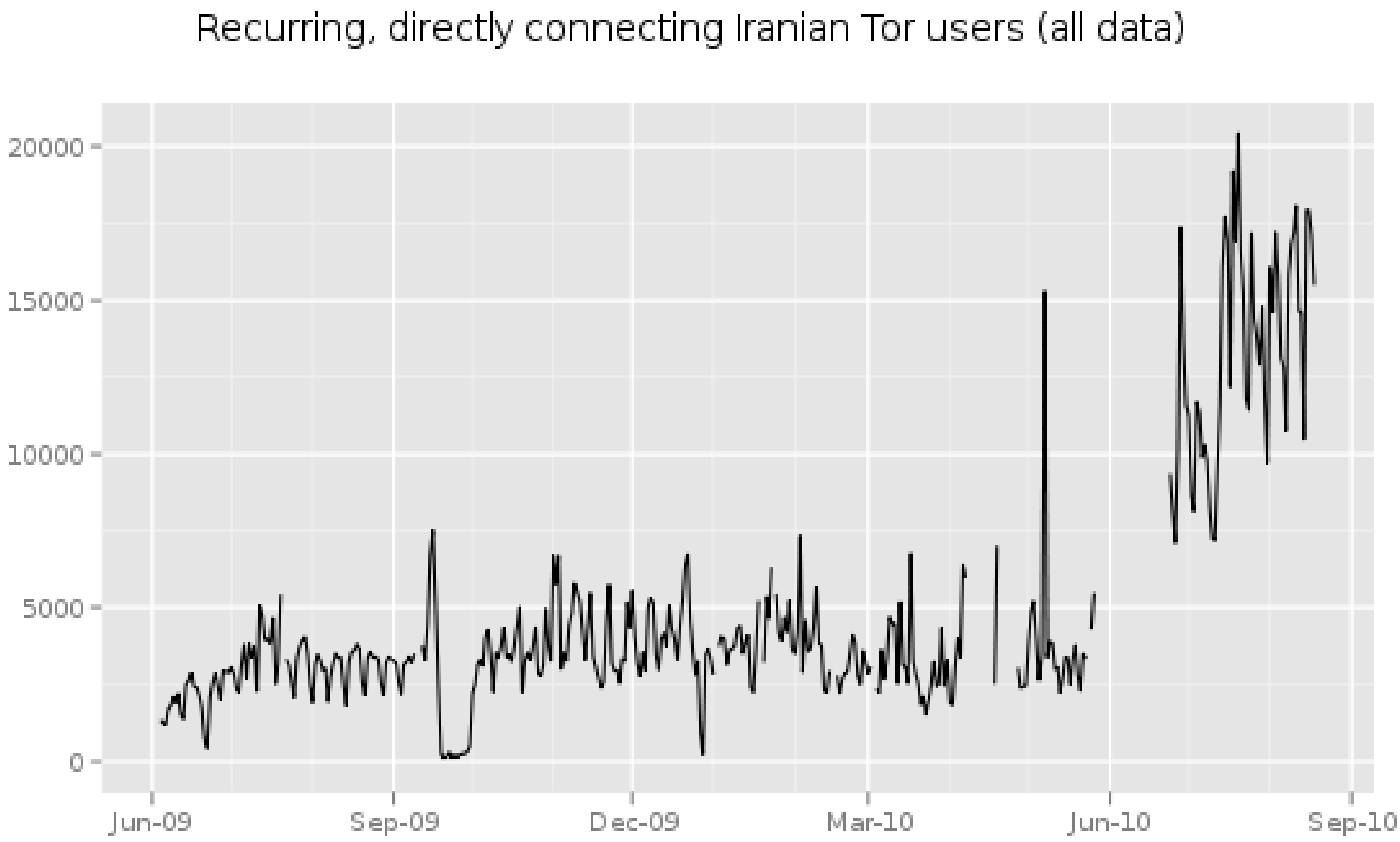
Next steps

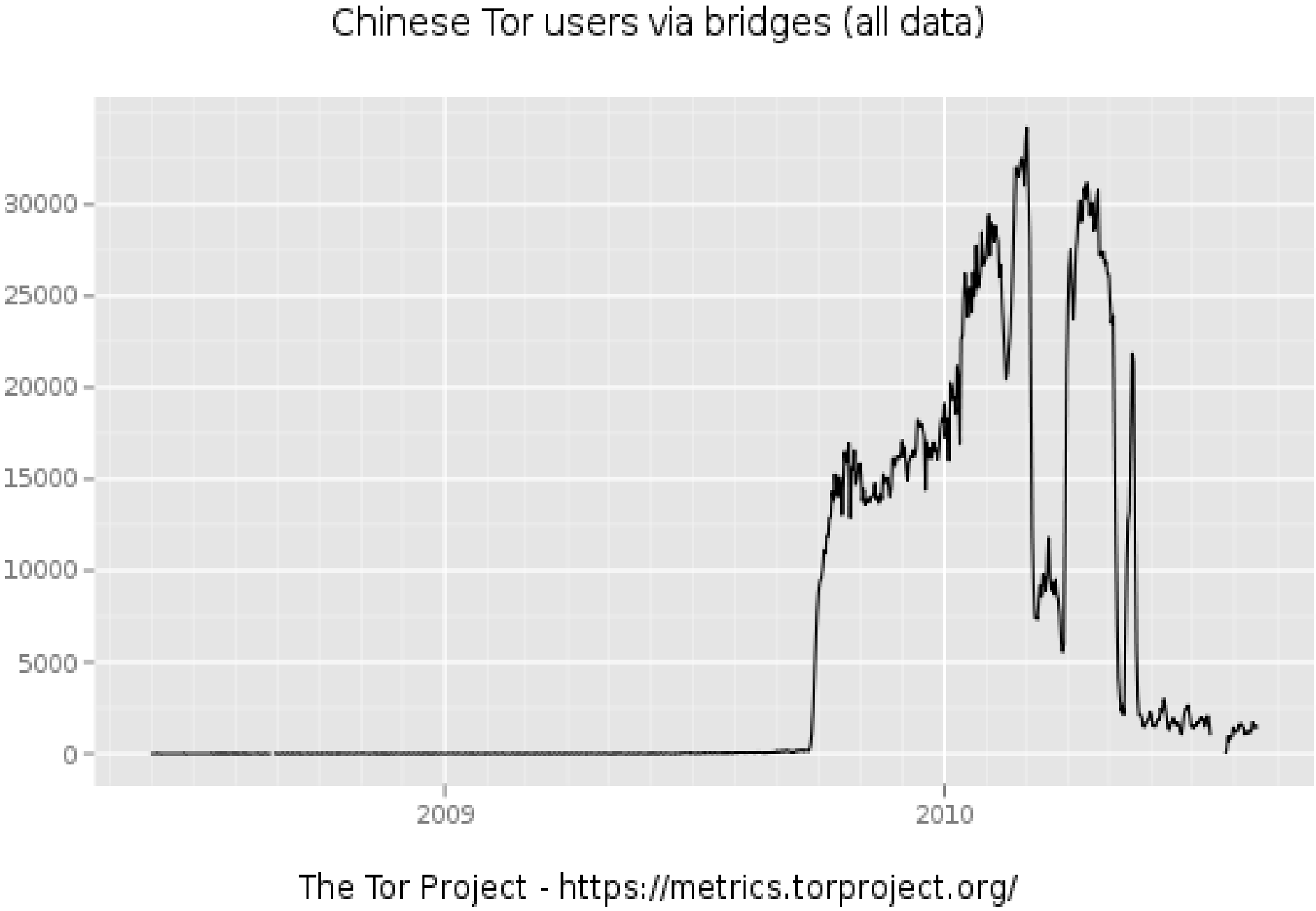
- Technical solutions won't solve the whole censorship problem. After all, firewalls are *socially* very successful in these countries.
- But a strong technical solution is still a critical puzzle piece.
- You should run a bridge! We only have 500.
- We'd love to help with some trainings, to help users and to make Tor better.

