Anonymity

Michalis Polychronakis

Stony Brook University
Privacy

“The right of an entity (normally a person), acting in its own behalf, to determine the degree to which it will interact with its environment, including the degree to which the entity is willing to share information about itself with others.” [RFC2828]

Anonymity

“The state of being not identifiable within a set of subjects, the anonymity set.” [Pfitzmann and Köhntopp]

Very different from privacy:

An anonymous action may be public, but the actor’s identity remains unknown (e.g., vote in free elections)
Web censorship: the net is closing in

Across the globe governments are monitoring and censoring access to the web. And if we’re not careful millions more people could find the internet fractured, fragmented and controlled by the state.
China toughens web censorship, encourages others to follow

Beijing sees internet controls as a 'question of sovereignty'

WATARU KODAKA, Nikkei staff writer
New European Copyright Enforcement Plans Loom Large Even as Users Revolt Against Filter Proposal

BY JEREMY MALCOLM | NOVEMBER 29, 2017

EFF has joined over 80 groups in writing once again [PDF] to European politicians about disastrous new EU copyright proposals. Along with human and digital rights organizations, media freedom organizations, publishers, journalists, libraries, scientific and research institutions, educational institutions including universities, creator representatives, consumers, software developers, start-ups, technology businesses and Internet service providers, we wrote

to share our respectful but serious concerns that discussions in the Council and European Commission on the Copyright Directive are on the verge of causing irreparable damage to our fundamental rights and freedoms, our economy and
INTERNET CENSORSHIP IS ADVANCING UNDER TRUMP
Anonymous communication

Sender anonymity
The identity of the party who sent a message is hidden, while its receiver (and the message itself) might not be

Receiver anonymity
The identity of the receiver is hidden

Unlinkability of sender and receiver
Although the sender and receiver can each be identified as participating in some communication, they cannot be identified as communicating with each other
The internet was not designed for anonymity

Packets have source and destination IP addresses

Using pseudonyms to post anonymously is not enough…
    Server always sees the IP address of the client

Need to hide the source IP address

Assuming no other PII is revealed (!) – *OPSEC is hard*
Stepping Stones: Anonymity

Proxies, relays, VPN servers

Server sees only the IP address of the relay
Since the relay cooperates, let’s also encrypt the connection to it

Sender anonymity against the server and network observers beyond the relay

Also: receiver anonymity against local observers
All they can see is client ⇔ relay connections
Encrypted tunnel hides the actual destination
Stepping Stones: Traffic Protection

Besides anonymity, the encrypted client ⇔ relay channel offers protection against local adversaries.

The definition of “local” depends on the location of the proxy:
- Users in the same LAN, employer’s admins, ISPs, governments, …

Protection against passive and active network adversaries (eavesdropping, MitM, MotS, …)
  - In addition to any end-to-end encryption (e.g., TLS)

Policy and censorship circumvention
  - Parental controls, company-wide port/domain/content blocking, hotel WiFi restrictions, government censorship, …
What about other adversaries?

The relay itself may be the adversary – can see it all!

Network observers beyond the relay can see it all!

Adversaries who couldn’t eavesdrop before, now can: just set up a rogue proxy/VPN server and lure users

End-to-end encryption is critical!
What about other adversaries?

A “global” adversary may be able to observe both ends

**Traffic analysis:** communication patterns can be observed even when end-to-end encryption is used
Eavesdropping vs. Traffic Analysis

Even when communication is encrypted, the mere fact that two parties communicate reveals a lot

Example: what can we learn from phone records?
  Who communicated with whom and when
  Activity patterns (periodic, time of day, occasional, …)
  Single purpose numbers (hotlines, agencies, doctors, …)

*It’s not “just metadata”…*

Network traffic analysis can reveal a lot
Passive traffic analysis

Frequency and timing of packets, packet sizes, amount of transferred data, …

Active traffic analysis

Packet injection, fingerprint injection through manipulation of traffic characteristics, …

Examples:

Message timing correlation to learn who is talking to whom
Visited HTTPS web pages through structural analysis (number/size of embedded elements etc.)
SSH keystroke timing analysis

“Traffic analysis, not cryptanalysis, is the backbone of communications intelligence.”

— Susan Landau and Whitfield Diffie
Mix Networks [Chaum 1981]

Main idea: hide own traffic among others’ traffic

Originally conceived for anonymous email: Trusted reemailer + public key cryptography

Additional measures are critical for thwarting traffic analysis: message padding, delayed dispatch, dummy traffic
Adding multiple mix relays allows for anonymity even if some relays are controlled by an adversary.

Deanonymization still possible if the adversary controls *all* relays of a circuit.

