UNIVERSITÀ DI PARMA

Online security and anonymity

Matteo Gianvenuti

Academic year 2024/2025

Summary

Introduction

The private companies, the organizations and the government aim to collect as much information about you as possible—just like in all the precious years, of course. Now and in the future, this will be much easier. This is because everything is going digital, mainly through hardware devices and software applications that do not pursue the user's interest. They are not designed to protect your data, block the tracking attempts, anonymize, or protect your identity; on the contrary, they do the opposite.

As immediate consequence, the problem of personal safety and so anonymity (especially online) nowadays is greater than ever, and it will continue to grow. Those who controls the information of an individual have the full control over them. In a world like this democracy no longer exists. There can be no opposition in a system where people are completely controlled.

For these reasons, in this course thesis, we will attempt to set up a secure workstation by analysing the security of various privacy-focused technologies, such as VPNs, Tor and Whonix.

What identify and expose you online?

- **IP address:** The IP address serves two main functions: network interface identification and location addressing. It allows the packets delivery from the source host to the destination host on the internet (also in a private network) then it uniquely identify you. This mean the online activity is linked to the IP address and from it to a person or at least to a device. The IP address is typically assigned by the internet provider (more properly the internet access provider or the Internet Service Provider also known as ISP).
  According to Italian laws, the ISP have to collect lots of information on the users, like: Information about who communicated with whom (e.g., IP addresses, connection timestamps, and duration), the internet access logs (the data must be retained for 12 months, extendable in case of a "crime") and much more.[1][2][3][4]

- **Browser and browser fingerprint:** The browser configuration is important because through the browser set up is possible to understand who is. How many other users have your same configuration? This is the fingerprint; thanks to it you can easily distinguish between thousands of users when you are looking for someone. It contains also a lot of information about the device.
  Even the browser itself is important, did you know that Google Chrome send to Google LLC all the URLs you visited? (until few years ago in EU) Instead, a privacy-focused browser does not collect any type of information on you, furthermore, with a good set up it can clean all your data after every session (the local data).

- **Cookies:** They are used to identify users and typically contain information about their identity, the site visited, and other related data. Cookies can track whether you have visited a specific site or even a particular page, including across different domains (e.g., third-party cookies). They monitor your behaviour.

- **Cross-site tracking and data aggregation:** Usually, companies share their information about you with several other companies for various economic and non-economic advantages. This may pose a security risk to you because sharing your data allows them to build a more precise and accurate digital profile of you.

- **Application tracking:** Did you know that more than the 80% of the applications in the common stores like Google Play Store and Apple App Store contains a lot of trackers? They try to get as much as possible information on you, like: OS, CPU data, memory data, GPS position, IP address, network data, age, gender, address, and so on.[5]

- **Operative System and device:** Of course, the operative system (OS), as well as device security is crucial for your online safety. The most commonly used OSs are Windows, macOS, Android (with integrated Google services) and iOS. All these OSs assign unique identifiers to users to track their activities, so they can be used to build detailed profiles of users, monitor their online behaviour, and even share data with third parties without explicit consent.

- **Social networks:** They are one of the most aggressive data collectors and even one of the most dangerous. Basing on how you interact with them, what you watch and in general what you do in the platform they can estimate a digital version of you; they may know you better than your friends if you gave them enough data. With these data they can manipulate and change the content you see and then change your mind. That is why I do not have any social media/network.
  Did you know that in the last Italian elections, a "truly anti-system" party was shadow banned? The elections were clearly influenced and manipulated (as it has always been done), not only by social media/networks but also by Google and similar companies.

- **Metadata:** All the files contain some metadata. They are information like author, device, timestamp, geolocation, application used to create the file, etc. Then, remember this the next time you will upload a file online or on a social media/network. There are several tools for metadata removing. Do not make the same mistake as John McAfee, Rocco Castoro, Robert King and Vice magazine.