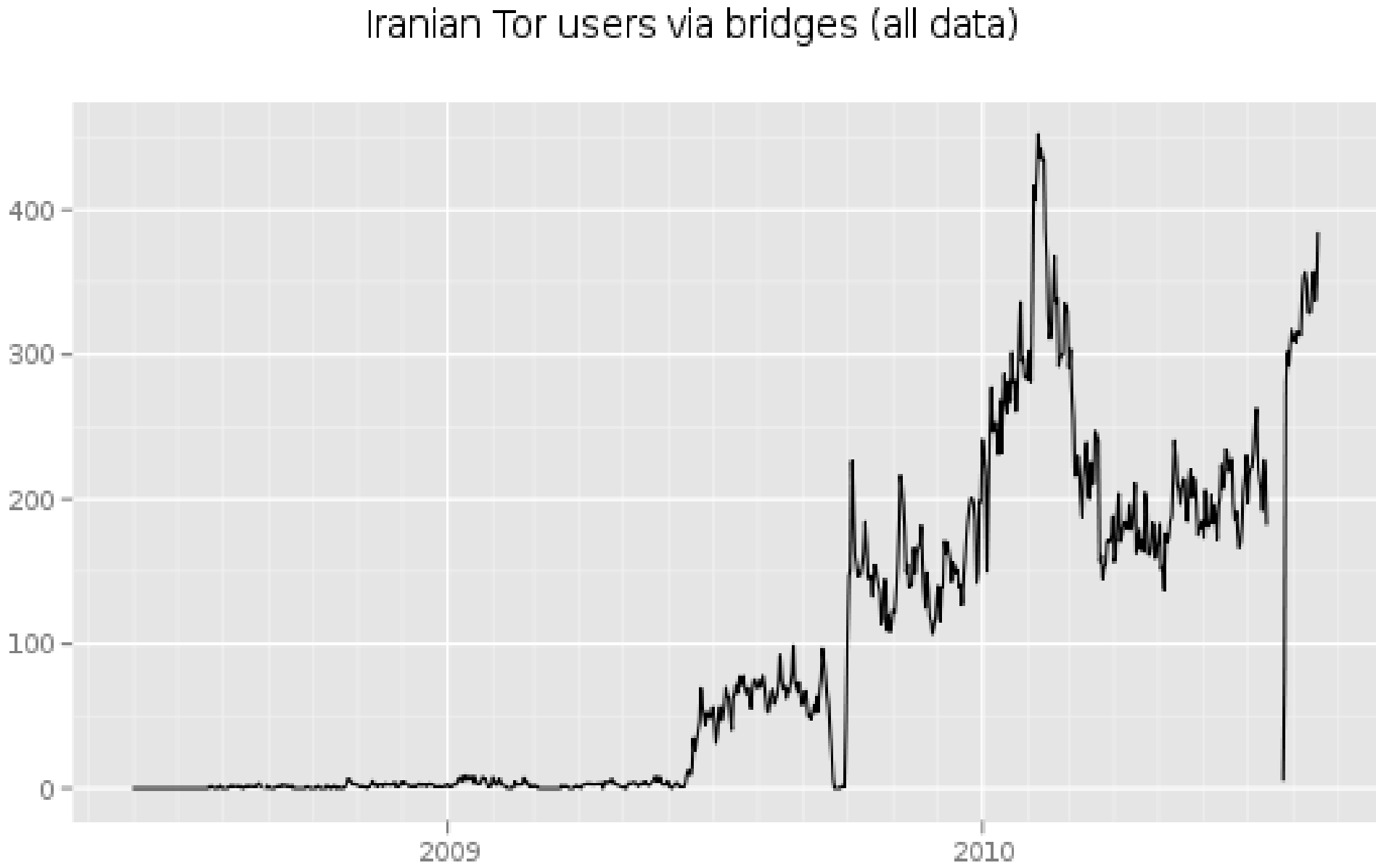
Our NSF EAGER

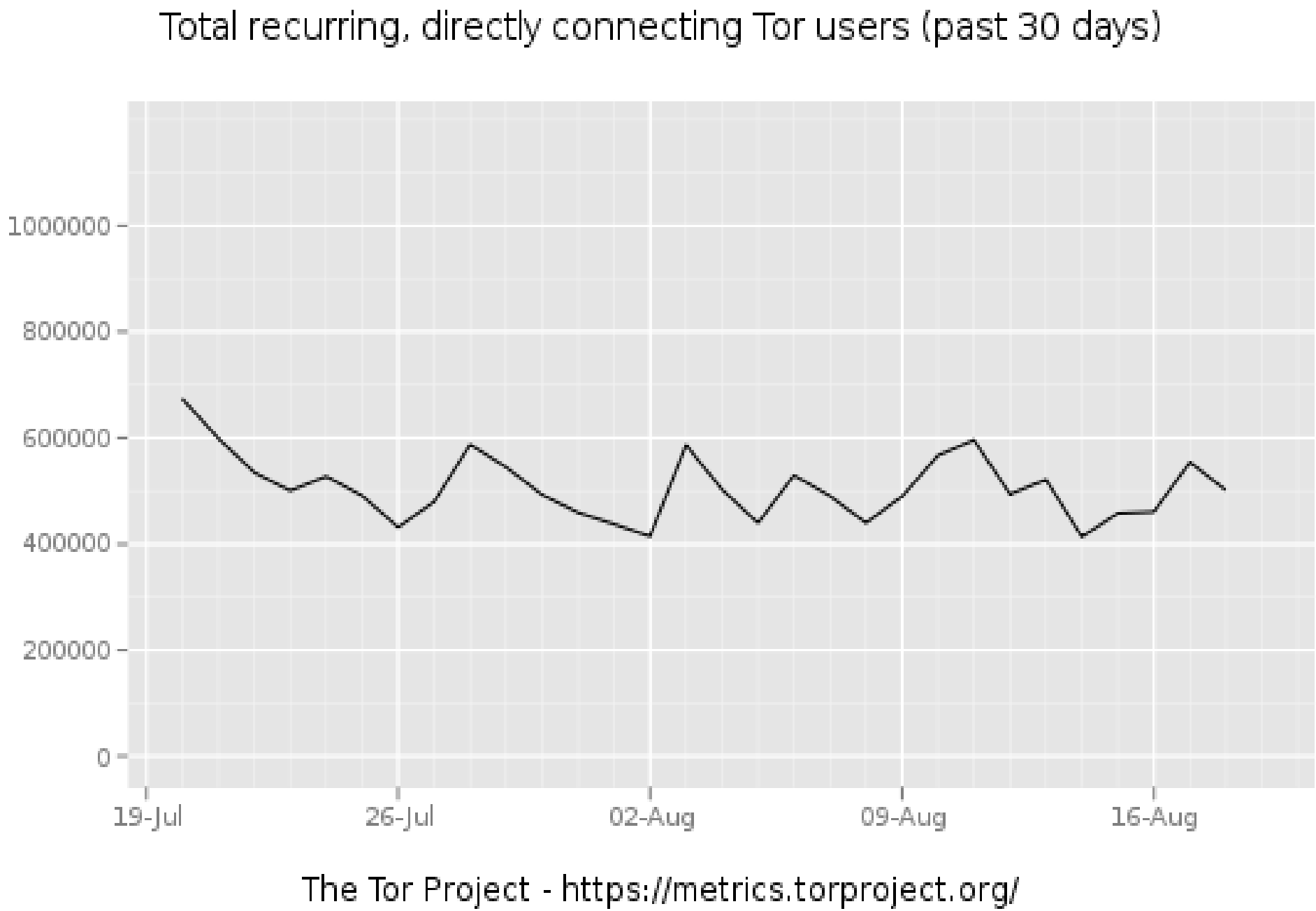
- 1) Invent and deploy new privacy-preserving algorithms to collect data about the Tor network, its performance, and its users
- 2) Publish this data, plus tools to analyze it
- 3) Figure out what else to measure and do it
- 4) Work with other research groups to make sure they get the data they need to solve the problems Tor actually has

