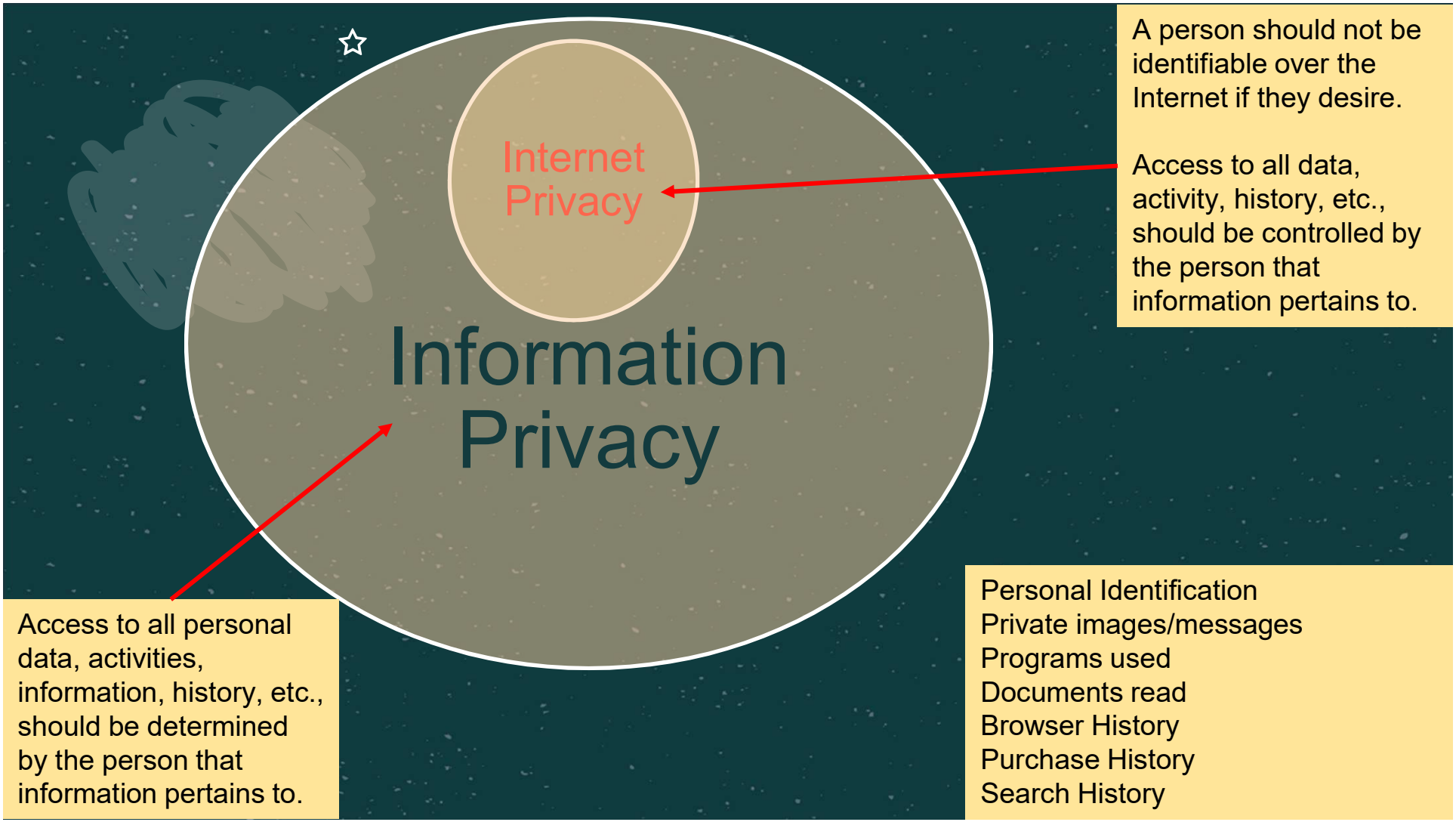

Internet Privacy
Information Security – Lecture 27
Aadil Zia Khan


Internet Privacy
Information Privacy
A person should not be identifiable over the Internet if they desire.
Access to all data, activity, history, etc., should be controlled by the person that information pertains to.
Access to all personal data, activities, information, history, etc., should be determined by the person that information pertains to.
Personal Identification
Private images/messages
Programs used
Documents read
Browser History
Purchase History
Search History

