Privacy, Anonymity, and Censorship

CS155 Computer and Network Security
Remote Access
Virtual Private Networks (VPNs)

**Problem:** How do you provide secure communication for non-TLS protocols across the public Internet?

VPNs create a fake shared network on which traffic is encrypted

Two Broad Types:
- Remote client (e.g., traveler with laptop) to corporate network
- Connect two remote networks across Internet
Several VPN protocols exist (PPTP, L2TP, IPsec, OpenVPN). Most popular is IPsec. OpenVPN is open source.
Wireguard

Recently introduced VPN that has gained significant following in the past 5 years over options like OpenVPN:

• Simpler protocol and much more performant than OpenVPN. Relatively few configuration options reduces opportunity for error

• Utilizes modern cryptographic primitives like Noise protocol framework, Curve25519, ChaCha20, Poly1305,
Cisco AnyConnect

Stanford and many other organizations use Cisco AnyConnect

Encapsulates traffic in TLS! Initial handshake uses normal TCP-based TLS for initial handshake and then DTLS (UDP-based TLS) to transport data.
Gooey Middle

VPNs support the idea of having a secure internal network and untrusted public Internet. Unfortunately, attacker has a ton of access once the network perimeter is breached.

Unfortunately, internal networks aren’t *that* secure. Computers are compromised all the time and attackers have free rein.
Google: assume internal network is *also* out to get you. Remove privileged intranet and put all corporate applications on the Internet.

Access depends solely on device and user credentials, regardless of a user’s network location

Protect applications, not the network
Privacy
What is Privacy?

Privacy is control over your own information. Freedom from intrusion into personal matters

Privacy is a person’s right or expectation to control the disclosure of his/her personal information, including activity metadata

Privacy is the “right to be let alone” — Louis Brandeis

Privacy means something like what the Founders meant by “liberty”

  Free speech, free association, autonomy, …
  freedom from censorship and constant surveillance

Privacy-motivating examples in U.S. History

  Martin Luther King Jr. “blackmailed” by FBI
  McCarthyism witch-hunt for communists
Direct Sharing

The Incredible Story Of How Target Exposed A Teen Girl's Pregnancy

GUS LUBIN
FEB. 16, 2012, 10:27 AM

Target broke through to a new level of customer tracking with the help of statistical genius Andrew Pole, according to a New York Times Magazine cover story by Charles Duhigg.

Pole identified 25 products that when purchased together indicate a woman is likely pregnant. The value of this information was that Target could send coupons to the pregnant woman at an expensive and habit-forming period of her life.

Plugged into Target's customer tracking technology, Pole's formula was a beast. Once it even exposed a teen girl's pregnancy:
First Response Early Result Pregnancy Test, 3 tests, Packaging May Vary
by First Response

4.5 out of 5 stars 486 customer reviews | 17 answered questions

47 Amazon Students rated this highly

List Price: $49.67
Price: $12.98 Prime & Free Returns. Details
You Save: $36.69 (74%)

Note: Available at a lower price from other sellers, potentially without free Prime shipping.

In Stock.
Ships from and sold by Amazon.com. Gift-wrap available.

Want it Tuesday, March 24? Order within 29 hrs 56 mins and choose One-Day Shipping at checkout. Details

Package Quantity: 1

1
$12.98 Prime
2
$41.00 Prime
3
$53.99

Customers Who Bought This Item Also Bought

Vitafusion Prenatal DHA and Folic Acid Gummy Vitamins, 180 Count
$20.25 Prime

One A Day Women's Prenatal One Pill, 30 Count
$13.48 Prime

Meyers Clinic Guide to a Healthy Pregnancy: The Pregnancy Experts... in Motherhood
$21.79 Prime

Summer's Eve Cleansing Wash, Morning Paradise, 19 Ounce
$3.99

Neocare 524580 Basal Digital Thermometer
$14.06 Prime

Nature Made Prenatal Multi Vitamin Value Size, Tablets, 250-Count
$16.79 Prime

Trojan Condom Pleasure Pack Lubricated, 40 Count
$18.12 Prime
Third Party Tracking

1. First of all, define "cat lady." Does one cat = cat lady? Two cats = cat lady? Does joking about being a cat lady à la sparkling, outgoing multimillionaire Taylor Swift automatically make one a cat lady? It is my personal belief that most female cat owners below the age of 40 fall into the "not a cat girl, not yet a cat lady" category.