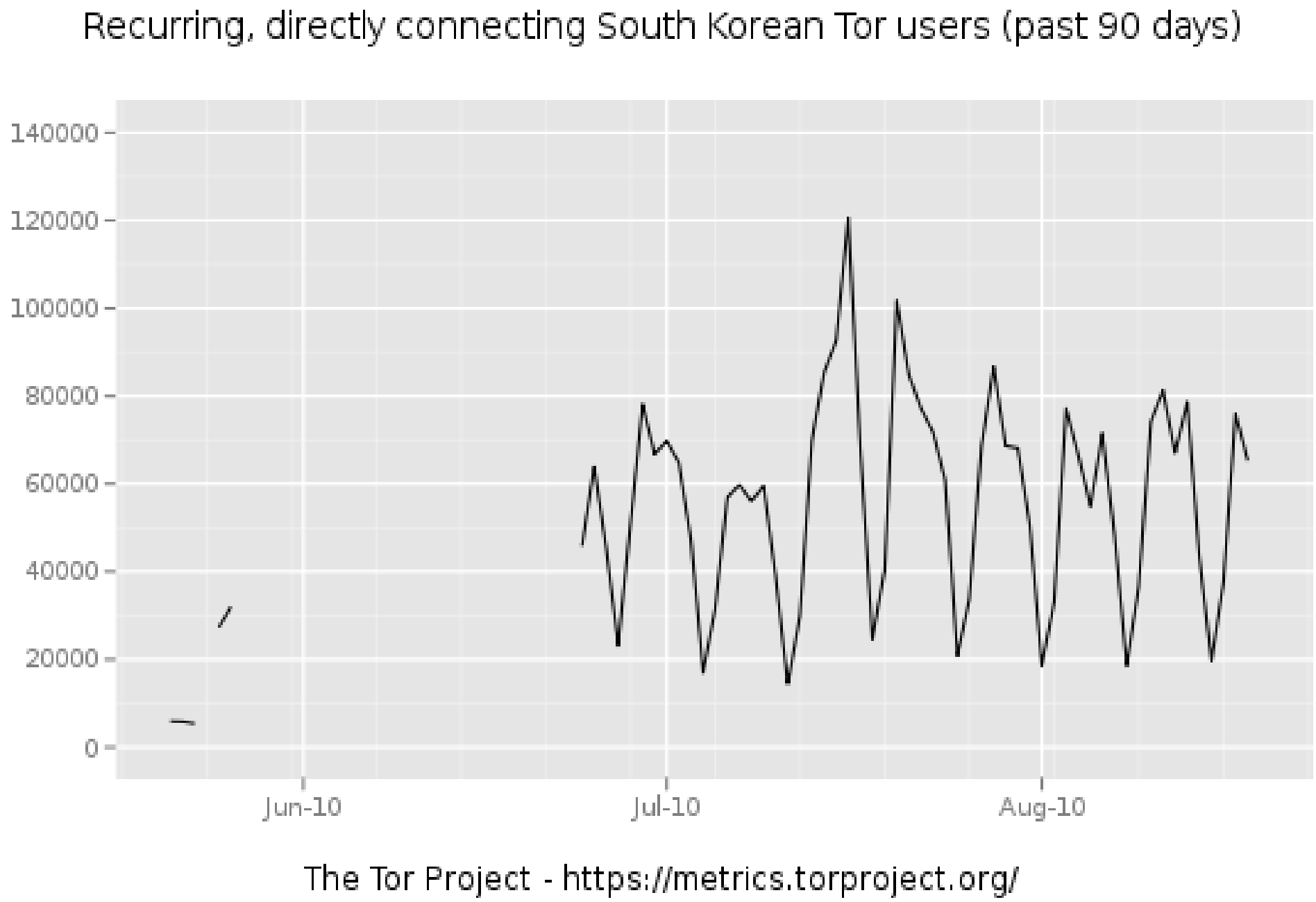

Recurring, directly connecting Chinese Tor users (all data)
35000
30000
25000
20000
15000
10000
5000
0
Jun-09
Sep-09
Dec-09
Mar-10
Jun-10
Sep-10
The Tor Project - https://metrics.torproject.org/

