Some Common Attacks Against Anonymous Web Networks

1. Traffic Analysis
2. Exit Node Exposure
3. Government Attacks
4. The Onions Strike Back

History of internet privacy and libraries

1. Patriot Act : enabled investigators to gather information when looking into the full range of terrorism-related crimes
2. Health Insurance Portability and Accountability Act (HIPAA) : creates national standards to protect individuals by removal of specific data identifiers (names, telephone numbers, email addresses, IP addresses, etc.)
3. Breeding's Library Technology Reports : privacy concerns related to web-based library information systems, including library circulation systems, digital certificates, data storage, web hosting, and trackers.

A Look at Privacy and Censorship Law

- Green Flag
    1. United Kingdom : Data Protection Act of 2018 - "guarantees" protection of sensitive personal data such as race, political and religious beliefs, health and biometrics, and sexual orientation and all users are informed and consent to data collection, that they are aware and have access to the data that is collected, and able to ask that data be erased/cease collection.
       Tor is freely available for public use
    2. France : The Right to be Forgotten- If an individual wants personal information, including photos and videos of them, removed from a website, the site is compelled to oblige
       Digital Last Will and Testament - Individuals are also granted the right to decide what will happen with their online data when they die
       Assurance of Net Neutrality - Net neutrality assures that Internet Service Providers provide equal access(no sped up/slow down/block sites) to all online information.
       one of the greatest rates of Tor use
    3. Germany : Bundesdatenschutzgesetz
       The collection, processing, and use of data are strictly prohibited, unless it is permitted by the law or the person
       Personal data must be collected from the person, not through third parties.

Any data standards must reflect a balance between the collecting agency and the user/individual.
Data anonymization should be used to mitigate the transfer of personal data across systems.
Any entity that collects/stores personal information must inform all affected persons that they are doing so.
Permission granted by an individual to sue data for a specific purpose is limited to that purpose
one of the most data secure countries in the world

4. Canada : National/provincial laws - limited the collection and dissemination of individuals' personal information by government agencies only
   Use of TOR is not very strict
5. Australia : data privacy laws - guarantee privacy , requires ISPs to collect and store data some of the most restrictive censorship laws of any developed country, which allows or forces ISPs to filter a variety of sites, including obvious ones (child pornography, sexual violence) as well as ones that may be more of a gray area (torrent sites

- Yellow flag
  1. United States : The Right to Privacy - That the individual shall have full protection in person and in property
     Ironically, the single biggest funder of the Tor Project also spends significant time and funding in attempt in finding ways to deanonymize it
  2. Nigeria and West Africa : Nigerian Cybercrimes Prohibition Act of 2015 - provide guidance on the policing of criminal activity
  3. East Africa : Access to Information and Protection of Privacy Act (misnomer) - have severely hampered access to information through censorship tactics and spying on the communications of citizens
     40%–45% of adults use the Internet
  4. South Africa : Protection of Personal Information Act of 2013 - expands the powers of the government to preserve the privacy of citizens and organizations
     much better off in terms of privacy than most countries on the African continent.
  5. Brazil : surveillance laws - allow for telecommunications to be intercepted for use in criminal investigations
     one of the largest pools of Tor users
  6. Mexico : Constitution - requires law enforcement to receive a warrant before interfering or intercepting the communications of private citizens
     Telmex, run by Carlos Slim, the fifth-wealthiest person in the world (as of late 2020) – has a virtual monopoly on all telecommunications in the country

At times restricted access to Tor

7. Spain : guarantees a right to privacy for citizens , very similar to other European countries and allows the government broad discretion in blocking access to the web as it deems necessary to preserve “public health and safety”
8. Italy : directed to store user data for years as part of national security efforts
   one of the greatest rates of Tor use

- Red Flag
  1. China : Data shared in a public forum is not held to the same standard and may be collected without consent
     Organizations/services that collect user data must provide the following information to all users: • How frequently will data be collected, and where will it be stored. • The types of data will be secured and potential risks. • Contact information for the data holder.Data may only be transferred or shared if the holder has consent from the user or the data has been entirely de-identified.
     cybersecurity law - critical information infrastructure operators” must be stored domestically (in China) - limits collection of Chinese citizens’ data by foreign organizations
     The Great Firewall of China - restrict access to a wide number of websites and services – including Tor.
     Mirror sites - different URL - Bridges are used to connect users to the Tor network when it is censored in their country
  2. Iran : does not have any laws that specifically protect the privacy of its citizens.
     most restrictive of all countries in terms of Internet privacy and the use of Tor
     one-fourth of surface websites are blocked (top 500 most-visited sites)
  3. Russia : Sensitive personal information – including racial demographics, political opinions, religious beliefs, health conditions, and sexual life – should generally be collected under no circumstances
     complicated relationship with Tor.

