INSTITUT DE FRANCE
Académie des sciences

Traçage numérique

7 mai 2020

Lorsqu'une personne contracte le COVID-19, identifier les personnes qu'elle a croisées fait partie des tâches utiles pour maîtriser l'épidémie. L'idée du « traçage manuel » est de remonter la chaîne des contacts récents du patient pour détecter les personnes à qui il a pu transmettre le virus, afin de les dépister à leur tour ou de les isoler par précaution si nécessaire.

Quand l'épidémie est trop développée, ce difficile travail de terrain devient délicat à réaliser, coûteux. On peut alors penser à proposer, en complément du traçage manuel, un « traçage numérique », par le biais d'une application hébergée sur les téléphones portables (*smartphones*). Ce système peut être imposé et peu respectueux de la vie privée, comme l'ont fait Taïwan ou la Corée du Sud, ou proposé sur **la base du volontariat**, tout en essayant autant que possible de **protéger la confidentialité des données** comme envisagé en France.

Le principe est d'utiliser le signal Bluetooth LE (pour Low Energy, ou BLE) des appareils pour repérer les personnes à proximité. Bluetooth est une norme de communication permettant l'échange bidirectionnel de données à courte distance en utilisant des ondes radio UHF sur une bande de fréquence de 2,4 GHz. Son intérêt est de simplifier les connexions entre les appareils électroniques, ici les *smartphones,* en supprimant des liaisons filaires. BLE permet de réduire la consommation énergétique d'un facteur allant de 20 à 100.

Si une personne tombe malade, il reste alors à prévenir ses contacts repérés et enregistrés par son téléphone et, ce faisant, selon les préconisations des épidémiologistes, leur proposer de consulter un médecin, se faire tester... Précisons que ceci vient en complément de l'utilisation des gestes barrières et du port des masques, et de traçage manuel, que l'efficacité dépend crucialement de la disponibilité de tests PCR ou LAMP en nombre suffisant, et que le dispositif n'est qu'une composante d'un ensemble plus vaste de mesures, dans le cadre d'une approche pilotée par une politique de santé.

Le système utilisant le Bluetooth est plus satisfaisant au regard du **respect de la vie privée** que les systèmes reposant sur la géolocalisation cellulaire ou GPS, à l'instar de ce qui s'est fait en Chine, et au sujet auxquels la plupart des pays européens, ainsi que la Commission européenne[1], ont déclaré leur opposition.

Selon des modèles prévisionnels réalisés par des chercheurs d'Oxford[2], il faudrait au minimum que 60% de la population installe et utilise convenablement une telle application pour qu'elle soit utile. Si l'on en croit cette étude, elle n'a d'intérêt qu'à condition **qu'une part très substantielle de la population l'adopte**. D'autres spécialistes pensent au contraire qu'elle constitue un complément utile au traçage manuel **largement indépendamment du nombre de personnes qui l'adoptent**[3].

Tous les systèmes en projet dont nous avons connaissance en Europe et en France en particulier comportent une composante commune (un serveur) et une composante décentralisée (un ensemble

[1] https://ec.europa.eu/commission/presscorner/detail/en/ip_20_670
[2] https://science.sciencemag.org/content/early/2020/04/09/science.abb6936
[3] Voir les minutes 40-45 de https://re.livecasts.eu/webinar-on-contact-tracing-applications/program

de *smartphones* qui peuvent communiquer entre eux à travers le Bluetooth). Ils sont donc tous **à la fois centralisés et décentralisés**.

Différents protocoles ont été développés par des acteurs du monde de la recherche et de l'industrie, citons le protocole DP3T[4] et sa déclinaison par l'alliance Google/Apple ainsi que le protocole ROBERT[5] proposé pour l'application STOPCOVID en cours de mise au point dans le cadre de l'initiative européenne PEPP-PT (pour Pan European Privacy Preserving Proximity Tracing) et dont le maître d'œuvre est Inria à l'intérieur d'un consortium incluant des grandes entreprises du numérique. **Ces deux protocoles sont disponibles**, comme le veulent les pratiques scientifiques de **science ouverte**, et peuvent ainsi être analysés et vérifiés par les spécialistes. S'il existe des différences entre les deux protocoles, les opposer sur une classification comme « centralisé » ou « non centralisé » tient plutôt du marketing car ils utilisent tous deux un serveur ; par contre, ils présentent chacun des avantages et des désavantages[6].