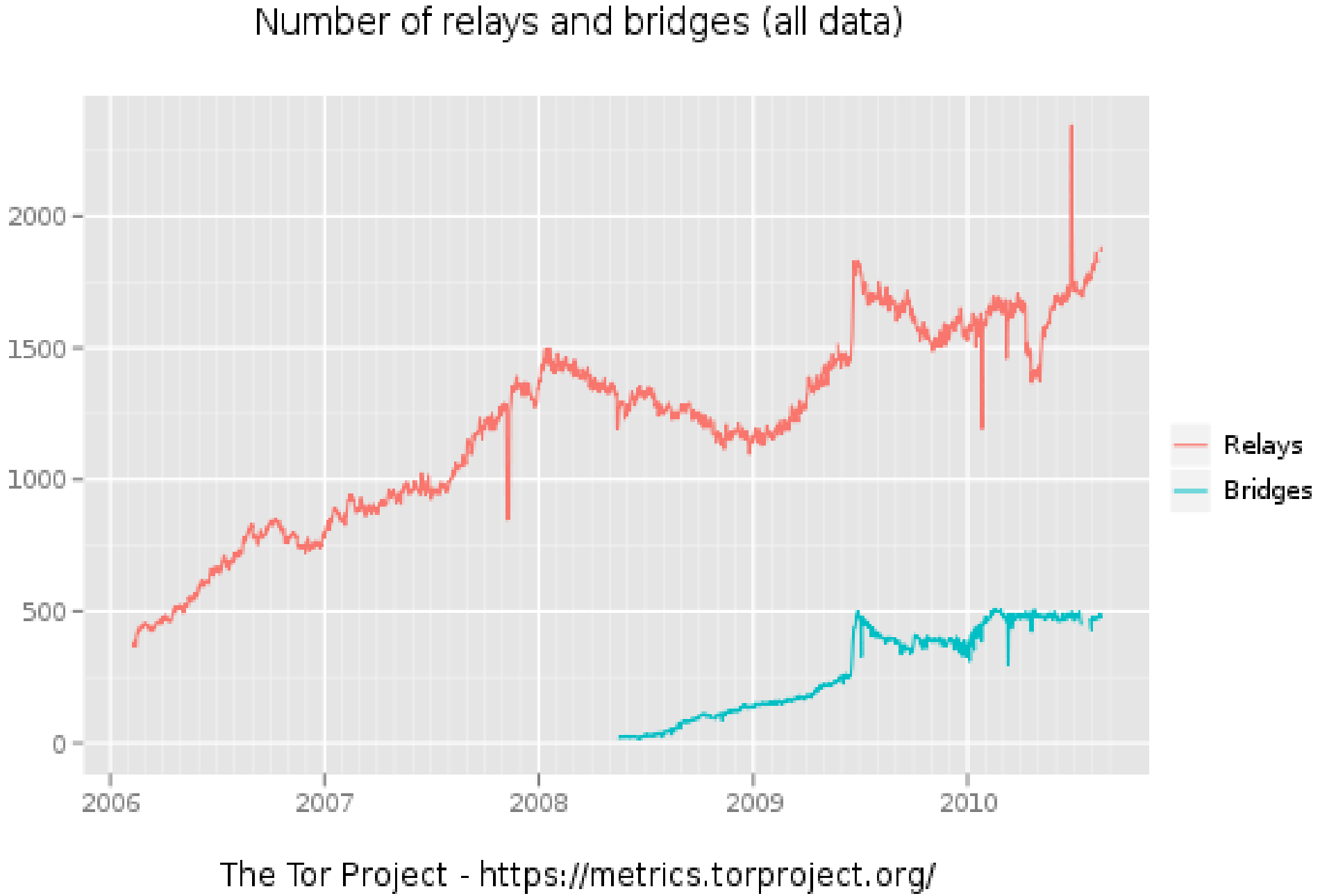

Recurring, directly connecting Iranian Tor users (all data)
20000
15000
10000
5000
0
Jun-09
Sep-09
Dec-09
Mar-10
Jun-10
Sep-10
The Tor Project - https://metrics.torproject.org/