- **Extra considerations:** In general, almost every entity that interacts with you collects data on you in an aggressive way. If I had to do a complete list of what identify you it may be infinite. Do consider, especially if you are not aligned with the establishment, all the data collected by almost all the big and small companies are completely available for the intelligence agencies. Furthermore, these agencies have built a massive surveillance system.[6]

Alternatives that protect you (from simple to complex)

Virtual Private Network (VPN)

The VPNs just hide your IP address but unfortunately this is not enough to protect your identity, the big companies have lots of tracking script, cookie, bot, and so on in almost all entity that you may interact with. Then even if you hide your IP address with the time, they can reach you by data aggregation and cross-platform-referencing. This is mostly possible thanks to the fingerprint analysis.

It is good to know that almost all the VPN providers does not really protect you, the most insidious even sell your traffic data and analyse it! Keep in mind that your VPN provider know exactly what you are doing.

Even if you use a good service provider with a no log policy like Riseup there still be some problems:

- Can you really trust your VPN provider?!
- Even if a connection is tunnelled it can be hijacked inside the browser and typically browser hijackers also contain spyware.
- Even if your VPN provider have a no log policy, they may be forced by the police, intelligence agencies, and so on to track your activity with or even without a warrant. Of course, they will not inform you of this. In this case they could also do a man-in-the-middle attack.
- With a VPN you only pass through one server this easily allows the traffic analysis. If a passive attacker is able to monitor and record the traffic on network links entering and exiting clients or servers can enable powerful statistical attacks—e.g., trackers can identify their targets by their keyboard typing style, their visited sites, minimal divergence from the standard time in the network requests (time attacks), etc. Some VPN providers provide double VPN, but it is still the same provider and lacks distributed trust by design, even if you trust it there may be force majeure as we saw in the previous point. So, the servers used for tunnelling must not know the user. This is only possible if there are multiple and independent servers as relays on a strict need-to-know basis and onion-layered encryption.
- Port shadow attacks: an exploit called “port shadow” allows the attacker to hijack connections, deanonymize users, or redirect traffic. This vulnerability has been recently discovered and affect some popular client as OpenVPN (RiseupVPN is based on this client).
- Website fingerprint attack: Any local observer on the network (e.g., ISP, WLAN) can make estimates of websites requested over the VPN by simply analysing the size and timing of the encrypted VPN data stream.
- And much more.[7]

The Onion Router (Tor)

Maybe this will surprise you, but obviously, Tor as a standalone tool is not enough to protect and anonymize you at 100%.

Let us see some vulnerabilities and cases in which were possible to identify the user:

- First of all, if the destination host does not use an encrypted connection (e.g., https) your data are completely exposed for who monitor the destination host network and even for the last node of the Tor network. In general, the applicative layer must be encrypted.
- If the intelligence agencies (like any other organization) maintain and control various Tor nodes, they could potentially become part of your Tor path, act as the guard node, or even worse, control the entire path and monitor everything you do—this is a hypothesis that is possible, but not very probable.
- If someone is able to monitor the whole network, they can do an end-to-end correlation attack—this is a weakness all the low-latency anonymous networks have. It means, Tor cannot protect against an attacker who simultaneously monitor traffic at both the entry and exit points of the Tor network. While Tor does provide protection against traffic analysis—especially when there is enough traffic to make it impractical—it cannot prevent traffic confirmation via end-to-end correlation. This attack was used to identify a Harvard student threatened to use bombs; the student was sending Tor traffic at the exact times the bomb threats were made, through the university network and an account of another student.
- The Tor browser was developed to provide privacy and anonymity for a wide variety of users. This can increase the volume of traffic on the Tor network, making it harder to identify users through statistical and correlation attacks. However, it only secures the traffic within browser. What about the rest of the device’s traffic? Using only the Tor browser is not enough to ensure you full protection.
- Tor as a standalone tool does not remove the fingerprint of the device and the applications. The Tor browser try to minimize the information leaked by the application to reduce the fingerprint, but it does not clear everything at all.
- The Tor browser is built on Firefox, in the previous years it had vulnerabilities (now patched but it may have more) that allowed the FBI and others to find users who thought to be safe and anonym online—thanks to an exploit called “EgotisticalGiraffe”, known from the Edward Snowden leaks in 2013. “EgotisticalGiraffe” exploits a type of confusion vulnerability in E4X, which is an XML extension for JavaScript. In particular, the Firefox browsers in many older versions of the Tor Browser Bundle were vulnerable to a JavaScript-deployed shellcode attack, as NoScript was not enabled by default. This vulnerability was used to extract users' MAC and IP addresses and Windows computer names.
- And much more. Keep in mind that mostly of the vulnerabilities are kept secret and exploited until someone reveal them.[8]