Countries with the Most Tor Users : Russia, Germany, France, United States, United Kingdom, Israel, Canada, Hong Kong, Denmark, Australia, Nigeria, South Africa, China, Brazil, Mexico, Uganda, Egypt

Countries with the Most Tor Bridge Users

1. China : fewest number of daily users
2. Iran : access content not approved by government as “culturally appropriate”
3. Russia : use the platform for broad political gain and ban it for use by its own population., controversy, invasion, and potentially illegal political activities on the web

4. United States : love/hate relationship with Tor

TOR Relays

1. guard or non-exiting relay is a fast relay on the network that does not provide any kind of exit off the network, just a way to move between different relays on the network.
2. If a guard relay is not fast enough, it will be demoted to a simple middle relay. Guard and middle relays are the simplest type of relay to run since they do not require a lot of setup or configuration
3. Exit relays are the final step in the relay chain for the Tor network. They accept packets from the middle or guard relays and direct it to the surface web for delivery
4. bridge relay is a specialized relay that can be used by individuals or organizations who need an extra layer of protection when connecting to the Tor network.
5. Snowflake relays do not require a relay to be operational 24/7 and works as extension on a Firefox or Chrome browser that operates when your computer is in use and allows you to continue to use the Internet normally while allowing individuals in high-censorship areas to connect securely to the Tor network

Library and information science (LIS) : Anonymous Web Research Topics

1. Natural Language Processing : examine content uploaded on the anonymous web, extends to nonlanguage content, like photos and images
2. Authorship Attribution : used in the digital humanities to help determine authorship of books where no author is known, content shared is something illegal, and was posted anonymously, it may be possible to identify the author through linguistic analysis
3. Extending and Incorporating the Tor Network in New Platforms : several emerging networks that employ the Tor protocol (using onion routing) as part of a new service that promises even more privacy, security, or user-friendly interface

Anonymous Web and COVID-19 : . During the peak of the pandemic in nations like India, these drugs became highly valued and extremely scarce, leading to a black market, The impact of an emergency situation on the use of the anonymous web should serve as a learning opportunity for future experiences.

Methods for Anonymous Web Research

1. Analysis of Traffic Data : illustrate relationships between major events and usage of the anonymous web.
2. Content Analysis : to analyze language use and communication, hoping to reveal some deeper meaning.
3. Text Mining : can reveal insight about the people that use these platforms
4. Ethnography : immerse themselves in a culture, in order to better understand the motivations behind beliefs and practices

Potential Hazards and Other Considerations

1. Most people using the platform could be considered to be a part of some sensitive population, and that they are using the platform specifically to preserve their anonymity, so extra care (even beyond normal) must be taken to ensure the sanctity of the study and its data
2. Allowing your priors/ beliefs about the anonymous to influence your analysis and cause you to draw false conclusions. It is evident from the titles of research articles pertaining to the anonymous web that many are drawn to the criminal elements on the platform
3. Bibliometric studies analyze the bibliographic data for a set of publications on a particular topic. The objective of this analysis is to identify important authors, publication venues, and topic themes, with this information being used to inform future research by suggesting new topics and best places to publish research

Tor :

1. Coinbase and Binance : Coinbase and Binance – the two most popular cryptocurrency exchanges
2. SciHub and Library Genesis :These platforms are very helpful to researcher, also very illegal
3. Facebook

Dictionary

Principles of Information Rights

Anonymity State in which one's identity is concealed or unknown.
Censorship Suppression of one's right to freely communicate or exchange information with others.
Privacy State in which someone or something is free from the observation of others. Privacy is distinct from anonymity in that, with anonymity, the product of a person's activities (e.g., a book) is known to others, but the identity of the person (e.g., the author of the book) is known only to that person or a small group of others. Conversely, if a book were kept private, then nobody would have access to it other than the author.
Security Freedom from risk or danger; preservation of the integrity of one's data or information.

Anatomy of the Computer

Binary Something that can hold two states; for instance, "off" and "on," "yes" and "no," or "0" and "1."
Bit The smallest unit of information that can be communicated – a binary digit.
Boolean Logic The branch of mathematical logic where all variables have binary options (commonly, "true" and "false").