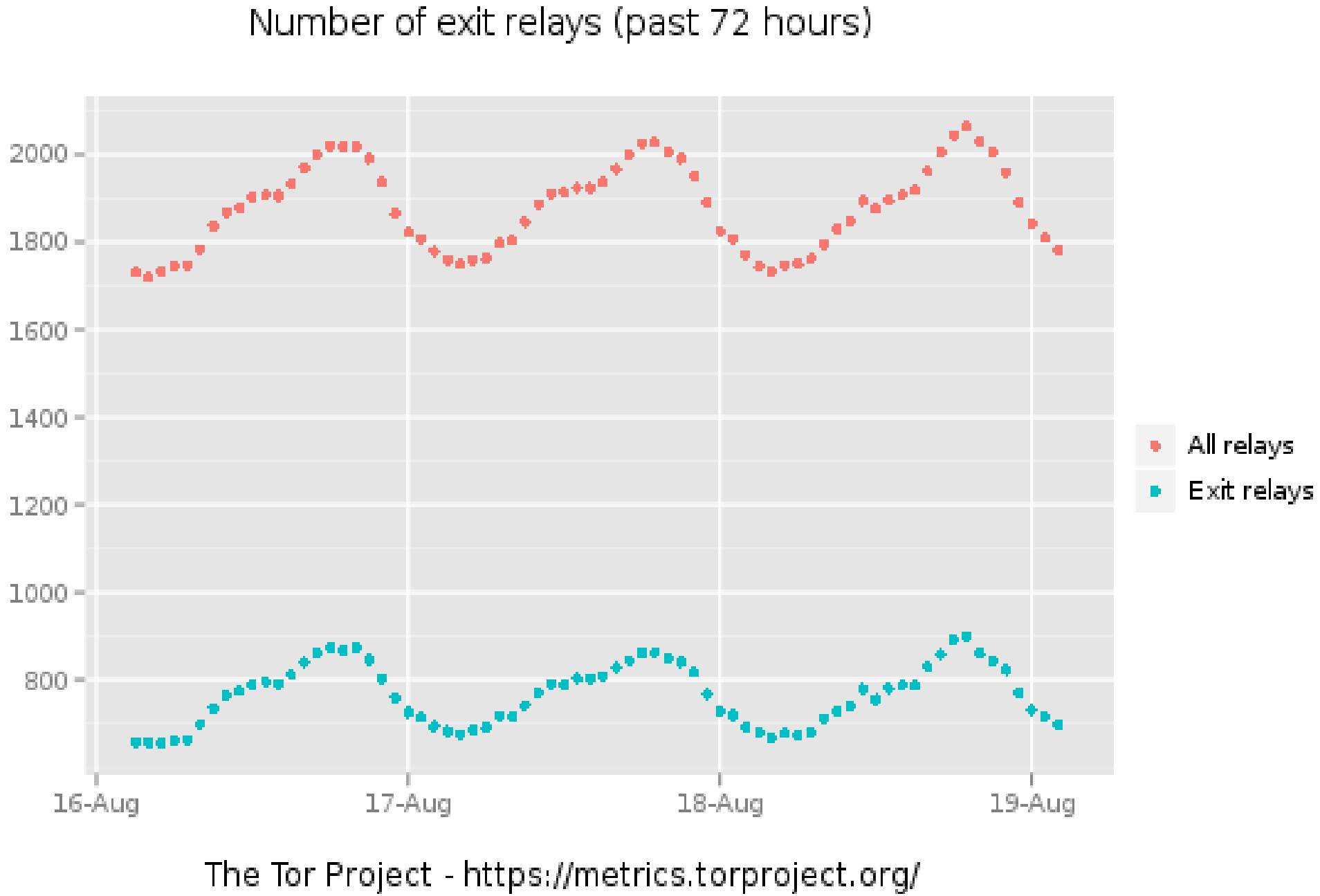

Chinese Tor users via bridges (all data)
30000
25000
20000
15000
10000
5000
0
2009
2010
The Tor Project - https://metrics.torproject.org/