2. Cat ladies mostly look like ... normal ladies. You know. Like regular women. Not like the old hat who sits in front of your local Shop Rite with aluminum foil on her head.
Third Party Cookies

- Site A’s page requests a third-party resource (image, script, iframe)
  - Normally, browser sends cookie associated with that third-party in that request

Cookie: ID=784c39
Referer: cnn.com/
Third-Party Web Tracking
Cookies and Code

(CNN)

GET / HTTP/3

GET /facebook-like.js HTTP/3
Cookie: User=Deepak, Referer=cnn.com

(Facebook)

• With this request, companies can link your cookie to your browsing data (e.g., through Referer header, Host headers, Origin, or just JavaScript)
Web Tracking
Cookies and Code

- What exactly is sent in the referer?
Trump escalates January 6 cover-up

The former President is trying to keep the House select committee probing January 6 from seeing a list of documents as he ramps up his political comeback

Brian Stelter’s ominous prediction: Imagine it’s 2022 and …

January 6 committee is losing patience with Trump’s former chief of staff Mark Meadows as it seeks his testimony

Washington Post report rebuts the January 6 alt-reality that Tucker Carlson promotes

Biden says US ‘continuing to suffer’ from Trump’s decision to pull out of Iran nuclear deal

Astros top Braves 9-5 in World Series Game 5

- Trivia: Can you name the only player to play in all 4 cities that the Braves have called home?
- Analysis: The Braves win the World Series. But they’re striking out with some fans

Students are fed up with raging adults at school board meetings

- A Texas lawmaker is investigating 850 books on race and gender that could cause ‘discomfort’ to students
- Opinion: When parents scream at school board meetings, how can I teach their children?

Southwest launches investigation into pilot reportedly using anti-Biden phrase on flight

- Reporter reveals what Lindsey Graham said during January 6 riot
- White House press secretary tests positive for Covid, last saw Elden Tuesday
- BREAKING: Japan’s Fumio Kishida defies expectations as ruling party keeps majority
- Aurora borealis puts on a gorgeous show
- ‘Step up or step out!’: Lawmaker calls out attorney general
- Police investigating desecration of Torah scroll at fraternity
- COP26 climate talks off to an ominous start after weak G20 leaders’ meeting
- Video shows passengers fleeing knife attack on train
### Third Party Cookies

Facebook, DoubleClick, etc. know much more about you than actual website does because they can track you across websites.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Top 1M</th>
</tr>
</thead>
<tbody>
<tr>
<td>google-analytics.com</td>
<td>67.8%</td>
</tr>
<tr>
<td>gstatic.com</td>
<td>50.1%</td>
</tr>
<tr>
<td>fonts.googleapis.com</td>
<td>42.8%</td>
</tr>
<tr>
<td>doubleclick.net</td>
<td>40.5%</td>
</tr>
<tr>
<td>facebook.com</td>
<td>33.7%</td>
</tr>
<tr>
<td>google.com</td>
<td>33.2%</td>
</tr>
<tr>
<td>facebook.net</td>
<td>27.4%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Domain</th>
<th>Top 1M</th>
</tr>
</thead>
<tbody>
<tr>
<td>ajax.googleapis.com</td>
<td>23.1%</td>
</tr>
<tr>
<td>googleapis.com</td>
<td>19.6%</td>
</tr>
<tr>
<td>googlesyndication.com</td>
<td>14.1%</td>
</tr>
<tr>
<td>googleadservices.com</td>
<td>12.8%</td>
</tr>
<tr>
<td>twitter.com</td>
<td>10.7%</td>
</tr>
<tr>
<td>fbcdn.net</td>
<td>10.5%</td>
</tr>
<tr>
<td>adnxs.com</td>
<td>10.5%</td>
</tr>
</tbody>
</table>
Ghostery found 13 trackers

- ClickTale
  Analytics, Analytics, Behavior Tracking
- DoubleClick
  Advertising
- Facebook Connect
  Widgets, Social
- Gravity Insights
  Analytics
- Krux Digital
  Beacons
- Livefyre
  Widgets, Commenting System
Do Not Track

Mozilla is a global, nonprofit organization dedicated to making the Web better. We emphasize principle over profit, and believe that the Web is a shared public resource to be cared for, not a commodity to be sold. We answer to no one but you and believe it is crucial to put you in control of your online experience. We are aiming to give you better insight and control into the ways your personal information is collected, used, stored and shared online.