Privacy Through Simple Access Control

- Facebook allows its users to specify who can see their friend, messages, photos, etc.
- Twitter allows you to control who can reply to your tweets
- Many Social media sites allow blocking

actual problem is much bigger than that - achieving privacy is not so simple

Let's Start With Private Account Information

- Visit https://haveibeenpwned.com/ and enter your email address to see if your information was stolen
- Results for the email I tested
  - Cit0day: In November 2020, a collection of more than 23,000 allegedly breached websites were made available for download on several hacking forums - data consisted of 226M unique email address alongside password pairs
  - Collection #1 : In January 2019, 773 million unique email addresses alongside passwords were distributed on a hacking forum
  - Covve: In February 2020, Covve contacts app data, including Email addresses, Job titles, Names, Phone numbers, Physical addresses, Social media profiles got leaked – the tested email never had the permission to be used by this app
  - People Data Labs: In October 2019, 622 million unique Exposed information included email addresses, phone numbers, social media profiles and job history data got leaked – the tested email never had the permission to be used by this app
  - LinkedIn: In May 2016, LinkedIn had 164 million email addresses and passwords exposed and offered for sale on a dark market site 4 years later
  - Exploit.In: In late 2016, 593 million email address and password pairs got leaked

Lesson Learnt

Change passwords frequently – very frequently

Use 2-Factor Authentication

Gmail: Put a plus sign (+) after your name but before the @-sign and then type anything you like there, and the email will still reach you – create different combinations for each site so you can identify which website got compromised

Let's Look At Privacy And Browsing

What Happens When You Access A Webpage

- Browser sends a DNS query to get the IP address of the server hosting the page
  - DNS packets are not encrypted – anyone who can see the packet can also see the website you want to access

What Happens When You Access A Webpage

- Browser then sends an HTTP or HTTPS request for that site
  - HTTP requests/responses are not encrypted - anyone who can see the packet can also see the website you want to access and the contents on the site
  - HTTPS requests/responses are encrypted but the requested domain name is not - anyone who can see the packet can also see the website you want to access

What Happens When You Access A Webpage

- Each HTTP/HTTPS request exposes some information
  - IP address of the requester if the user has bought an IP address, else the IP address of the ISP's/organization's NAT behind which the user's machine exists
  - User Agent – browser name and version and plugins, operating system
  - Time zone
  - Country
  - Screen settings – very large setting could imply you have impaired vision
  - Language
  - CPU, RAM, and other hardware details

What Happens When You Access A Webpage

- Each HTTP/HTTPS request may also include cookies
  - Cookies are used to uniquely identify website visitors
  - Help to remember things like your account login info, or what items were in your online shopping cart, etc.
  - They can be misused to link all your visits, searches, and other activities on a site together
  - This usage is a privacy violation, and browsers allow you to block, limit, or delete cookies

What Happens When You Access A Webpage

- Request headers, cookies, and all additional information sent with the request together form the digital fingerprint of the user
- A digital fingerprint is essentially a list of characteristics that are unique to a single user, their browser, and their particular hardware setup
- Tracking sites can stitch all the small pieces together to form a unique fingerprint of user and his device

What Happens When You Access A Webpage

- Advertising code and invisible trackers on the requested site might also cause your browser to make hundreds of requests to other hidden third parties
- Each request could contain your digital fingerprint
    - Also finds out the websites you are visiting