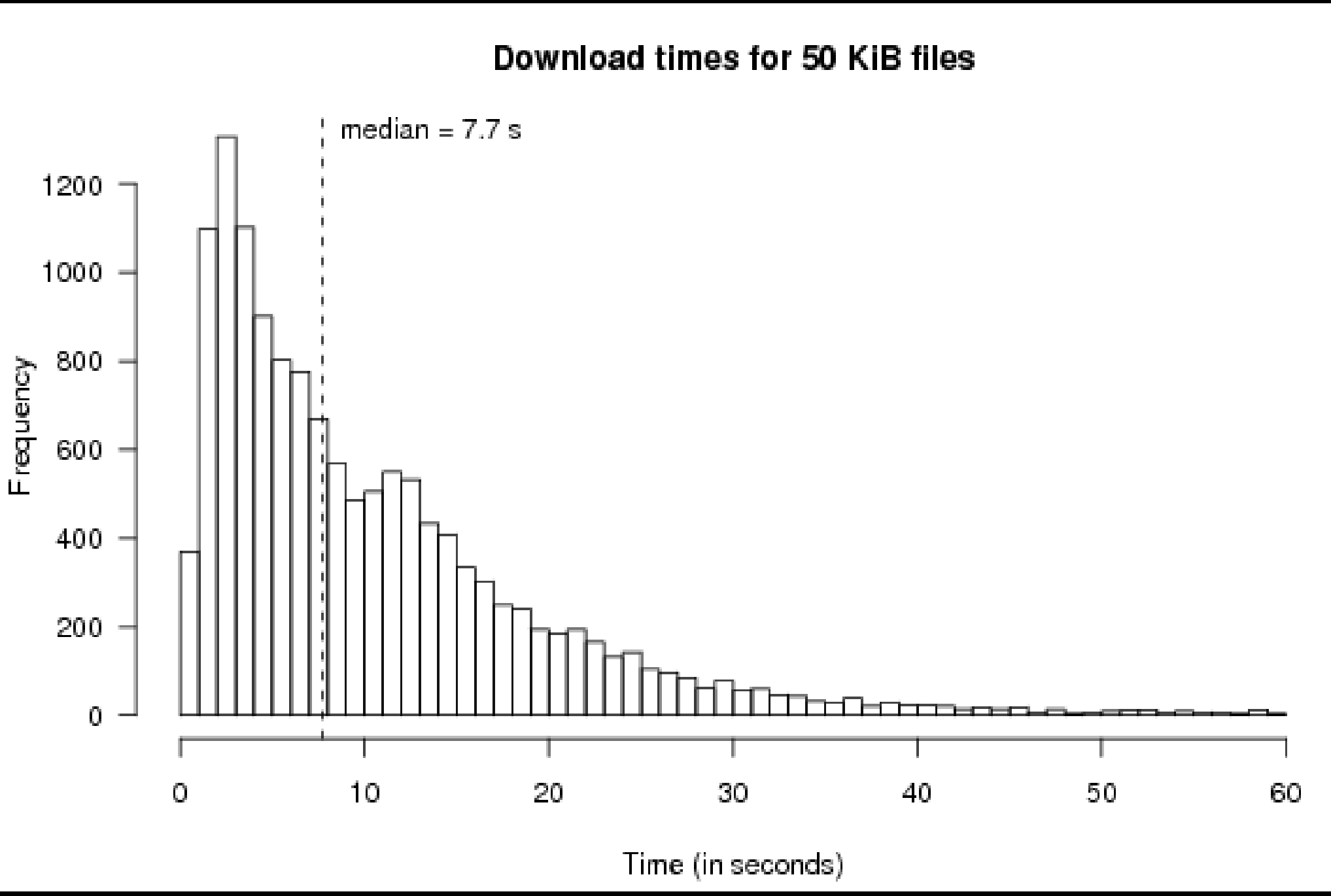

Iranian Tor users via bridges (all data)
400
300
200
100
0
2009
2010
The Tor Project - https://metrics.torproject.org/

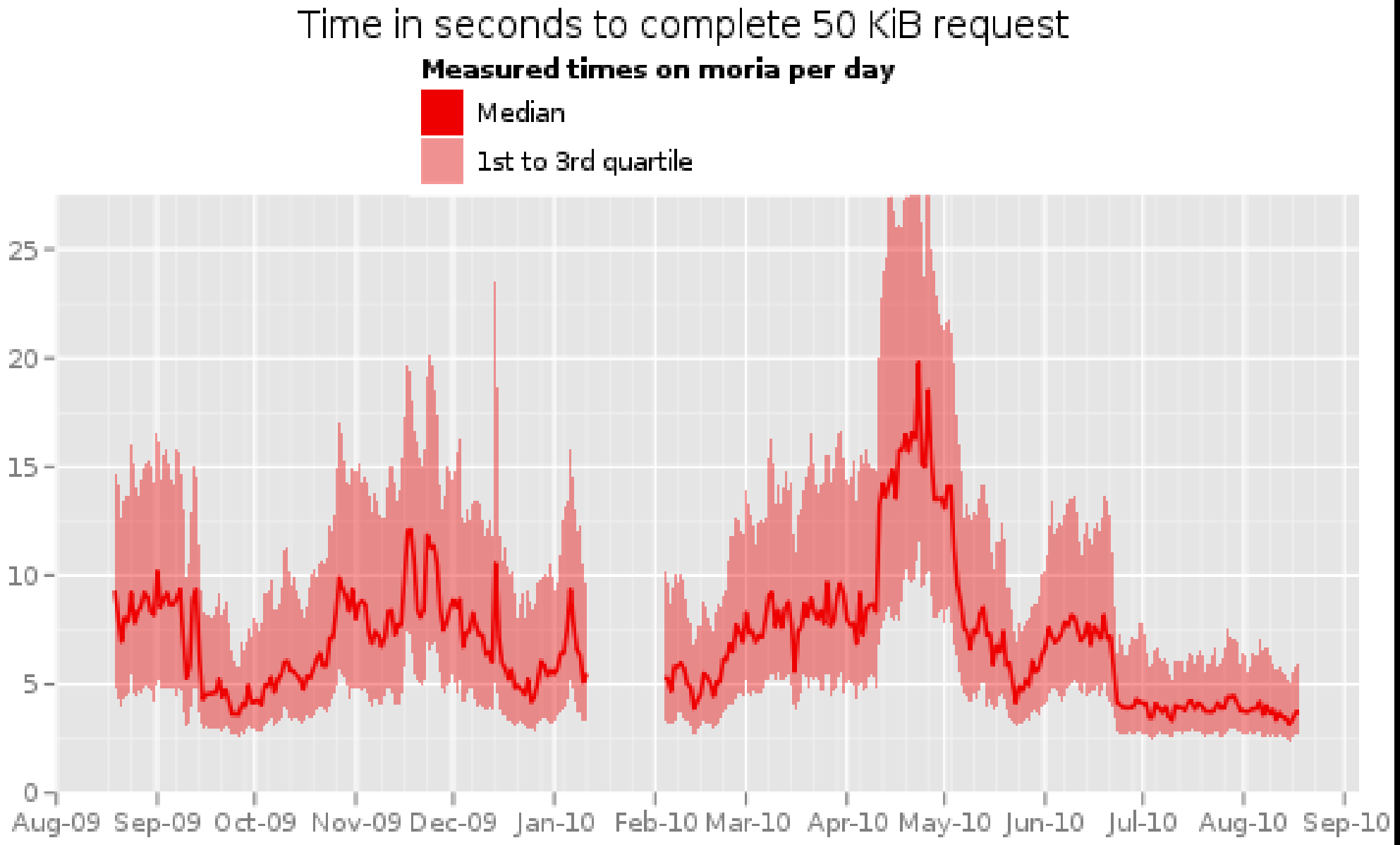

Total recurring, directly connecting Tor users (past 30 days)
1000000
800000
600000
400000
200000
0
19-Jul
26-Jul
02-Aug
09-Aug
16-Aug
The Tor Project - https://metrics.torproject.org/