Mozilla Firefox offers a Do Not Track feature that lets you express a preference not to be tracked by websites. When the feature is enabled, Firefox will tell advertising networks and other websites and applications that you want to opt-out of tracking for purposes like behavioral advertising.
2024 — The Year of the End of Third Party Cookies?

- Firefox:
  - Third-Party Cookies from known trackers are dropped
  - Third-party cookies use separate cookie jar per site, so they can't be used to track users across sites

- Safari: Blocks third-party cookies

- IE: blocks some third-party cookies based on baked-in blacklist

- Edge does not block third-party cookies by default

- Chrome announced that they will drop support for third party cookies by the end of 2024
Google Topics

<table>
<thead>
<tr>
<th>User's browser</th>
<th>User's browser</th>
<th>Site that displays ads</th>
<th>Adtech code</th>
<th>Adtech code</th>
<th>Adtech code</th>
</tr>
</thead>
<tbody>
<tr>
<td>User visits websites</td>
<td>Browser infers topics of interest</td>
<td>User visits site displaying ads</td>
<td>Topics are retrieved</td>
<td>Ad is requested</td>
<td>Ad is displayed</td>
</tr>
</tbody>
</table>

The user visits websites about a range of topics, for example: "Country Music", "Makeup & Cosmetics", "Vegetarian Cuisine". The browser calculates the most frequently visited topics from the user’s recent browsing history. The user visits a site whose adtech platform needs to select an ad for them. The adtech platform gets topics of interest to the user by calling the Topics API function `browsingTopics()`. The adtech platform uses the topics provided by the Topics API as part of the input to help select an ad. An ad is displayed to the user.

Topics are selected from a taxonomy consisting of hierarchical categories such as /Arts & Entertainment/Music & Audio/Soul & R&B and /Business & Industrial/Agriculture & Forestry.

The (maximum) three topics returned for a user are chosen at random from the top five for the past three epochs (with a 5% chance of getting a random topic).
Web Tracking
Browser Fingerprinting

- Websites can also fingerprint you effectively with browser fingerprinting, which is a technique that leverages all your settings to identify you, and stores this in a cookie on your browser.
  - https://amiunique.org/
- So long as JavaScript can run (by third-parties), you run the risk of being “followed” on the web.
You’ve gone incognito

Pages you view in Incognito tabs won’t stick around in your browser’s history, cookie store, or search history after you’ve closed all of your Incognito tabs. Any files you download or bookmarks you create will be kept.

However, you aren't invisible. Going Incognito doesn’t hide your browsing from your employer, your Internet service provider, or the websites you visit.

LEARN MORE
Bankruptcy

Bankruptcy helps people who can no longer pay their debts get a fresh start by liquidating assets to pay their debts or by creating a repayment plan. Bankruptcy laws also protect financially troubled businesses. This section explains the bankruptcy process and laws.

About Bankruptcy

Filing bankruptcy can help a person by discarding debt or making a plan to repay debts. A bankruptcy case normally begins when the debtor files a petition with the bankruptcy court. A petition may be filed by an individual, by spouses together, or by a corporation or other entity.

All bankruptcy cases are handled in federal courts under rules outlined in the U.S. Bankruptcy Code.

There are different types of bankruptcies, which are usually referred to by their chapter in the U.S. Bankruptcy Code.

- Individuals may file Chapter 7 or Chapter 13 bankruptcy, depending on the specifics of their situation.
- Municipalities—cities, towns, villages, taxing districts, municipal utilities, and school districts may file under Chapter 9 to reorganize.
- Businesses may file bankruptcy under Chapter 11 to liquidate or Chapter 13 to reorganize.
Privacy Enhancing Technologies

Methods for protecting personal data

Most Common/Successful? TLS.