Byte Equal (generally) to eight bits, this is the smallest unit of information that can be used to encode computer characters (e.g., letters, numbers, and symbols).

Compression A reduction in the number of bits needed to encode a message. Compression can be achieved through statistical means of reducing redundancy or removing unnecessary information.

Computer Something that performs computations; historically, this was humans, in the last century it has become a role for digital technology.

Digital Communicates using digits, such as binary data.

Packets (data) Data packaged with a set of instructions that control how the data is communicated and received. Fundamental to modern data communication on networks (such as the Internet).

Random Access Memory (RAM) A computer's short-term memory. Allows the computer to operate and retain content while the user is interacting with it.

Solid-State Drive (SSD) A computer's long-term memory. Stores data even when a computer is not in operation.

Networking

ARPANET Precursor to the modern Internet, a network designed to facilitate communication among government and research facilities.

Cookie A piece of data that "attaches" itself to a user as they navigate the web and which allows sites, advertisers, and other third-party entities to track users' behavior.

Domain Name System (DNS) A system devised to provide domain names that identify websites, associating the names that users' input with the actual IP addresses that connect sites.

Encryption Encoding of information into a representative ciphertext that can then be securely communicated across a network.

Exit Relay The final relay in a Tor path that connects to the destination user/server.

Garlic Routing Variant of onion routing, used by the I2P network, that packages encrypted messages together before sending them in order to prevent common attacks used against other networks.

Information Theory The study of how information is stored and exchanged/communicated.

Internet-of-Things A network of physical "things" that communicate with one another using the Internet in order to better perform some role. For instance, a futuristic example of Internet of Things may include a network that can detect an increase in temperature, pull your curtains, and start your air conditioner and ceiling fan.

Internet Protocol (IP) Address An identifier for every device that connects to the Internet or associated network. Analogous to how a physical home address identifies the location of a domicile (hence the name).

Internet Service Provider (ISP) Intermediary service that facilitates the communication of data across the Internet.

Network An interconnected collection of devices that can communicate among one another.

Network Routing Protocol Set of data that defines rules for how other data is communicated. IP addresses are a component of the routing protocol.

Nodes Also known as relays. Points along anonymity networks, like Tor, where information is encrypted and/or routed.

Onion Routing Method of data communication where data is encrypted and relayed through a series of nodes in order to facilitate anonymous communication.

Routing Method through which data is communicated among points on a network.

Traffic Analysis Method of collecting information about data shared on networks by observing traffic patterns (how much data was sent, what frequency, etc.).

World Wide Web (WWW) System of linked resources, stored on servers that are accessed through a public network; it is the content and organizational system that makes the Internet worthwhile for users.

Information Systems and Retrieval

Filter Bubble A state in which search results have become so personalized based on past searching behavior that users are presented only with information that does not challenge their preconceived perceptions and biases.

Google Analytics Analytics tracking service, operated by Google, that tracks website traffic.

Information Access The ability to find and obtain information necessary to satisfy their needs.

Information Retrieval The process by which information is found and obtained.

System Analysis and Design The study and process of analyzing the functioning of a system and the needs of its users, and designing solutions to satisfy the identified needs and gaps.

Code and Coding

Compiler Computer program that translates code inputted in one language (e.g., a human-readable language like JAVA) into another language understood by the computer (machine-readable language).

Java Programming language developed in the mid-1990s that has gained popularity for its ease of use and broad potentials; generally used to create interactive web applications.

Machine-Readable Language Content coded in a programming language that can be interpreted and acted upon by a computer.

Python Programming language popular due to its simplicity and broad use potentials; also commonly used for data mining and statistical analysis.

SQL Programming language used to manage data within system databases.

Levels of the Web

Anonymous Web A term that is (pragmatically) synonymous with the dark web but avoids the negative connotations that have been skewed by media portrayals.

Dark Web The "deepest" part of the web, accessible only via specialized software.

Darknet Another term used to refer to the dark web with a decidedly negative connotation; often used when referring to illegal activities on dark web platforms.

Deep Web The layer of the web that is hidden behind some sort of authentication screen (e.g., a log-in page).

Surface Web The top level of the web; content that can be accessed by simply navigating to a site without requiring any log-in/authentication of the user.

Types of Anonymous Web Platforms and Alternatives

Brave Web browser designed to optimize user privacy and block web ads while compensating content creators through the use of the Basic Attention Token (BAT), a native cryptotoken.