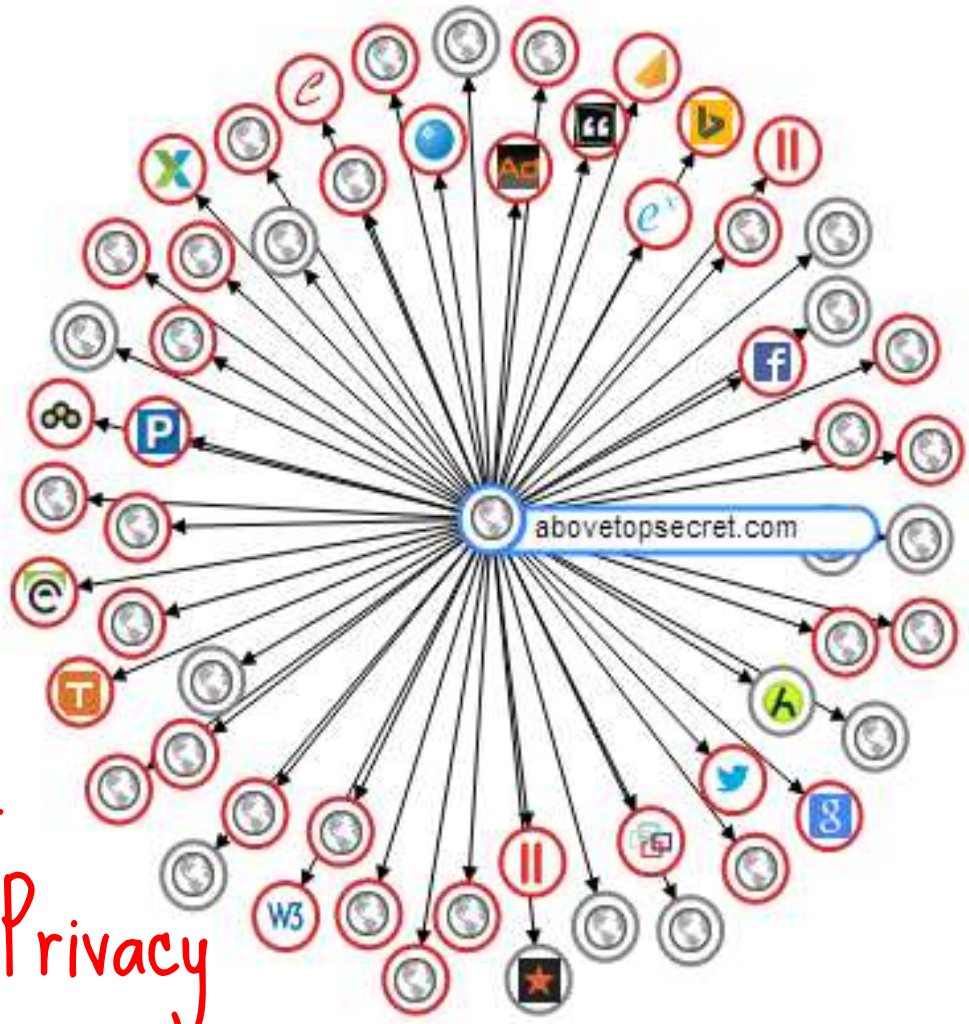
- This is “Cross Site Tracking” - companies collecting browsing data across multiple websites
    - This is much worse than a single website tracking its visitors’ activity on its own pages

abovetopsecret.com

When you visit this site, the following sites are informed:

- google-analytics.com
- fonts.googleapis.com
- gstatic.com
- ajax.googleapis.com
- gorillanation.com
- rubiconproject.com
- doubleclick.net
- 4dsply.com
- scorecardresearch.com
- realmedia.com
- adblade.com
- facebook.com
- exponential.com
- craveonline.com
- tribalfusion.com
- adnxs.com
- googlesyndication.com
- pubmatic.com
- casalemedia.com
- optmd.com
- fastclick.net
- dotomi.com

Trackers - Threat To Privacy



How Private Is Our Browser - Let's Test It

- Visit https://coveryourtracks.eff.org/ and start the test
  - It will display all the privacy leaks

How Safe Is The Website - Let's Test It

- Visit https://themarkup.org/blacklight specify the website you want to test and start
  - It will display all the user-tracking technologies
- I tested www.pakwheels.com and www.umt.edu.pk
  - UMT had no tracker ☺
  - Pakwheels had
    - 7 Ad trackers belonging to the companies OneSignal, Alphabet, and Facebook - sending data to companies involved in online advertising
    - 2 cookies set for New Relic and Alphabet
    - Canvas fingerprinting - this technique identifies users even if they block third-party cookies by drawing a unique image on the canvas
    - Session recorder (script belonging to the company Hotjar) - tracks user mouse movement, clicks, taps, scrolls, or even network activity
    - Keystrokes were not captured (good news)