Principe général du système STOPCOVID[7]

Une personne, soucieuse de participer à la lutte contre la propagation de l'épidémie, télécharge **de manière volontaire** l'application sur son *smartphone*. Celui-ci reçoit alors un ensemble de **crypto-identifiants** (ou une méthode pour les générer toutes les 15 minutes) c'est-à-dire des **pseudonyme**s.

Le détenteur du *smartphone*, en laissant le Bluethooth activé, permet à son application de construire un **historique des crypto-identifiants rencontrés**, « à proximité », pendant une durée significative, lors des déplacements (ces crypto-identifiants sont stockés sur son *smartphone*).

Si la personne est **diagnostiquée positive**, elle fait remonter son historique de crypto-identifiants rencontrés sur un serveur d'une autorité de santé (par exemple), sans divulguer au serveur ses propres crypto-identifiants. Aucun lien n'est fait entre le téléphone de la personne et son historique. Chacun de ces crypto-identifiants est donc potentiellement « à risque ». En effet il correspond, sans qu'aucun lien ne soit possible avec une personne, à un *smartphone* qui a été en proximité d'un *smartphone* porté par une personne qui a été ultérieurement diagnostiquée positive par l'autorité de santé.

Par ailleurs, chaque *smartphone* ayant téléchargé l'application vérifie auprès du serveur central, « de temps à temps » (toutes les heures, tous les jours, etc.) si **ses propres crypto-identifiants sont parmi ceux à risque**. Si c'est le cas, cela signifie que le *smartphone* a été à proximité lors des jours précédents d'un *smartphone* porté par une personne qui s'est avérée ultérieurement être malade et donc qu'il se peut que le propriétaire du *smartphone* ait été contaminé.

Cette information peut alors déclencher une notification déterminée sur la base d'une évaluation du risque (dont le calcul est défini avec les épidémiologistes) en utilisant l'information relative au nombre de personnes déclarées infectées à proximité desquelles le *smartphone* s'est trouvé. La notification est accompagnée de recommandations telles qu'un respect scrupuleux des gestes barrières, un suivi journalier des symptômes, une suggestion de consultation, de test, etc. Ceci relève du choix d'une politique de santé de l'Etat.

---

[4] DP3T =Decentralized Privacy-Preserving Proximity Tracing : https://github.com/DP-3T/documents/

[5] ROBERT = ROBust and privacy-presERving proximity Tracing,
Une présentation sous forme graphique, simple à comprendre, du protocole ROBERT se trouve en :
https://github.com/ROBERT-proximity-tracing/documents/blob/master/ROBERT-infography-FR.pdf

[6] Par exemple, avec ROBERT, l'accès au serveur donne des informations sur le graphe social des contaminés, et avec DP3T, des attaques extérieures pour trouver qui est contaminé sont plus faciles.

[7] https://www.inria.fr/fr/contact-tracing-bruno-sportisse-pdg-dinria-donne-quelques-elements-pour-mieux-comprendre-les-enjeux

L'approche STOPCOVID et le cadre européen de protection de la vie privée

En accord avec le RGPD (Règlement Général sur la Protection des Données) et sous le contrôle des autorités indépendantes comme la CNIL, les informations circulent sous la forme de « crypto-identifiants », des données pseudonymisées, en général générées de manière éphémère et associées à un terminal et non à une personne.

Dans cette approche tout repose sur **le volontariat** à installer l'application sur son smartphone.

Il n'y a *a priori* aucune donnée relative au statut des personnes positives sur le serveur central. Il s'y trouve une liste de crypto-identifiants des *smartphones* s'étant trouvés à proximité des *smartphones* des personnes déclarées malades.

Dans le smartphone d'une personne que l'on aurait croisée, il n'y a **aucune donnée concernant son propre diagnostic médical**, aussi encrypté soit-il. Il n'y a qu'une liste des crypto-identifiants de tous les *smartphones* rencontrés.

Les paramètres du modèle de transmission et les données statistiques anonymes collectées par l'application et le(s) serveur(s) sont entre les mains de l'autorité de santé qui fixe l'utilisation de ce système.

Les limites de l'approche STOPCOVID

Il y en a un certain nombre. Par exemple, **les mesures de distance obtenues par Bluetooth ne sont pas de bonne qualité** et des travaux (notamment autour du protocole Bluetooth) sont en cours. Cela a conduit également plusieurs équipes internationales à mener des tests d'étalonnage pour proposer des modèles statistiques pour mieux estimer les distances. C'est par exemple le cas des équipes allemandes dans le cadre de l'initiative européenne PEPP-PT citée plus haut. D'autre part les signaux Bluetooth peuvent passer les murs et peuvent donc **générer de faux contacts**. Par ailleurs tout le monde n'a pas de smartphone. Pour pallier ce problème des adaptations sur objets connectés comme des bracelets sont envisagées.