Main drawback: prohibitively high latency for interactive communication.
Low-latency anonymous communication network

Layered encryption: each relay decrypts a layer of encryption to reveal only the next relay
Worldwide volunteer network of ~7K relays
~3M daily users

Three-hop circuits by default
  Entry node, middleman, exit node
  Longer circuits can be built
  Multiple connections can be multiplexed over the same Tor circuit

Directory servers point to active Tor relays
  10 directory servers hard-coded into the Tor client
  Monitoring for mass subscriptions by potential adversaries (sybil attack)
Applications

User-friendly Tor Browser
   Additional measures to thwart web tracking and fingerprinting

TAILS (The Amnesic Incognito Live System) Linux distribution
   Forces all outgoing connections to go through Tor

Onion services: hide the IP address of servers
   .onion pseudo top-level domain host suffix
   Not always easy: misconfigurations and leaks may reveal the real IP address of the server

SecureDrop (originally designed by Aaron Swartz)
   Platform for secure anonymous communication between journalists and sources (whistleblowers)

Many more: OnionShare (file sharing), Ricochet (IM), …
Step 1: Bob picks some introduction points and builds circuits to them.
**Step 2:** Bob advertises his hidden service -- XYZ.onion -- at the database.

**Onion addresses are self-authenticating:** derived from the service’s public key
**Hidden Services: 3**

**Step 3:** Alice hears that XYZ.onion exists, and she requests more info from the database. She also sets up a rendezvous point, though she could have done this before.

- **Tor cloud**
- **Tor circuit**
- **IP1-3** Introduction points
- **PK** Public key
- **cookie** One-time secret
- **RP** Rendezvous point

---

© Tor Project - https://www.torproject.org/docs/hidden-services.html.en
Step 4: Alice writes a message to Bob (encrypted to PK) listing the rendezvous point and a one-time secret, and asks an introduction point to deliver it to Bob.
Step 5: Bob connects to the Alice's rendezvous point and provides her one-time secret.
Hidden Services: 6

Step 6: Bob and Alice proceed to use their Tor circuits like normal.
1 Million People use Facebook over Tor

People who choose to communicate over Tor do so for a variety of reasons related to privacy, security and safety. As we've written previously it's important to us to provide methods for people to use our services securely – particularly if they lack reliable methods to do so.

This is why in the last two years we built the Facebook onion site and onion-mobile site, helped standardise the "onion" domain name, and implemented Tor connectivity for our application.
Censors want to block Tor

Directory servers are the easy target
   Block any access to them

Response: Tor bridges
   Tor relays that aren't listed in the main Tor directory
   Only a few at a time can be obtained on-demand (e.g., through email to bridges@bridges.torproject.org)
   Once known, adversaries may block them too...

Pluggable Transports
   Censors may drop all Tor traffic through deep packet inspection
   Hide Tor traffic in plain sight by masquerading it as some other innocent-looking protocol (HTTP, Skype, Starcraft, …)
Critical Tor flaw leaks users’ real IP address—update now

TorMoil threatens Mac and Linux versions of Tor browser; Windows and Tails not affected.

DAN GOODIN - 11/3/2017, 6:30 PM

Having a compartmentalized setup is important!

Running Tor Browser on a separate device is not enough

The device should be configured so that it cannot access the public internet, but instead always route all its traffic through Tor
FBI Admits It Controlled Tor Servers Behind Mass Malware Attack

BY KEVIN POULSEN 09.13.13 | 4:17 PM | PERMALINK
For a year, gang operating rogue Tor node infected Windows executables

Attacks tied to gang that previously infected governments with highly advanced malware.

by Dan Goodin - Nov 14, 2014 10:30am EST
How Hackers Abused Tor To Rob Blockchain, Steal Bitcoin, Target Private Email And Get Away With It

Across October and November of last year, some unlucky users of the world’s most popular Bitcoin wallet, Blockchain.info, and one of the better-known exchanges, LocalBitcoins, had their usernames and passwords silently pilfered. They were robbed of significant sums, probably tens of thousands of dollars worth of the virtual currency, possibly more. Security-focused email services, Riseup and Safe-mail were also targeted by the same crew. And according to the man who witnessed the attacks go off last year, Digital Assurance director Greg Jones, it looks like buyers and sellers of dark markets were the targets.

The attackers used a tried-and-tested method to begin with, setting up a number of malicious exit relays on Tor. Legitimate exit relays act as the final jump from the anonymising Tor network, which loops users through a number of randomly-chosen servers across the world to protect their identity, onto the clear web. But any nefarious type who runs a malicious relay can use an encryption removal technique known as SSL stripping, where connections are...
Detecting Traffic Snooping in Tor using Decoys

Expose unique decoy username+password through each exit node
Wait for unsolicited connections to the honeypot server using any of the exposed bait credentials
Detected Rogue Exit Nodes

30-month period: detected **18 cases** of traffic eavesdropping that involved **14 different Tor exit nodes**
Online Privacy and Anonymity: What Can We do?

Technical solutions exist

- Encryption
- Self-hosted services
- Anonymous communication

... But they are not enough

- Privacy vs. usability tradeoff
- Wrong assumptions
- Implementation flaws

Many users are not even aware of privacy issues, let alone solutions

Protect the right of individuals to control what information related to them may be collected

With technical means, not promises...