Recurring, directly connecting South Korean Tor users (past 90 days)
140000
120000
100000
80000
60000
40000
20000
0
Jun-10
Jul-10
Aug-10
The Tor Project - https://metrics.torproject.org/

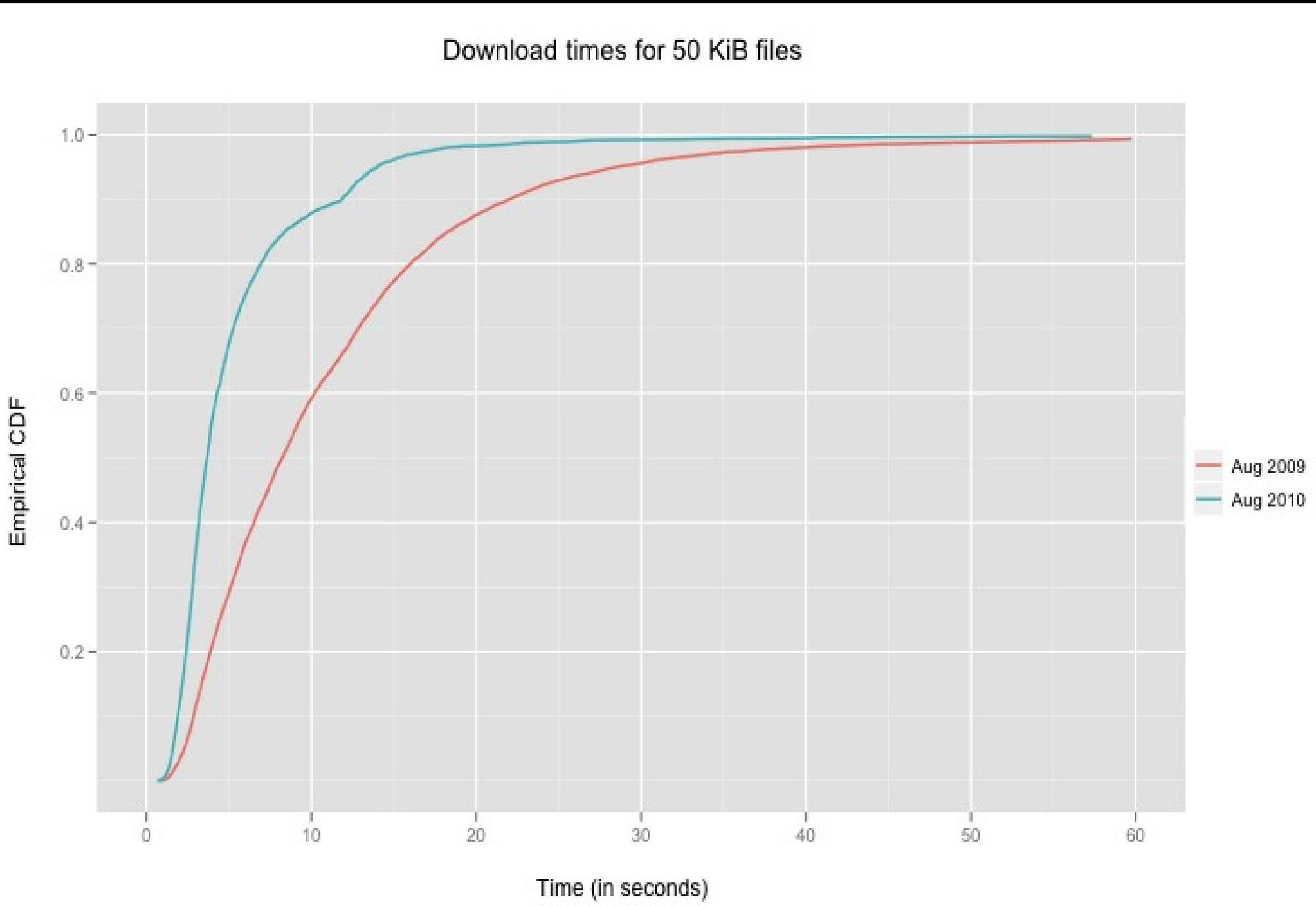

Number of relays and bridges (all data)
2000
1500
1000
500
0
2006
2007
2008
2009
2010
Relays
Bridges
The Tor Project - https://metrics.torproject.org/


Number of exit relays (past 72 hours)
2000
1800
1600
1400
1200
1000
800
16-Aug
17-Aug
18-Aug
19-Aug
All relays
Exit relays
The Tor Project - https://metrics.torproject.org/


Download times for 50 KiB files
median = 7.7 s
Frequency
0
200
400
600
800
1000
1200
0
10
20
30
40
50
60
Time (in seconds)



The Tor Project - https://metrics.torproject.org/


Download times for 50 KiB files
Empirical CDF
1.0
0.8
0.6
0.4
0.2
0
10
20
30
40
50
60
Time (in seconds)
Aug 2009
Aug 2010