DuckDuckGo A privacy-enhancing search engine that does not track and store users data and thus also mitigates issues like filter bubbles.

Firefox Popular browser that offers elevated privacy relative to Internet Explorer and Google Chrome, but less so that anonymous web platforms.

Freenet An anonymous web platform proposed by Ian Clarke in the late 1990s and implemented in the early 2000s that supports censorship-resistant peer-to-peer communication.

I2P An anonymous web platform, released by the Invisible Internet Project (I2P) in the early 2000s, that is intended to facilitate private peer-to-peer communications and host eepsites: Small, used-owned sites that operate similar to blogs.

Tor The Onion Router, the original anonymous web platform developed in the mid-1990s, was designed to facilitate anonymous communication around the world, but has evolved to allow users to access both surface website and .onion sites, unique to the Tor platform.

Bridges Bridges are relays within the Tor network that are not included in a public directory of Tor relays. These relays are used to subvert efforts to shut down or block the Tor network. These bridges are used by many of the sites that exist on the network and can also be used by individual users on the network, though it does slow loading times considerably.

Eepsites I2P's alternative to the surface websites; function like small blogs or MySpace pages where uses can upload and share content.

HTTP/HTTPS Hypertext Transfer Protocol (HTTP) and Hypertext Transfer Protocol Secure (HTTPS) are the foundation of data communication over the Internet, serving like the "to:" statement on an email or letter. Using HTTPS, the communications sent are encrypted with Transport Layer Security that protect the communication between sender and receiver while prevent eavesdropping by third-parties.

Onion Sites Websites that are unique to the Tor network and can only be access via the Tor browser. These sites host the "hidden services" that are discussed in relation to criminal activities on the dark web.

The Cryptoverse

Binance Binance is one of the most popular platforms for the exchange of cryptocurrencies, with one of the largest selections of coins. The platform has received intense scrutiny from the United States, with the platform being originally headquartered in China, and then the Cayman Islands, due to a lack of regulatory compliance.

Bitcoin The first, full-fledged cryptocurrency released for public use. Bitcoin was developed by an anonymous person/group in order to provide a decentralized, private and secure alternative to fiat currencies (those traditionally used for exchange, like the U.S. dollar or British pound. Exchange of Bitcoin is managed through a ledger system, where users compete to solve a complex puzzle in order to validate transactions, in exchange for which they receive a share of Bitcoin. This process of validation requires large amounts of energy (as the puzzles are complex and thousands of computers are competing to be the first to complete them), which is one of the major concerns with Bitcoin and similar "proof-of-work" coins.

Blockchain A collection of cryptographic blocks, or records, that are held on a ledger and allow for information to securely be exchanged or managed without compromising the anonymity of the owner of that information. To avoid ownership of information being exchanged more than once before the ledger can be updated, transactions are validated by chaining blocks together and confirming only one transaction on the ledger. For instance, if Person A has one apple and promises to sell it to both Person B and Person C, but Person D, who is in charge of a

ledger of exchanges, recognizes that the promise was made to Person B first, then Person B pays and receives for the apple and Person C neither receives nor pays for anything. This protects Person C from being sold something that does not actually exist (since the apple has already been promised to Person B).

Coinbase The most popular crypto exchange platform in the United States – essentially the Etrade of the cryptoverse. Offers over 50 popular crypto coins for trade and allows users to earn cryptocurrency in exchange for watching educational videos or "staking" their crypto: essentially, investing it back into the currency developers to help strength the currency and grow its popularity.

Cryptocurrency A digital asset managed on a blockchain ledger. Owners of a cryptocurrency actually own "keys," which are essentially passwords that validate their ownership of a certain amount of the asset on the ledger.

Dogecoin A meme-, or joke, coin developed in 2013. Dogecoin was designed as a joke and was never intended to actually have any value. However, traders, ostensibly following the praise of the coin by Elon Musk, have purchased the coin in such quantities that, in early 2021, the value of the coin peaked over seventy cents for a short period. This was a significant increase over the value of about one-half of one cent in December 2020. As of summer 2021, the price is back down to about 20 cents.

Ethereum The second-most popular cryptocurrency, after Bitcoin. Ethereum was initially released in 2015 using the same "proof-of-work" scheme used by Bitcoin; however, as of 2021, it is in the process of upgrading to "Ethereum 2.0" by transitioning to a "proof-of-stake" scheme, which validates transactions based on the consensus of those holding the large amounts of the currency. This scheme will significantly reduce the energy consumed by the network and shorten transaction speeds.