Whonix

Whonix is an operating system focused on anonymity and security. It is architecture is based on the Whonix-gateway and Whonix-workstation. This means, it is used as guest on another secure OS—like Qubes, Fedora, etc with LVM, Luks, etc for the physical-offline security. Shortly, with the cited architecture every application run in a virtual machine in an isolated network, the workstation; the workstation's traffic is forwarded to the gateway and from here to the internet through Tor—it also supports many others anonymising networks. It protects the whole device/workstation traffic basing on the setup.

Whonix can practically solve all the previously cited problems and more:

- Whonix is based on Kicksecure to keep the system protected by viruses and malwares.
- It uses Kloak to prevent the exposing of your typing style.
- Java, JavaScript, Flash, browser plugins and mis-configured applications cannot leak the user's real external IP address. So, it gives a best as possible fingerprint protection.
- Tor and Tor Browser are not running inside the same virtual machine which means an exploit in the browser cannot affect the integrity of the Tor process.
- It prevents and mitigates time attacks.
- Distinct pre-installed applications are routed through different network paths for maximum security.

Whonix may need a custom setup basing on your needs and the applications you use. There still be some considerations to keep in mind:

- Is possible to understand you are using Tor because the guard nodes are public, so some servers/services may limit your options.
- Of course, you should not share personal data, metadata, etc on the internet.
- In the unlikely event someone control all the Tor nodes Whonix may not protect you.
- It is open source, that is an advantage because you check and trust it by yourself, but someone may search for bugs in the source code to attack you.

"I use Qubes and a Whonix gateway literally everyday, …" cit. Edward Snowden.

He is free this means it works with the proper setup. Otherwise, they would have taken him.

There are many other things to say, but the space for this course's thesis is finished. For more details you can check the references.[9]

Conclusion

The world is moving more and more towards a sort of "matrix" in that everyone is controlled at all. This is possible thanks to ignorance and complicity. The people use systems insecure by design. Modern surveillance threatens not only individual privacy but also the freedom to dissent.

Pay attention at the services you use; you make it possible.

Bibliography and notes

[1] D.Lgs. 109/2008: https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/1607282.

[2] D.Lgs. 196/2003: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2003-06-30;196.

[3] GDPR: https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:32016R0679.

[4] Law 167/2017: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:legge:2017;167.

[5] You can check it through the anti-trackers integrated in the mobile browser DuckDuckGo or even better through TrackerControl form the F-Droid store.

[6] From "Errore di sistema" / "Permanent record" the Edward Snowden autobiography.

[7] Mostly of the VPN section content came from the following link and previous knowledge. The time attack came from Cybersecurity Luca Veltri's course. Minor things came from other sites. https://www.whonix.org/wiki/Whonix_versus_VPNs.

[8] Tor vulnerabilities: Cybersecurity Luca Veltri's course, https://github.com/Attacks-on-Tor/Attacks-on-Tor, https://en.wikipedia.org/wiki/Tor_(network) and https://www.schneier.com/blog/archives/2013/10/how_the_nsa_att.html.

[9] Whonix: https://www.whonix.org/, https://www.whonix.org/wiki/About, https://www.whonix.org/wiki/Features, https://www.whonix.org/wiki/Keystroke_Deanonymization#Kloak, https://www.whonix.org/#security and https://www.whonix.org/wiki/Anonymity.

Academic year 2024/2025