Comme dans toutes les utilisations d'applications sur *smartphone*, notamment avec Bluetooth, des possibilités de dé-pseudonymisation par des acteurs malveillants existent. Le protocole Robert peut néanmoins être amélioré pour offrir une meilleure résistance à ces attaques potentielles. **Des atteintes à la vie privée ne pourront jamais être complètement écartées**[8]. Il faudra vérifier qu'elles restent raisonnables compte tenu des gains en terme de santé publique que l'application apporte. Il est bon de rappeler dans ce contexte que nous utilisons quotidiennement des applications sur nos *smartphones* **bien plus invasives** au niveau des données personnelles, comme les publicités de tiers et le traçage qui les accompagne.

Une des difficultés est que les gains sanitaires sont difficiles à évaluer, et les règles pour qualifier un contact de suspect (durée, distance) encore incertaines à cause d'une épidémie aux propriétés mal connues. Les épidémiologistes en s'appuyant sur des modélisations mathématiques pourront au fil de l'eau les estimer au mieux des connaissances acquises.

Il semble que, si l'on met en regard les avantages, **des vies peuvent être sauvées**, et les inconvénients, **des atteintes à la vie privée sont possibles**, du déploiement d'une telle application, la balance penche du côté de son adoption par le plus grand nombre de citoyens. Évidemment, cette adoption doit

[8] Voir l'excellente description de certaines d'entre elles **ici**.

s'accompagner de toute une panoplie de précautions, essentielles d'ailleurs pour gagner son adoption par tous. Il faut veiller aussi à ce que son adoption pour une période limitée dans le temps ne favorise pas à l'avenir l'usage de ce même type de technologie pour d'autres fins, comme le soulignent les membres de la Commission Nationale Consultative des Droits de l'Homme (CNCDH) dans leur avis du 28 avril dernier sur le suivi numérique des personnes[9].

Une telle application **n'est pas une application de « pistage »** : elle n'utilise que le Bluetooth, en aucun cas les données de géolocalisation cellulaire ni de géolocalisation GPS.

Ce n'est pas non plus une application de surveillance : elle préserve convenablement l'anonymat. Si les données du serveur sont correctement protégées et leur utilisation limitée, sa conception minimise les risques que quelqu'un, par exemple l'État, ait accès à la liste des personnes diagnostiquées malades ou à la liste des interactions sociales entre les personnes. La seule information qui est notifiée à un utilisateur est que son *smartphone* s'est trouvé dans les jours précédents à proximité du *smartphone* d'au moins une personne qui a, depuis, été diagnostiquée malade et s'est déclarée comme telle dans l'application.

Ce n'est pas une application de délation : dans le cas où une personne est notifiée, elle ne sait pas qui est à l'origine de la notification. Lorsque c'est elle qui est déclarée malade par l'autorité de santé, elle ne sait pas qui est notifié.

L'utilisation des données générées par l'application est **limitée exclusivement au traçage des contaminations du COVID-19** et **les données sont détruites après quelques semaines dans les téléphones et sur le serveur**. Ses utilisateurs **choisissent de l'installer. Ils choisissent d'activer le Bluetooth.** Ils peuvent, à tout moment, le désactiver ou **désinstaller l'application.** Ils **choisissent** aussi de se déclarer comme **potentiellement contaminés.**

Cette fiche a été conçue et rédigée par la cellule de crise Coronavirus de l'Académie des sciences. Créée à l'initiative de Pascale Cossart, Secrétaire perpétuel de l'Académie, celle-ci réunit des académiciens experts du domaine : Jean-François Bach, Pierre Corvol, Dominique Costagliola, Pascale Cossart (coordinatrice), Patrick Couvreur, Olivier Faugeras, Daniel Louvard, Félix Rey, Philippe Sansonetti, Alain-Jacques Valleron.

Les informations qui figurent sur cette fiche ont été produites collectivement et sont susceptibles d'évoluer. Elles seront éventuellement réactualisées en fonction des avancées des connaissances scientifiques.

[9] Pour voir cet avis :
https://www.cncdh.fr/sites/default/files/avis_2020_-_3_-_200424_avis_suivi_numerique_des_personnes.pdf