ECOLE DES MINES DE SAINT-E´ TIENNE

Covid-19 pandemic and proximitytracing

Submitted by:
Aninda MAULIK,
CPS2

Supervisor:
Prof. P. JAILLON

January 13, 2021

Contents

1. Synthesize and draw all exchanges describe in protocol specification for the two proposal

The

1.1. proximity contact recording

1.2. exposure verification

1.3. infection declaration

2. Explain why people say DP-3T is decentralised protocol and not Robert:

DP-3T:- Decentralized Privacy-Preserving Proximity Tracing (DP-3T, stylized as dp3t) is an open protocol developed in response to the COVID-19 pandemic to facilitate digital contact tracing of infected participants. [1]

open protocol:- An open protocol allows vendors' equipment to interoperate without the need for a proprietary interface or gateway. They talk the same language and no translation is needed. A closed protocol is one that is proprietary and not open to communication other products without an interface or gateway (such as SNA).[2]

decentralised protocol:- A decentralized protocol is a protocol where client and host nodes combine to create a general network. Both the client and hosts nodes must be supported by the software used for the protocol. The host nodes are connected to form a type of backbone for the network, providing a gateway to the network for client nodes.[3]

Robert protocol:It's a communication protocol — in informatics this means a.
"procedure" that describes how an application should work. [4]

Communication protocol:- A communication protocol is a system of rules that allow two or more entities of a communications system to transmit information via any kind of variation of a physical quantity. The protocol defines the rules, syntax, semantics and synchronization of communication and possible error recovery methods. [5]

First of all I will explain what is centralised model attempt. It is an attempt to minimise data by generating and keeping track of ephemeral identifiers distributed to users which can be used to construct the contact graph of a user only in the case they are infected. The generation of identifiers and generation of contact graphs are done on a server which is often assumed to be controlled by a government or another "trusted" entity. Any identifiers that an infected individual uploads to the system that he or she has observed can be resolved by the server into a persistent identifier that can be used to single-out an at-risk user. This model assumes that the entity running the server shall not misuse the data and capabilities of the server in cases beyond managing infection progression, for example, at the request of law enforcement, border control or intelligence agencies. Such protection relies on the protection of the central server which can potentially be re purposed into a 'data grab' model (i.e. a model that relies on a disproportionate collection of personal data in time of crisis, and assumes legal protections will be sufficient to protect populations which is often not the case).

Now I will explain what is Decentralised models . This a model that are designed to keep as much data on user devices as possible. Methods are introduced to strictly

control data flows in order to avoid accumulating any contact data on a centralised server. This means that a server exists but only to enable people to use their own devices to trace contacts. The server is not trusted with personally identifiable information at all, cannot use any identifiers to single out an individual, nor does it provide any individual with the identifiers they should broadcast, and therefore is much less vulnerable to function creep than all other solutions.

Now, DP3T has no central authority except for server that is essentially a shared memory of all users.Its essentially all upto the user to make proximity tracing work.It proposes a privacy-friendly, decentralized proximity tracing system that reveals minimal information to the backend server. It proposes three different protocols to support exposure detection and tracing. These protocols provide developers with choice regarding the trade-off between privacy and computation cost but share a common framework. In all three protocols, smartphones locally generate frequently-changing ephemeral identifiers (EphIDs) and broadcast them via Bluetooth Low Energy (BLE) beacons. Other smartphones observe these beacons and store them together with a time indication and measurements to estimate exposure (e.g., signal attenuations). The proximity tracing process is supported by a backend server that distributes anonymous exposure information to the app running on each phone.This backend server is trusted to not add information (i.e., to not add fake exposure events) nor remove information (i.e., to not remove exposure events) and to be available. The backend acts solely as a communication platform and does not perform any processing. It is untrusted with regards to protecting users' privacy. In other words, the privacy of the users in the system does not depend on the actions of this server. Even if the server is compromised or seized, their privacy remains intact. If patients are diagnosed with COVID-19, they will be authorized by health authorities to publish a protocol-specific representation of their EphIDs for the contagious period to aid in decentralized proximity tracing. We are aware that each country, and in some cases each country's regions, will have existing processes and systems in place to manage mass testing, to communicate between testing facilities and laboratories, and to inform patients. We further note that some implementations of the system might skip the authorisation step altogether and rely on self-reporting. However, it strongly advise implementing one of the proposed authorisation mechanisms to achieve stronger security guarantees. When authorized, users can instruct their phones to upload a representation of the EphIDs to the backend. The backend stores the uploaded representations. To protect COVID-positive users from network observers, all phones equipped with the app generate dummy traffic to provide plausible deniability of real uploads. Other smartphones periodically query the backend for information and reconstruct the corresponding EphIDs of COVID-19 positive users locally. If the smartphone has recorded beacons corresponding to any of the reported EphID s, then the smartphone's user might have been exposed to the virus. The smartphone uses the exposure measurements of the matched beacons to estimate the exposure of the phone's owner.

