Contact tracing: an overview of the challenges

Bruno Sportisse, Inria CEO

(Updated translation from a French tribune, April 18)

The French digital ecosystem is fully involved in the fight against Covid19. From the outbreak of the disease, researchers, developers and firms have been sharing their skills and expertise with healthcare personnel and doctors in their efforts to find medical solutions (vaccines and treatments). At Inria, the National Institute for Research in Digital Science and Technology which I have the honour of heading, more than twenty projects are underway with our partners in French public research organizations (CNRS, CEA, INSERM, INRAE, Universities, …) as well as with companies. The goal of these often highly operational projects is to help hospitals in the digital management of the crisis and to support medical research.

The spontaneous mobilisation of scientists and engineers from the digital ecosystem has been nothing short of exceptional. This statement is also an opportunity to pay tribute to them.

Among the projects, the issue of "contact tracing" has become one of France's primary concerns since the announcement on 8 April 2020 by Olivier Véran and Cédric O, the Ministers for Health and for Digital Affairs respectively, that work was underway to build the prototype for a French application, STOPCOVID, as part of a comprehensive lockdown exit strategy. Leadership of the project, which involves both public and private stakeholders, has been entrusted to Inria.

Through Inria, France is participating in the PEPP-PT initiative alongside German, Italian and Swiss teams. Together with our Fraunhofer partners, the Inria teams are today publishing the ROBERT protocol – ROBust and privacy-presERving proximity Tracing – which represents the current state of our reflections on the technical architecture of a contact-tracing application that respects European values. This protocol is available on Github (https://github.com/ROBERT-proximity-tracing/), in line with standard scientific practice.

My aim here is to outline what this protocol contains using words that are understandable to all, but above all to explain its assumptions and the spirit in which it was devised. In all the urgency of the exceptional circumstances we are experiencing, it is important to report calmly on a difficult topic that has multiple dimensions.

To start with, and in light of the legitimate questions being raised and the potential confusion that may surround this issue, I think it is useful to state what an application based on this protocol is not, with the credible assumption that we live in democratic societies, with regulation bodies that do their jobs.

This application is not a "tracking" app: it only uses Bluetooth, never GSM or geolocation data.

Nor is it a surveillance app. To be even clearer: it has been designed in such a way that NOBODY, not even the government, has access to the list of people diagnosed as positive or to the list of social interactions between people. The only information I would receive would

be a notification that in the last few days, my smartphone has been in close proximity to the phone of at least one person who has since tested positive and has reported this to the app.

This application is not an “informant”. If I receive a notification, I do not know who was at its origin. When I myself report that I have tested positive, I do not know who is notified.

It is not a compulsory application. Users choose to install it. They choose to enable their Bluetooth. They can disable Bluetooth or uninstall the app at any time.

This having been clarified, what does the application do?

“Contact tracing” refers to the ability to inform a person, through an app on their smartphone, that in recent days (typically two to three weeks) they have been in contact with people who have tested positive for Covid19. This person, or “contact case”, therefore presents a risk of being a carrier of the virus and accelerating the spread of the epidemic. The digital technology used to qualify this risk is based on the ability of two smartphones to recognise that they are close to each other through Bluetooth technology, which only works at short distances (a few metres). Many projects thus prefer the term “proximity tracing”, which is more precise in describing the role played by smartphones, and is a term that I will adopt from now on. No geolocation technology (indicating place and time) is used.

Thus defined, proximity tracing obviously raises many questions, particularly about its use: how does a person diagnosed as positive use this information? Further to this information, what instructions should be followed by/for the people identified at risk, as part of a global health strategy? We will return to these points later.

What is its purpose?

Medical staff have longstanding field investigation experience in tracing the chains of propagation of an epidemic. As such, this type of app should therefore be viewed only as a complement to these practices. A significant example of this is the Trace Together app used in Singapore.

More recently, scientific work carried out by an epidemiologist at Oxford, Christophe Fraser, has shown, *based on simulations*, that the use of proximity tracing is a useful aid in breaking the spread of the epidemic. His multidisciplinary team simulated progress of the disease over 250 days in a fictitious (modelled) city of one million inhabitants and showed the impact of using the app according to its level of diffusion (from 0 to 80%). To summarise the results of this simulation very simply, the act of downloading the application by one or two people (depending on the case) leads to a reduction in transmission of the virus to just one person. Naturally, many parameters are likely to influence its impact (such as transmission “false positives”, or the extent to which people adopt protective behaviour). This research highlights a trend, one which is in line with common sense.

At all events, all of the teams working on these issues are fully aware that proximity tracing is merely one component in a much larger set of measures in the framework of an approach

guided by healthcare policy. I personally do not know anyone who believes in technological solutionism in this respect.

What are the components of this system?

As well as adopting a humble approach, it is important to be transparent when presenting all the dimensions of this system.