Comes with browser. Also used for protecting email. It just works, without you having to configure anything. Protects contents of communication from passive eavesdroppers and active MITM attacks.

Tools that provide confidentiality also provide some privacy. You probably don't want your landlord or coffee shop customers to learn things about you.
Encouraging HTTPS Adoption

2014: HTTPS used as a page rank indicator

Early 2018: Mozilla announces that new features will require HTTPS

Late 2018: New Chrome HTTPS indicators
Chrome Page Loads over HTTPS

Google Transparency Report
STARTTLS as seen by Gmail

- Outbound emails: 98%
Protecting Metadata

TLS only protects content. What doesn’t TLS protect against?

We may want to protect metadata:

• Who is visiting what websites? Who is sending messages to whom?
• Gov’t might not like that you’re visiting Human Rights Watch website
• Gov’t might not be amused that you’re sending messages to Human Rights Watch
• We may want to hide the existence of the message (maybe sending an encrypted message at all is going to cause you problems)
What is Anonymity?

Anonymity (“without name”) means that a person is not identifiable within a set of subjects

**Unlinkability** of action and identity

- For example, sender and his email are no more related after adversary’s observations than they were before

- Who talks to whom

**Unobservability**

- Adversary cannot tell whether someone is using a particular system and/or protocol
Why Anonymity?

To protect privacy:
- Avoid tracking by advertising companies
- Viewing sensitive content
  - Information on medical conditions
  - Advice on bankruptcy

Protection from prosecution
- Not every country guarantees free speech

To prevent chilling-effects
- It’s easier to voice unpopular or controversial opinions if you are anonymous
Anonymity is Hard

Internet anonymity is hard...

Right there in every packet is the source and destination IP address

ISPs store communications records
   - Law enforcement can subpoena these records

Wireless traffic can be trivially intercepted

Tier 1 ASs and IXPs are compromised — NSA, GCHQ, “Five Eyes”
Anonymity

Difficult if not impossible to achieve on your own

You generally need help.

**State of the art technique:** Ask someone else to send it for you
Naive approach .... VPNs
Naive approach .... VPNs

Lulzsec fiasco
Posted on September 23, 2011

We have received concerns by users that our VPN service was utilized by a member or members of the hacktivist group ‘lulzsec’. Lulzsec have been ALLEGEDLY been responsible for a number of high profile cases such as:

- The hacking of the Sony Playstation network which compromised the names, passwords, e-mail addresses, home addresses and dates of birth of thousands of people.
- The DDOS attack which knocked the British governments SOCA (Serious Organised Crime Agency) and other government websites offline.
- The release of various sensitive and confidential information from companies such as AT&T, Viacom, Disney, EMI, NBC Universal, and AOL.
- Gaining access to NATO servers and releasing documents regarding the communication and information services (CIS) in Kosovo.
- The defacement of British newspaper websites The Sun & The Times.
Naive approach .... VPNs

“...received a court order asking for information relating to an account associated with some or all of the above cases. As stated in our terms of service and privacy policy our service is not to be used for illegal activity, and as a legitimate company we will cooperate with law enforcement if we receive a court order”
Tor is a successful privacy enhancing technology that works at the transport layer

Millions of active users.

Normally, a TCP connection reveals your IP address

Tor allows TCP connections without revealing your IP
Tor ("The Onion Router")

Tor operates by tunneling traffic through multiple “onion routers” using public key cryptography.
Who Knows What?

**Entry node:** knows Alice is using Tor, and identity of middle node, but not destination

**Exit node:** knows some Tor user is connecting to destination, but not which user

**Destination:** knows a Tor user is connecting to it via the exit node

Tor does not provide encryption between exit and destination (use HTTPS!)
Does Tor Provide Anonymity?

Tor provides for anonymity in TCP connections over the Internet, both unlinkably (long-term) and linkably (short-term).

What does this mean?