In approaches in which both identifier generation and COVID-exposure estimation are centralized, such as ROBERT a central server estimates a user's likelihood of COVID-exposure, instead of the user's smartphone in decentralized designs. Depending on the system, the server notifies the at-risk users or users query the server about their

status.
In all these systems, the central server holds a long-term pseudo-identifier for every user and uses it to derive ephemeral pseudo-identities (EphIDs) that are pushed to the smartphones.
In the ROBERT system, the backend tracks the exposure of each user of the system. COVID-19 diagnosed patients upload their observed EphIDs to the backend server. The backend server associates these observed EphIDs to long-term pseudo-identifiers of at-risk users. The backend uses the associated data to update the exposure for each of the at-risk users.Smartphones of ROBERT users regularly query the backend to request their exposure status. The backend answers whether the exposure passed the threshold or not.

Thus,its clearly states that DP3T is not centralised and Robert is a more centralised protocol.

3. 3

3.1. What are main privacy goals expressed for each protocol?

ROBERT

Data The protocol is supposed collect minimal data that is required by the proximity-tracing app to notify users that they have been in close proximity of COVID-19 virus carriers.
Tracking In particular, proximity tracing applications should not collect any geo-location data.The protocol should rely on the collection of temporary pseudonyms.
PrivacyThe Proximity tracing applications along with the protocol should not let any information leak to any 3rd party or government.In today's generation, information is the highest valued currency. This app+protocol should not get into money-making.
Spying The app and the protocol is supposed to help people and not to spy on them. There are several kinds of spying techniques available in regular apps. Such techniques should not be used with this protocol.Once this pandemic is over, users should be notified to delete this app.

DP3T

Data:The central server just understands the current number of potential Covid patient. The central server wouldn't know, "who", "from where", "gender".In this way, there is no way for the central server to abuse the data.The central server doesn't even notify the updated count all the time, to the health facilities. The health care facilities can just accept Covid potential patient, if patient walk-in.
Tracking There is no way to understand "who", "from where", "gender", etc, unless an user has pro-actively declared him/herself as Covid positive. Spying The protocol has no intention of spying, another
design-implementation would establish this fact. The proximity tracing application would be dismantle itself, triggered by the protocol.

3.2. How these two protocols fulfills these goals?

ROBERT

Identification:The authority receives some information only during the phase when

pseudonyms of contacts of a user diagnosed with COVID-19 are sent, and the authority receives only a user's pseudonym during the "exposure status request" phase. In addi- tion, only a temporary pseudonym is used, and the only information the authority can derive from such pseudonym is that the corresponding person has been exposed (i.e., in proximity to a user diagnosed with COVID-19). Therefore pseudonyms do not allow the authority to track users.The authority does not learn the real identities of any user, whether diagnosed with COVID-19 (i.e., tested positive), or exposed.

Location:Also, the authority cannot infer the "proximity graph" of user1, user2 or User3. If the authority wants to do physical tracking, it will need to deploy sniffing devices or compromise the user's mobile device by exploiting a vulnerability on his phone, which could lead to targeted surveillance but not mass surveillance. In addition, this kind of physical tracking is already feasible with other solutions such as video-surveillance, GSM boundary, or dedicated tracking devices.

Privacy of an individual:Knowing whether or not the person is diagnosed with COVID-19 would violate his/her privacy. The ROBERT protocol ensures that when a user tests positive, his/her application gets disabled, and the person can be asked – by her health professional (doctor/general practitioner) – to self-isolate for two weeks. When she recovers, her health professional will re-enable her application.

replay protection:Ephemeral pseudonyms are generated by the authority using a secure cryptographic function applied to a secret key associated with your application and timing information. These ephemeral pseudonyms are then sent securely to your application and stored for future usage. Only the authority and your application can access them. Ephemeral pseudonyms are then broadcast on Bluetooth to inform nearby applications. To limit the risk that the message containing your pseudonym be reused by a malicious user, this message is only valid during a few seconds

Authentication:The installation of the application involves a registration phase in which a secret key is generated and shared between the user's mobile device and the authority. The key is then used to generate ephemeral pseudonyms and to digitally sign the messages exchanged with the authority. Digital signatures are mathematical schemes used to verify the authenticity of messages. It gives the recipient a very strong reason to believe that the message was created by a known sender, and its content has not been altered in transit. Moreover, communications between the application and the authority are all encrypted.

