Tor and circumvention: Lessons learned

Roger Dingledine
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
Estimated 250,000 daily Tor users
Threat model: what can the attacker do?

Alice

Anonymity network

Control part of the network!

Bob

watch Alice!

watch (or be!) Bob!
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.

“It's privacy!”

Private citizens
Anonymity serves different interests for different user groups.

“Anonymity" -> "Businesses"

"Businesses"

"It's network security!"

"Anonymity" -> "Private citizens"

"Private citizens"

"It's privacy!"
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.

“It's traffic-analysis resistance!”

Governments

Anonymity

Human rights activists

“It's reachability!”

Businesses

“It's network security!”

Private citizens

“It's privacy!”
Regular citizens don't want to be watched and tracked.

Hostile Bob

Incompetent Bob

Indifferent Bob

"I sell the logs."

"Oops, I lost the logs."

"Hey, they aren't my secrets."

The AOL fiasco

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

(the network can track too)
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

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

“What does FBI Google for?”

“What about insiders?”

“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...

- Citizen Alice
  - Alice's small anonymity net
  - Municipal anonymity net
  - Investigated suspect
  - "One of the 25 users on AliceNet."

- Officer Alice
  - Municipal anonymity net
  - "Looks like a cop."

- AliceCorp
  - AliceCorp anonymity net
  - Competitor
  - "It's somebody at AliceCorp!"
... so, anonymity loves company!
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.

Alice1

E(Bob3, “X”)

E(Bob2, “Z”)

E(Bob1, “Y”)

Evil Relay

Bob1

Bob2

Bob3

Alice2

Alice3
... 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
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) I 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
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 Libyan 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/
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 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?
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.
I CAN HAZ FREEDOM?

Tor
TorProject.org
How you can help

Teach your friends about Tor, why it's useful, what it does and what it doesn't do.
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.
Design more bridge distribution ideas
Work for us over the summer?
More discussion Tuesday night: www.tossug.org
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

Super-encrypt 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
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/
Download times for 50 KiB files

Empirical CDF

Time (in seconds)