Proximity Tracing Applications: The misleading debate about centralised versus decentralised approaches

Inria, France*

April 18, 2020

Proximity tracing applications

The traditional way to conduct contact tracing so far is through interviews with individuals diagnosed as carriers of the virus. During this process, personal data about diagnosed people and people who have been in contact with them is collected and processed by the health authority. This procedure is effective to combat the spread of a virus but it is cumbersome and time consuming.

The main goal of proximity tracing mobile applications is to notify people that they have been in close proximity of COVID-19 virus carriers even though the carriers did not have symptoms and were not even tested at the time of interaction.

As stated by the EDPB [1] "envisaged technical solutions need to be examined in detail, on a case-by-case basis". Indeed, each technical solution for proximity tracing applications is based on a complex architecture, including at least applications installed on users' smartphones and back-end system(s) controlled by the authorities.

Data minimisation principle

Proximity tracing applications should follow the *data minimisation principle* of the GDPR. If the only objective of the application is to notify people who have been in contact with diagnosed COVID-positive individuals, then the application should collect and process only data needed for this purpose, hence *data about proximity of people*. Any other information, such as users' location, phone numbers, names or any other information that can be used to identify the individual, should not be collected.

Need for a back-end system

Several approaches described as *decentralised* have been proposed. However, a "fully decentralised" approach is not realistic for proximity tracing. Indeed, only relying on data exchanges between applications to inform who is at risk or not would not be a secure and reliable solution (the reception of these notifications would depend on the current proximity between people leading to slow and incomplete information delivery). Therefore, most proposals actually involve a back-end server sharing information with the applications. In a such architecture, two design choices are of the utmost importance (1) how to distribute the data between the server and the device and (2) where the status of the user (at risk or not) is verified.

With regard to the first choice, in order to comply with the data minimization principle of the GDPR, the contact information collected by the application should be stored only locally (on the mobile phone). To provide a quick and reliable notification to at risk users, only data allowing to establish proximity contacts of individuals diagnosed positive should be sent to the back-end server. Most solutions follow this approach.

*Collaborative Inria work led by the PRIVATICS team.

Therefore, the debate about "centralised" versus "decentralised" approaches concerns essentially the second design choice, i.e., where the status of the user is verified: this verification is performed on the device of the user in the so called "decentralised" solutions and on the central server otherwise.

If the verification is performed locally, all applications must receive information about the users diagnosed positive to verify if any of them are part of their contact list. It has been shown [3] that this is a serious source of privacy risk because it makes it easier for malicious users to detect that a person has been diagnosed positive. This may lead to stigmatization which is seen by many experts as a major risk for contact tracing applications [2]. On the other hand, performing this verification on the central server does not represent a significant privacy risk if the server knows only pseudonyms and is not able to relate them to the users. In particular, it does not pave the way to mass surveillance (which is the second ethical issue identified in [2]), unless malicious authorities massively deploy Bluetooth receivers. It is important also to notice that no information about the fact that a user has been diagnosed positive should be stored on the server. This back-end system must nevertheless be secured, and regularly audited and controlled by trusted independent authorities (such as Data Protection Authorities and National Cybersecurity Agencies of EU Member states). The privacy impact of the choices between local and central computations is therefore not as simple as it may look.

Risk analysis approach

Considering that all solutions involve a central server and local applications, the debate should therefore not focus on the catchwords "centralized" and "decentralized". It should rather rely on a precise analysis of the privacy and ethical impacts of each solution. This analysis should address the following key issues:

- What data is collected and how is it protected? Is it possible to relate the data to persons? Where is this data stored?
- What are the possible sources of risk (e.g. authority, users of the application, malicious third parties, etc.)?
- What are the capabilities of these sources of risk? For example, can they modify the application or install wireless antennas?
- What could be the intentions of these sources of risk? For example, is it possible that they may want to compromise the system or to know which users have been diagnosed positive?
- How hard is for such sources of risk to reach their goals given their capabilities?
- What is the robustness, quality and reliability of the application and its added value to combat COVID-19?
- Considering the answers to the previous questions, what would be the privacy and ethical impacts of potential breaches, including their likelihood and severity and scope? These impacts should be balanced with the added value of the application.

Such questions can be answered only by analysing each technical solution on a case-by-case basis, ideally together by technical experts and legal experts. In addition, an impact analysis should take into account the views of civil society and other disciplines such as ethics and sociology.

Conclusion

The debate about proximity tracing applications is of high importance to all the EU Member States and all the fundamental rights of individuals residing in those states. We underline the importance of this debate and encourage *to compare technical solutions based on privacy risk assessment rather than on ill-defined catchwords such as "centralised" vs "decentralised"*.

References

[1] EDPB open letter. Ref:OUT2020-0028. https://edpb.europa.eu/sites/edpb/files/files/file1/edpbletterecadvisecodiv-appguidance_final.pdf.

[2] Ethical issues of anti-pandemic applications. http://www.lecre.umontreal.ca/les-enjeux-ethiques-des-applications-anti-pandemie/.

[3] Serge Vaudenay. Analysis of DP3T. Cryptology ePrint Archive, Report 2020/399, 2020. https://eprint.iacr.org/2020/399.