DP3T

Data No entity can observe or keep record of a global view of the social graph of a population, in anonymized form or otherwise.The entities in the system receive the minimum amount of information tailored to their requirements. None can abuse the data for other purposes, nor can they be coerced or subpoenaed to make data available. Protects non-infected users No entity, including the backend server, can learn informa- tion from non-infected users.

Graceful dismantling The system will organically dismantle itself after the end of the epidemic. Infected patients will stop uploading their data to the central server, and people will stop using the app. Data on the server is removed after 14 days.

3.3. Point out the pros and the cons for each protocol?

Robert Pros: communications between the application and the authority are all encrypted.The type of instructions that will be transmitted as well as how often that in- formation will be shown on the user's mobile phone depend on the health professional and epidemiologists decisions.Users query the server to learn their exposure status. The identity of at-risk users is only protected when servers cannot de-anonymize users through their permanent app identifiers and network identifiers.
Cons: Fake exposure events. Triggering false alerts is easy in all centralised designs.Suppressing at-risk contacts. Hiding at-risk contacts is possible in any proximity tracing system. Prevent contact discovery. Any proximity tracing system based on Bluetooth BLE is susceptible to jamming attacks by active adversaries.

DPT3 Pros:Data use: Data collection and use limited to the purpose of the data collection: proximity tracing. This implies that the design should avoid collecting and using any data, for example geolocation data, that is not directly related to the task of detecting a close contact between two people. Controlled inference: Inferences about individuals and communities, such as information about social interactions or medical diagnosis, is controlled to avoid unintended information leakage. Each authorised entity should only be able to learn the information strictly necessary to fulfill its own requirements. Protect identities: The system collect, store, and use anonymous or pseudonymous data that is not directly linkable to an individual's identity where possible. Erasure: The system respect best practices in terms of data retention periods and delete any data that is not relevant.
Cons:A fake exposure event could make a person believe that they are at risk, even though they have never been exposed to a diagnosed user. Attackers could try to generate fake exposure events to trigger false alerts, e.g. by relaying or broadcasting EphIDs at large scale. This would violate the authenticity requirement of the system. There is a risk that either a COVID-19 positive user or the backend server could prevent other individuals from learning they are at risk, e.g., by modifying the app's local storage. This violates the integrity of the system and would lead to an increased health risk for at-risk individuals who rely on the system for alerts. A malicious actor could disrupt the system, e.g. by jamming Bluetooth signals, and prevent contact discovery.

4. Imagine you have to implement a contract tracing system using one of the two protocol

I would chose Robert protocol. I understand that DPT3 is better secured, but I believe that Robert protocol could be the best solution during such crisis period.

4.1. Describe the global infrastructure (with required equipment placed) of your solution (take into account all different users populations: general population, administration, medical teams)

- Let me consider myself to be App developer, so, my first consideration would be putting across the best design. Best design can be put across my choosing the right colours.There's a complete domain about Colour Science, maybe, I should consult a colour Science expert to do a proper design. There's a reason, that I am giving preference to colour and design, and my reason is to create a willingness by the user to use this app. The most popular apps, like Facebook, makes sure that the design is good. I used to be an addictive Facebook user. In order to lose this addiction, I moved to using Facebook in phone's browser which would help me interact with ugly interface thereby helping me to reduce the Facebook time. A couple of months ago, one fine day, I found out that the browser experience has become exactly the same as the app. Based on this, it is very important to note that designs matter. Just another point, the application should be very easy to use. People don't like writing a lot of information, specially elderly people, so I would prefer letting users allow the app to take information from social-ID. Now, if I have to talk about below the poverty-line users, illegal immigrants who have just been introduced to smart phones during this pandemic, then they need to type manually in their preferred language. So, my app should support multiple language because poverty doesn't have a cast,creed,language or religion. This concludes the first page of the app.
- The second page of the app would be filled with the medical details the user has, and this information would be derived from the social-ID which is linked to medical profile, most of the time. This page would also have more details with the title, How are you feeling today.. We would try to include check boxes in regards to the Covid-19 symptoms like the follows

[6]