Cybercrime

Cybercriminal One who commits a crime involving the use of a computer network. Cybercrimes, like the production of ransomware and hacking of important accounts and databases, is believed to be a multi-trillion dollar "industry."

Doxing The acquisition and dissemination of previously private information about individuals or organizations. This information may be acquired through social engineering methods like phishing and shared in an attempt to shame or discredit the target's reputation.

Hacking In the literal sense, one who uses their computer acumen to overcome some obstacle in a non-traditional way. Hacking can be a completely legal act. However, hacking is often associated with criminal activities, such as hacking into a company's database

through the "non-traditional" way of stealing log-in credentials of an employee.

Laundering Investing or employing resources that have been acquired through illegal means into a legal avenue in order to "clean" it. For instance, someone who tried to deposit $1 million cash into their bank account may be met with intense scrutiny of the origins of these funds, but they would not receive the same scrutiny if they received this money in Bitcoin (especially given the anonymity of the currency) and then sold it off for $1 million over a period of several months.

Phishing A type of social engineering – the manipulation of people into revealing private information – where targets are lured into divulging private information as the product of some scam warning or offered service to them. For instance, one may receive an email that an account of theirs has been deactivated and they must log-in immediately – using the link provided in the email – in order to recover their account. The link may lead the target to a spoofed page that looks similar to the actual website and will collect the target information when entered so that it can then be used by the phisher.

Silk Road A defunct marketplace on the Tor network that is notorious for offering a variety of illegal or questionable items that could be purchased with cryptocurrency. Shut down by FBI in 2013 and again in 2015.

Torrents Peer-to-peer file sharing platforms that are often used to send copyrighted or other illegal content. Most commonly associated with the program BitTorrent.

Policy and Law

Acceptable Use Policy A policy that outlines the expected, appropriate behavior of users of a computer network. Like a "no running" sign by a pool, but much longer and full of legal jargon.

Children's Internet Protection Act (CIPA) An act of the United States' Congress designed to protect vulnerable populations like children from accessing obscene content. This act mandates filtering of public Internet for any entity (e.g., schools and libraries) that receive discount Internet services through the government's E-rate program.

Health Insurance Portability and Assurance Act (HIPAA) An act of the United States' Congress that prescribes the appropriate privacy and security measures that should be employed when collecting, storing, and sharing the medical information of patients.

Net Neutrality The position that the Internet should be a neutral platform, agnostic toward what types of communications and activities are occurring on the network. This is opposed to the view that Internet Service Providers should have the capacity to prioritize certain websites or services. In this latter case, it would be feasible, for example, for an ISP

to make an agreement with a retailer like Walmart to speed up access to its website while slowing access to Target, Amazon, etc.

Stop Online Piracy Act (SOPA) A failed act of the United States' Congress that was proposed in 2011 as a measure to curb online copyright infringement and trafficking of counterfeit or illegal items and services. The act would have expanded the power of law enforcement to police content uploaded to the Internet and held content aggregators (e.g., Google and YouTube) responsible for creating measures to combat all instances of piracy.

USA PATRIOT Act An act of the United States' Congress that broadened the power of law enforcement to surveil the activities of the public, particularly in their use of communications technologies like phones and the Internet.

Literacy

Data Literacy The ability to create, comprehend, and communicate data and insights that can be gleaned from that data (e.g., through statistical analysis of the data).

Metaliteracy The ability to comprehend and communicate one's own experiences and acquisition of competencies throughout the literacy learning process. It is not only about learning how to be information literate, but understanding what information literacy means and how it can be used to collaborate and improves the lives of others.

Privacy Literacy The ability to comprehend and communicate principles of online privacy and employ these principles to improve one's own privacy.

Security Literacy The ability to comprehend and communicate principles of online security and employ these principles to improve one's own security.

Other Privacy Tools

HTTPS Everywhere Available as a plug-in for many popular web browsers (and included as standard in the Tor network), HTTPS Everywhere requests that each site a user visits use Hypertext Transfer Protocol Secure, if possible, in order to further enhance the privacy of the connection with the user.

Password Manager Available as a software or plug-in for web browsers, a password manager securely stores passwords for the user, which enables the user to use more complex passwords while not worrying about forgetting/misplacing them.

Virtual Private Network (VPN) Serves as a tunnel that connects the user directly with the website/service they are trying to access, bypassing third parties that might intervene or observe their activity.