• There’s no long-term identifier for a Tor user
• If a web server gets a connection from Tor today, and another one tomorrow, it won’t be able to tell whether those are from the same person
• But two connections in quick succession from the same Tor node are more likely to in fact be from the same person
Performance: message bounces around a lot (can be slow)

Attack: government can coerce server operates in one country
  Defense: use mix servers in different legal jurisdictions

Attack: adversary operates all of the mixes
  Defense: have lots of mix servers (Tor has ~7,000 onion routers today). Use diverse set.

Attack: adversary observes when Alice sends and when Bob receives, links the two together

A side channel attack – exploits timing information
  Defenses: pad messages, introduce significant delays
  Tor does the former, but notes that it’s not enough for defense
How do you protect against an adversary creating a large number of onion routers and performing timing observation at entrance and exits?

Limit the servers used for initial connection to a subset of trusted nodes:
- Have long and consistent uptimes…
- Have high bandwidth…
- Are manually vetted by the Tor community

Tor client selects 3 guard relays and uses them for 3 months
Relays must self-elect to be exit nodes. Why?

- Legal problems

- If someone does something malicious or illegal using Tor and the police trace the traffic, the trace leads to the exit node.
As described, Tor protects the identity of the client, but not the server.

What if we want to run an anonymous service?
- a website, where nobody knows the IP address?

Tor supports Hidden Services…
- Allows you to run a server without disclosing the IP or DNS name

Many hidden services
- Duck Duck Go, Tor Chat, Wikileaks
**Step 1:** Bob picks some introduction points and builds circuits to them.
Step 2: Bob advertises his hidden service -- XYZ.onion -- at the database.
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.
Steps in Tor Hidden Services:

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.
Step 6: Bob and Alice proceed to use their Tor circuits like normal.
This hidden site has been seized
by the Federal Bureau of Investigation,
in conjunction with the IRS Criminal Investigation Division,
ICE Homeland Security Investigations, and the Drug Enforcement Administration,
in accordance with a seizure warrant obtained by the
United States Attorney’s Office for the Southern District of New York
and issued pursuant to 18 U.S.C. § 983(j) by the
United States District Court for the Southern District of New York
Who uses anonymity systems?

“If you’re not doing anything wrong, you shouldn’t have anything to hide.”

- Implies that anonymous communication is for criminals

The truth: who uses Tor?

- Journalists, Law Enforcement, Human Rights Activists, Business Executives, Intelligence/Military, Normal People
Internet Censorship

Government censors
   Block websites containing “offensive” content
   Commonly employ blacklist approach

Observed techniques
   IP blocking, DNS blackholes, forged RST packets

Popular countermeasures
   Mostly proxy based — Tor, Freenet, Ultrasurf, …
   Problem: Cat-and-mouse game
Internet Censorship

Pervasive censorship
Substantial censorship
Selective censorship
Changing situation
Little or no censorship
Anyone can look up the IP addresses of Tor relays
  - Public information in the consensus file

Many countries block traffic to these IPs
  - Essentially a denial-of-service against Tor

Solution: Tor Bridges
  - Tor proxies that are not publicly known

Tor Bridges
Here are your bridge relays:

bridge 87.172.228.191:443
bridge 129.59.73.69:9001
bridge 128.197.31.73:443

Bridge relays (or "bridges" for short) are Tor relays that aren't listed in the main directory. Since there is no complete public list of them, even if your ISP is filtering connections to all the known Tor relays, they probably won't be able to block all the bridges.

To use the above lines, go to Vidalia's Network settings page, and click "My ISP blocks connections to the Tor network". Then add each bridge address one at a time.
Obfuscating Tor Traffic

Bridges alone may be insufficient to get around all types of censorship
- DPI can be used to locate and drop Tor frames

Countries would passively detect and block bridges
- Single use bridges

Tor adopts a pluggable transport design
- Tor traffic is forwarded to an obfuscation program
- Obfuscator transforms the Tor traffic to look like some other protocol
  - BitTorrent, Skype, HTTP, streaming audio, etc.
Decoy Routing (Telex)

- **User’s Computer**: Request for permitted site, Request for prohibited site
- **ISP Infrastructure**: Router
- **Websites**:
  - NotBlocked.com: unobjectionable site (permitted by censor)
  - Blocked.com: prohibited site (blocked by censor)

Request for **permitted** site  Request for **prohibited** site
Decoy Routing (Telex)