Let's Play A Prank

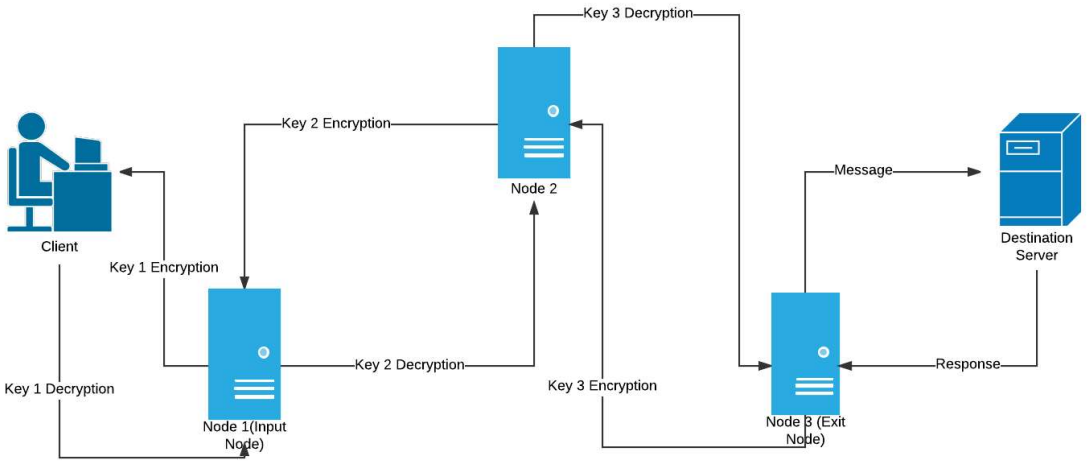
- Borrow your friends phone or computer
- Search for something very weird
- Browse some websites on that

your friend will now keep seeing ads related to that ☺

Achieving Browsing Privacy

- Use a Tracker Blocker or Ad Blocker – they prevent tracking companies from reading your browser fingerprint
- Most trackers run on JavaScript so disable that
    - Disabling JavaScript limits the functionality of many websites
- Using a fingerprint resistant browser – ToR browser

Onion Routing Using Tor Browser



- The client encrypts the HTTP request using keys 1, 2 & 3 and sends it to the Node 1 (Input Node)
- Node 1 decrypts the message using Key 1 and realizes that it doesn't make sense so sends it to Node 2
- Node 2 decrypts the message using Key 2, realizes that its still encrypted and sends it to the exit node
- Node 3 (exit node) decrypts the message using Key 3, finds a GET request for youtube.com and passes it onto the destination youtube server
- The server processes the request and serves up the desired webpage as a response - response passes through the same nodes in the reverse direction where each node encrypts using their specific key
- It finally reaches the client in the form of a triple encrypted response which is then decrypted

Can I Throw Caution To The Wind When Searching?

Are Searches Private?

- Search engines track a user's searches through user's signature, account, or IP address
  - Geolocation is essential for relevant results
- Search engines retain information in order to provide better services

Try searching something weird – soon all advertisements (on your computer and maybe other computers on your LAN) would be related to that search

How To Preserve Search Privacy

- Use search engines/browsers focused on privacy
- Search engines such as Startpage.com, and Disconnect.me anonymize Google searches
- DuckDuckGo combines the search results from various search engines
- Use Tor Browser

Should I Open Emails?

- Someone send you an email – you read it but don't reply – next day he tells you that he knows you read the email
  - How did this privacy breach occur???

Web Beacon / Tracking Pixel

- Web Beacon / Tracking Pixel
  - Initially web beacons were small digital image files that were embedded in a web page or email
    - Now we can have scripts
  - The image could be a single transparent pixel of the same color as the background
  - When a user opens the page or email where such an image is embedded, web browser or email reader automatically downloads the image by requiring the user's computer to send a request to the server where the source image is stored
  - This request provides identifying information about the computer, allowing the host to keep track of the user
    - Sender of an email - or even a third party - can record information like the time that the email was read, the IP address of the computer that was used to read the email, the type of software used to read the email, and the existence of any cookies previously sent

How To Protect Against Tracking Pixels / Web Beacons

- Configuring the email reader software to avoid accessing remote images
- Disconnect from the Internet after downloading email but before reading the downloaded messages
- Use a text-based email reader (such as Pine or Mutt) - plain-text email messages cannot contain web beacons because their contents are interpreted as text instead of embedded HTML code, so opening such messages does not initiate any communication
- Some email readers offer the option to disable all HTML in every message (thus rendering all messages as plain text)
- Remember sometimes gmail doesn't load images and user must explicitly choose to load images - many email readers and web-based email services do not load images when opening a hypertext email that comes from an unknown sender to prevent tracking pixels

slidesgo