To start with, it has **technological limitations**: Bluetooth technology was not designed to be precise in its estimation of distances between two smartphones. Results may depend on many parameters, such as people's physiology, smartphone position, type of smartphone, battery status, etc. This limitation has led several international teams to conduct calibration tests in order to propose statistical models that correct these errors. This is the case, for example, of the German teams in the framework of the European PEPP-PT initiative (to which I will return later).

Another limitation concerns the **virus transmission model**, which remains very uncertain: aerosol transmission or transmission via droplets (which are larger), with an impact on the time the virus spends airborne? Via surfaces? How should the viral load be estimated? Etc. All proximity-tracing apps are based on risk functions that have been defined with epidemiological researchers based on the current state of the art. This knowledge is still very incomplete and is likely to change very quickly according to feedback.

Lastly, one key and highly sensitive subject relates to the **information sharing protocol**. This is a sensitive subject because it is linked to both political and democratic concerns: what information will be shared and who will it be shared with? Which authorities/organisations will be responsible for operating these exchange systems? Who are we putting our trust in? What are the security levels and what possible attacks are we prepared to withstand in a credible manner?

Firstly, none of the projects are planning to set up a peer-to-peer network, where everything would be based around a supposedly "independent" (I will come back to this point) community of terminals/smartphones sharing information with each other. The main reason for this is the potential impact of the security breaches which could occur with this sort of approach

Accordingly, all the systems proposed comprise a shared component (a server) and a decentralised component (a set of smartphones able to communicate with each other via Bluetooth): **all the systems currently being assessed are therefore both centralised** (I will return to the use of this term below) **and decentralised**.

In this context, debates around the supposed advantages of a decentralised system over a centralised system do not seem to me to meet the criteria of scientific rigour. The term "centralised" is often used with the deliberate intent of implicitly stigmatising a state as wanting to survey people. Supposedly decentralised approaches, which those reluctant to place their trust in a central authority may be more comfortable with, can have major weaknesses from the point of view of information privacy. This can be demonstrated through

scientific analysis, which, by definition, is verifiable, and which lends itself to discussion, as opposed to ideological concerns or semantic preconceptions.

Secondly, what information will be shared?

In all projects which respect the European privacy framework (the GDPR, which is governed by independent authorities, such as the CNIL in France), information circulates in the form of "crypto-identifiers": pseudonymised data, normally generated only for a brief window of time (typically 15 minutes), which is linked to a terminal and *not an individual*. In other words, as a smartphone travels, it will encounter fleeting crypto-identifiers (those of the smartphones it encounters).

Existing projects differ in terms of the type of information sent to these central servers, operated either by a health body, a trusted partner or even a company.

Today, as has been mentioned earlier, the French and German teams from Inria and the Fraunhofer Society have released a joint protocol named ROBERT, which stands for ROBust and privacy-presERving proximity Tracing.

I would like to take this opportunity to mention a number of my colleagues from Inria with whom I have had the pleasure of working over the past 3 weeks, alongside Eric Fleury, who is also on board for this adventure: Nataliia Bielova, Antoine Boutet, Claude Castelluccia, Mathieu Cunche, Cédric Lauradoux, Daniel Le Métayer and Vincent Roca (in alphabetical order). Experts in high level IT concepts (concerning cryptographic protocols and transmission protocols) and with a firm grasp of the relevant societal and ethical issues, these people are the pride of French research in computer science and informatics. The name of their team, PRIVATICS, says it all. It would be remiss of me not to thank my German colleagues as well: Claudia Eckert, Alexander Küchler, Martin Schanzenbach and Julian Schütte from AISEC, the Fraunhofer Society's cybersecurity institute. This is also a great example of the Franco-German cooperation.

This protocol is based on the following principles, which are explained in detail in the example below:

- An individual keen to help tackle the spread of the virus *voluntarily* downloads the application onto his or her smartphone. This person's smartphone is then sent a set of crypto-identifiers (or a method of generating them every 15 minutes).
- By leaving their Bluetooth switched on, the owner of the smartphone will enable this application to build up a history of the crypto-identifiers encountered "nearby" for a significant length of time while out and about (these crypto-identifiers being on the smartphones of other individuals who have also downloaded the application).
- If the individual has tested positive for coronavirus, a history of the crypto-identifiers they have encountered will be relayed to the server of a health body (for example), without divulging to the server their own crypto-identifiers. There will be no link between the individual's phone and their history. Each of these crypto-identifiers is therefore potentially "at risk" (corresponding, without any possible link to an individual, to a smartphone which was in the vicinity of another smartphone carried by an individual who tested positive for coronavirus at a later stage).