User's Computer | ISP Infrastructure | Websites
---|---|---
Browser | Telex Client | NotBlocked.com unobjectionable site (permitted by censor)
| | | Blocked.com prohibited site (blocked by censor)

Censor

Request for permitted site | Request for prohibited site
Decoy Routing (Telex)

User's Computer \[\text{Browser} \quad \text{Telex Client}\] \quad \text{ISP Infrastructure} \quad \text{Website}

- **NotBlocked.com** unobjectionable site (permitted by censor)
- **Blocked.com** prohibited site (blocked by censor)

Request for **permitted** site → Request for **prohibited** site
Decoy Routing (Telex)

User’s Computer

- Browser
- Telex Client

ISP Infrastructure

- Router
- Invisible Tag
- Telex Station

Websites

- NotBlocked.com (unobjectionable site, permitted by censor)
- Blocked.com (prohibited site, blocked by censor)

Requests:
- Green arrow: Request for permitted site
- Red arrow: Request for prohibited site
- Green text: Request for **permitted** site
- Red text: Request for **prohibited** site
Decoy Routing (Telex)

- **User's Computer**
  - Browser
  - Telex Client

- **ISP Infrastructure**
  - Router
  - Invisible Tag

- **Websites**
  - NotBlocked.com: unobjectionable site (permitted by censor)
  - Blocked.com: prohibited site (blocked by censor)

- **Requests**
  - Green arrow: Request for permitted site
  - Red arrow: Request for prohibited site

- **Telex Station**
  - Key symbol
Email Protection

Your email provider may be required to turn over your (securely stored) email

- Warrant (for content)

Metadata

- National Security Letter (NSL), Court Order

What if you want to protect email content?
Modern implementations: GnuPG, Keybase

Each user has:
- A public encryption key, paired with a private decryption key
- A private signature key, paired with a public verification key

How does sending/receiving work?
How do you find out someone's public key?
PGP Operations

To send a message:
- Sign with your signature key
- Encrypt message and signature with recipient's public encryption key

To receive a message:
- Decrypt with your private key to get message and signature
- Use sender's public verification key to check sig
PGP Public Keys

How do you obtain Bob's public key?
Get it from Bob's website? (😖)
Get it from Bob's website, verify using out-of-band communication
  Keys are unwieldy fingerprints
  A fingerprint is a cryptographic hash of a key

What if you don't personally know Bob?
  Web of Trust (WoT)
  Social Network (Keybase)
Lost PGP Key

What if Bob's machine compromised?

His key material becomes known

Past messages can be decrypted and read

You also have sender's signature on messages sent, so you can prove identity of sender

Sender must trust recipient's ability/desire to keep her statements private
1. Use authenticated Diffie-Hellman to establish a (short-lived) session key $E_K$

$$SS = (g^y)^x$$
$$E_K = H(SS)$$

A diagram shows the process:
- Alice sends $\text{Sign}_{\text{alice}}(g^x)$ to Bob.
- Bob sends $\text{Sign}_{\text{bob}}(g^y)$ to Alice.

Bob's output:
$$SS = (g^x)^y$$
$$E_K = H(SS)$$
2. Then use symmetric encryption on message M ... and authenticate using a MAC

Alice

\[ SS = (g^y)^x \]
\[ EK = H(SS) \]
\[ MK = H(EK) \]

Bob

\[ SS = (g^x)^y \]
\[ EK = H(SS) \]
\[ MK = H(EK) \]
3. Re-key using Diffie-Hellman

\[ SS' = (g^{y'})^{x'} \]
\[ EK' = H(SS') \]
\[ MK' = H(EK') \]
\[ MK = H(EK) \]
OTR: Off the Record Chat

4. Publish old MK

\[ SS' = (g^{y'})^{x'} \]
\[ EK' = H(SS') \]
\[ MK' = H(EK') \]
\[ MK = H(EK) \]

Alice

Bob

“Deniability”
Note this is suited to interactive communication, not so much email.

But, OTR provides
- message confidentiality
- authentication
- perfect forward secrecy
- deniability

OTR has since lost popularity. Signal Protocol now de facto standard.