- The idea is not to make the user, panic, about the symptoms and efforts would be put to ease on the questions. Several testing would be done before the app release. Moreover, based on the user information, he/she would be notified if that person needs to visit a testing center, stay in home quarantine for some number of days. Efforts would be also put, to display the current appointment status of the testing centers and facities would be provided to add an appointment for self.
- The third page of the app, should have a simple interface, where government authorised body can post information about the advantages of using natural ingredients during this pandemic. We have to make sure, that this interface is free from any kind of advertisement.This app should be non-profit app. There would be other information like Total infected count, Recovered count, deceased count in regards to area of coverage starting from 2 kms to the end, as much as possible.
- There would be a fourth page, where the user needs to go to before getting out of the house. In this page, camera activation facility to make payments, storing-

deleting options in regards to keeping credit-card information for contact-less pay- ments to shops, train stations and so on.

- All the user details would be stored in a central server, which would not be accessible by ay government bodies. The user details would be deleted from the central server after 30 days.
- There would be a fifth page, where the user would be notified if he/she has come in contact with a Covid 19 and would be suggested to go to a testing center. This page would also contain the option to self proclaim being Covid-19 positive. This page would also contain the option to facilitate priority based appointment for such users.
- This part was all about the user experience.
- About the technical part, the app should be compatible with ios, android, windows and linux based phones. Ios, android phones captures user-location irrespective of user's knowledge. The idea would be to use these features and also notify the users about the vulnerabilities. Some techies find ways to stop the location services in their phone, sosuch techies should be notified that this app works best with location services, so they can choose to switch on the location services.

4.2. What are the main security points you should take into account in your solution and explain how you solve it.

- As a contact-tracing app developer, I have chosen the Robert Protocol.
- I would set the settings of the Robert Protocol to standards.
- I completely understand the vulnerabilities. A regular phone is made vulnerable by the phone manufacturer, network provider and theapps. So, if we understand, the biggest vulnerability is choosing to use a smart phone.
- A contact app designer doesn't have any self-ulterior motive.
- Anyway,the backend design would be put in place to support the standard settings of the Robert protocol.
- So, we have a concept of introducing premium versions of same apps. Given the opportunity, I would design the premium version of this app with the same facilities but with DPT3 protocol. The premium version would be available only to the army personnels. Right now, I understand that a single app cannot be designed for different security protocols. So, I would strongly recommend that my app with Robert Protocol is not used by army personnels.
- My points are:
  - Data: All the data of the user would be in the device itself. The cypher text version of the data would be stored in the central server. Absolute efforts would be given to prevent any data-leak or eaves-dropping by any un-authorised personnel or government bodies.Moreover, the data would be also deleted from the user's phone and the central server at the end of 30 days, and it would be made mandatory to fill in the same details as before by the user.
  - Trust and Privacy: This app would not be able to gather any information from other apps. However, text message services can be accessed only with user permission.

5. In the two protocols, a server is required at exposure verification and infection declaration. Imagine you are the owner (administrator) of the server and be able to observe all communications with your server at any time. Imagine you are a telecom operator and be able to to observe all communications, you are also able to localize cell phone

5.1. in these two situations, explain how the flows could be exploited, what the attacker could learn and what protocols' properties cannot be respected

- Location data: The decentralized design prevents location tracking of Covid-19 postive patients during their contagious period.In centralised systems, keys are genenerated on the server. Access to these server side keys, enables linking EphIDs to app indentifier which in turn facilitates the user identification and enables tracking their further movements.
- COVID-19 positive: The centralised and decentralised systems have the same privacy limitation which could be exploited to get the infected user list. The centralised system hides Covid-19 positive user's identity well, thus EphIDs cannot be used to pin-point Covid positive user. However, network eaves- dropping and centralised system's dummy traffic bypassing can be used.
- data collection: Bluetooth data could reveal family,community information.
- Fake alerts: In centralised system, a user can trigger fake Covid alert
- Identification protection: Covid 19 patient's ,thereby at-risk user's identity can be hidden in proximity-tracing system.

References

[1] https://en.wikipedia.org/wiki/Decentralized$_{P}$rivacy − P reserving$_{P}$roximity$_{T}$racing.
[2] https://searchnetworking.techtarget.com/answer/What-is-open-and-clos
[3] https://www.chainbits.com/cryptocurrency-terms/decentralized-protoco [4] https://www.inria.fr/sites/default/files/2020-04/ROBERT-infography-E [5] https://en.wikipedia.org/wiki/Communication$_{p}$rotocol
[6] https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.