Tor and circumvention: Lessons learned

Nick Mathewson
The Tor Project
https://torproject.org/
What is Tor?

Online anonymity 1) open source software, 2) network, 3) protocol

Community of researchers, developers, users, and relay operators

Funding from US DoD, Electronic Frontier Foundation, Voice of America, Google, NLnet, Human Rights Watch, NSF, US State Dept, SIDA, ...
The Tor Project, Inc.

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

Not secretly evil.
Estimated ~250,000 daily Tor users
Anonymity in what sense? “Attacker can’t learn who is talking to whom.”
Threat model: what can the attacker do?

Alice

Anonymity network

Bob

watch Alice!

Control part of the network!

watch (or be!) Bob!
Anonymity isn't cryptography: Cryptography just protects contents.

Alice

“Hi, Bob!”

<gibberish>

attacker

Bob

“Hi, Bob!”
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.

“It's privacy!”
Anonymity serves different interests for different user groups.

- "It's privacy!"
- "It's network security!"

Private citizens

Businesses
Anonymity serves different interests for different user groups.

“IT's traffic-analysis resistance!”

Governments

Anonymity

Private citizens

“IT's privacy!”

Businesses

“IT's network security!”
Anonymity serves different interests for different user groups.

- **Private citizens**: “It's privacy!”
- **Governments**: “It's traffic-analysis resistance!”
- **Businesses**: “It's network security!”
- **Human rights activists**: “It's reachability!”
Regular citizens don't want to be watched and tracked.

Hostile Bob

Incompetent Bob

Indifferent Bob

“I sell the logs.”

“Oh, I lost the logs.”

The AOL fiasco

“Hey, they aren't my secrets.”

The network can track too

Name, address, age, friends, interests (medical, financial, etc), unpopular opinions, illegal opinions...

Blogger Alice

8-year-old Alice

Sick Alice

Consumer Alice

Oppressed Alice

Alice

Alice

Alice

Alice
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.

"Why is alice.localpolice.gov reading my website?"

"Why no, alice.localpolice.gov! I would never sell counterfeits on ebay!"

"Is my family safe if I go after these guys?"

"Are they really going to ensure my anonymity?"
Governments need anonymity for their security

“Do I really want to reveal my internal network topology?”

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

“What about insiders?”

“What does FBI Google for?”

“Somebody in that hotel room just checked his Navy.mil mail!”
Journalists and activists need Tor for their personal safety

Activist/Whistleblower Alice

Monitoring ISP

Monitored website

Filtered website

Monitored network

“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...

Citizen Alice → Alice's small anonymity net → Municipal anonymity net → Investigated suspect

Officer Alice → Municipal anonymity net → Investigated suspect

AliceCorp → AliceCorp anonymity net → Competitor

“One of the 25 users on AliceNet.”

“Looks like a cop.”

“It's somebody at AliceCorp!”
... so, anonymity loves company!

Citizen Alice

Officer Alice

AliceCorp

Shared anonymity net

Investigated suspect

Competitor

“???”

“???”

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

Evil Criminal Alice

- Compromised botnet
- Stolen mobile phones
- Open wireless nets
- .....
Current situation: Bad people on the Internet are doing fine

- Trojans
- Viruses
- Exploits
- Botnets
- Zombies
- Espionage
- DDoS
- Extortion
- Spam
- Phishing
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
New or returning Tor clients per day

[Graph showing the number of new or returning Tor clients per day from June 1 to June 29, 2009. The green line represents China, and the red line represents Iran.]

https://torproject.org
Another Iran user count

Talked to chief security officer of one of the web 2.0 social networking sites:

10% (~10k) of their Iranian users in June 2009 were coming through Tor

90% (~90k) were coming from proxies in the Amazon cloud
Iran and DPI

We made Tor's TLS handshake look like Firefox+Apache.

When Iran kicked out Smartfilter in early 2009, Tor's old (non-TLS) directory fetches worked again!

Jan 2011, Iran blocked Tor by DPI for SSL and filtering our Diffie-Hellman parameter. Socks proxy worked fine the whole time.
Directly connecting Iranian Tor users

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

Servers publish self-signed descriptors.

Authorities publish a consensus list of all descriptors.

Alice downloads consensus and descriptors from anywhere.
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
“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
How do you find a bridge?

1) https://bridges.torproject.org/ will tell you a few based on time and your IP address

2) Mail bridges@torproject.org from a gmail address and we'll send you a few

3) We mail some to a friend in Shanghai who distributes them via his social network

4) You can set up your own private bridge and tell your target users directly
Number of directory requests to directory mirror trusted

https://torproject.org

China
Chinese Tor users via bridges
Chinese users via bridges

The Tor Project - https://metrics.torproject.org/
Directly connecting Chinese Tor users

The Tor Project - https://metrics.torproject.org/
Directly connecting Tunisian Tor users

The Tor Project - https://metrics.torproject.org/
Directly connecting Egyptian Tor users

The Tor Project - https://metrics.torproject.org/
Directly connecting Saudi Tor users

The Tor Project - https://metrics.torproject.org/
Directly connecting Saudi Tor users

The Tor Project - https://metrics.torproject.org/
Saudi users via bridges

The Tor Project - https://metrics.torproject.org/
Censoring attacker's goals

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.
What we're up against

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?
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

Assume the users aren't attacked by their hardware and software
No spyware installed, no cameras watching their screens, etc
Users can fetch a genuine copy of Tor?
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 affect 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?
Trust and reputation

See January 2009 blog post by Hal Roberts about how some circumvention tools sell user data

Many of these tools see circumvention and privacy as totally unrelated goals
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, ...
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
Next steps (policy)

Technical solutions won't solve the whole censorship problem. After all, firewalls are *socially* successful in these countries. But a strong technical solution is still a critical puzzle piece.

You should run a bridge! We only have ~750. We'd love to help with some trainings, to help users and to make Tor better.
BridgeDB needs a feedback cycle

Measure how much use each bridge sees
Measure bridge blocking
Then adapt bridge distribution to favor efficient distribution channels
Need to invent new distribution channels
Need more and changing bridge addresses
Redirecting a whole /16 ?
Promote clients to bridges?
Measuring bridge reachability

Passive: bridges track incoming connections by country

Active: scan bridges from within the country

Clients self-report blockage (via some other bridge)

Measure remotely via FTP reflectors

Bridges test for duplex blocking
Other components

Traffic camouflaging

Superencrypt so no recognizable bytes?
Shape like HTTP?

We're working on a modular transport API

Client-side automation for usability

Performance / scalability

Especially for low bandwidth
Questions?

See also:

- https://torproject.org/
- http://freehaven.net/anonbib/
- Specs, design papers, open proposals, etc
- Public mailing lists
- The code is open source
Time in seconds to complete 50 KiB request

Measured times on all sources per day

- Median
- 1st to 3rd quartile

The Tor Project - https://metrics.torproject.org/
Total relay bandwidth

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