- Furthermore, in the context of the model proposed for use in France, each smartphone onto which the application has been downloaded will “check in” with the central server (every hour, every day, depending on the configuration) to determine whether or not their crypto-identifiers are among those at risk. If this is the case, this would indicate that, in the past few days, the smartphone had been in the vicinity of a smartphone carried by an individual who tested positive at a later stage.
- Notifications will be issued based on a risk assessment (for which the calculation must be defined in conjunction with epidemiologists, as I have already mentioned) using proximity information. The flexibility of the system is key when it comes to managing a public health crisis, factoring in ever-changing medical knowledge, or even having the system learn this (while always based on anonymised statistical data) in order to make it more efficient and to reduce the occurrence of false positives, for example. Provided that the health body is in control of all of this, of course (a key point, which I will return to later).
- This information can then trigger various actions (this is not the subject of this article): scrupulously following shielding measures, tracking symptoms on a daily basis, doctor’s appointments, tests, etc. This will depend on the public health policies of the country in question.

In the context of this approach, a number of strong choices have been made, which I would like to return to, given that this research has been carried out under pressure but very much in a framework of shared values. Everything depends on consent and voluntary action; it is essential that we are not able to infer that our neighbour is positive and that they may have infected us; when such an application is deployed, a server must not collect a list of all of the people infected (this is a really important point); and decisions on public health policies are the responsibility of sovereign states.

For example, on the central server (if we are to accept this term), there will be NO data relating to the status of individuals who have tested positive. Instead, the server will contain a list of the crypto-identifiers for the smartphones located in the vicinity of other phones belonging to individuals who have tested positive.

Here’s another example of a strong choice: on my neighbour's smartphone, there will be NO data concerning my medical diagnosis, even in encrypted form. Instead, there will be a list of the crypto-identifiers for all of the smartphones encountered.

Another example of a responsible public health policy: both the parameters of the transmission model and the anonymous statistical data will be in the hands of the health body – they will be in charge of use of the system, not a private company, however innovative it might be.

Other options are available, forming the basis of other approaches which differ from those adopted by Inria and Fraunhofer: crypto-identifiers for individuals who have tested positive could pass through central servers before being sent to all smartphones. Smartphones within which the crypto-identifiers encountered and the crypto-identifiers of individuals who have tested positive could be matched up. This is a system that can be described as being very much decentralised – but it can also be described as “decentralised centralisation”: on each

smartphone, there will be a list of all of the crypto-identifiers for the individuals who have tested positive. Another advantage of this system is that it is easily enabled by the API unveiled by Apple and Google a week ago, which is expected to be available by mid-May and which will be the first of its kind in the history of computing.

At all events, it will be up to governments to decide whether or not to use a protocol in accordance with their policies. It is our responsibility as scientists, meanwhile, to give them the resources they need to make such decisions.

Is this protocol definitive and has it been fully finalised?

The short answer is no. It is a work in progress, and will be subject to scrutiny.

Firstly, as is the case with all scientific projects, the protocol will be submitted for peer review. This will require an open approach: the scientific paper has been made available to the scientific community via Github (https://github.com/ROBERT-proximity-tracing/). Potential breaches, attacks and suggestions will be put forward. Much has already been done, including through exchanges with the ANSSI, France's National Cybersecurity Agency, which must take great credit for its remarkable level of engagement.

Furthermore, as some have already noted, contact-tracing apps require a number of prerequisites, meaning they cannot be deployed in their current state on all smartphones used by the people of France. For example, Bluetooth has to be able to function effectively, even when the smartphone is on standby. What this means is that we need to work with the designers of operating systems in order to open up this possibility.

New versions will therefore appear in the future, but an initial software implementation is currently being developed based on the ROBERT protocol. As is the case with everything linked to the STOPCOVID project, it will be made available opensource, using an MPL (Mozilla Public License). Thanks to our history, Inria is only too aware of the impact the opensource model can have on improving software technology and making it more transparent, not to mention allowing it to be used by other countries that may not have the capacity to develop this type of software. This is also very much the focus of the W3C's Open Web Standards – it is my honour to be president of W3C's European hub, which underlines Inria's commitment to these values.

Another challenge is interoperability

This is the reason for France's involvement, through Inria, in the PEPP-PT initiative, which we discussed earlier, alongside teams from Germany, Italy and other countries. We might not always agree on the decisions taken or on the hypotheses proposed (e.g. who is most likely to carry out an attack? A democratic country or a hacker, perhaps not even a particularly smart one, and certainly not always an ethical one). We might not all place the same emphasis on the issues of digital and technological sovereignty. Regardless, we work together on a scientific and technological playing field, to build solutions which respect the values that we share and which are interoperable (the applications deployed will have shared components but will be national given the importance of forming part of a national healthcare system).

During these unprecedented times for our nation, France can turn to its research and innovation ecosystem to successfully carry out projects capable of simultaneously meeting the required levels of effectiveness for public health policies, respecting individual freedoms and maintaining or even reinforcing our digital and technological sovereignty.