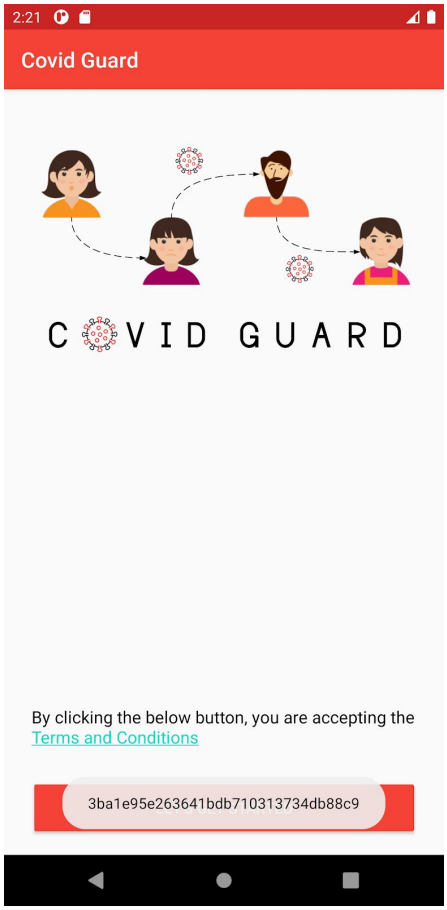
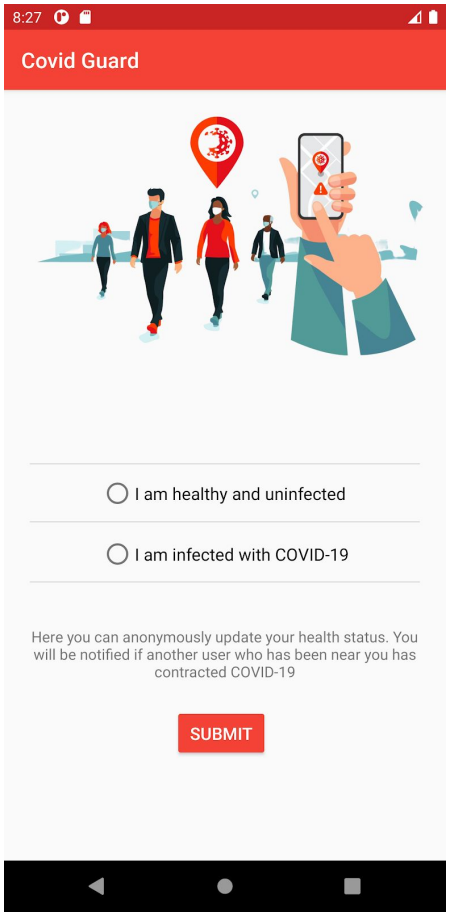
Weekly Blog 3

Apart from the three main papers which I have read (listed below), I developed two of our app screens on Android Studio. The screenshots can be seen below:





Things I learned:

1. I learned how activities are managed in Android Studio. For example, the registration screen is an activity which can either be made through XML coding or through the UI provided by Android Studio.
2. I learned about the lifecycle of an app. Namely:
    a. onStart(): when the app is launched
    b. onPause(): when the app is running in the background
    c. onResume(): when the app is in use.
    d. onStop(): when the app is exited using the back button.
    e. onDestroy(): when the app is killed after onStop() or due to less memory.

3. I learned about how to use onClick() functions as part of an event. For instance, what should happen during a button click.
4. I learned how to place images from the local through uploading it in the drawable folder.
5. I learned about the various types of layouts such as Constraint layout, linear layout to name a few.

Contact Tracing Mobile Apps for COVID-19: Privacy Considerations and Related Trade-offs

1. What research question does the paper address?

   The paper mainly tries to address the privacy from three perspectives, from snoopers, from authorities, and contacts.

   a. Privacy from snoopers: In this scenario, the authors compare the privacy concerns of sending out the name and phone numbers of people who have been affected by COVID. The authors examine this approach with that of the Singaporean app, “TraceTogether.” In TraceTogether, the user’s information is anonymised using a unique temporary ID, which is refreshed now and then.

      In our project, we similarly use an approach where the user’s identity is not comprised. Instead, a unique temporary identifier known as the Rolling Proximity Identifier (RPI) is used for the matchmaking process. Therefore, we believe there exists privacy from snoopers in this regard.

   b. Privacy from authorities: In this regard, the app TraceTogether makes use of a centralised approach for performing the contact tracing. More specifically, the data of both the diagnosed and non-diagnosed users are uploaded to a central server, presumably hosted by a government agency. However, the government can obtain the identity of users through this approach.

      In our project, we have the advantage that the architecture that’s followed is a decentralised one. More specifically, the diagnosed users are the only people who upload their RPIs to the server. On the other hand, the diagnosed users never upload their RPIs to the server. Instead, they download the data from the server to see if they have come in contact with a COVID positive contact.

c. Privacy form contacts: Again, as mentioned before, a centralised approach is followed by the TraceTogether app of Singapore. Here, the data and match-making are performed in the central server. The downside here is the faith that needs to be put over the government.

   In our approach, we have the upper hand that the RPIs of users are regularly refreshed, and a match is only made upon the diagnosed user's data upload to the server. More specifically, the users are only notified they have come in contact with an infected person and not informed of the latter's identity to the user.

2. What was the conclusion of the research?

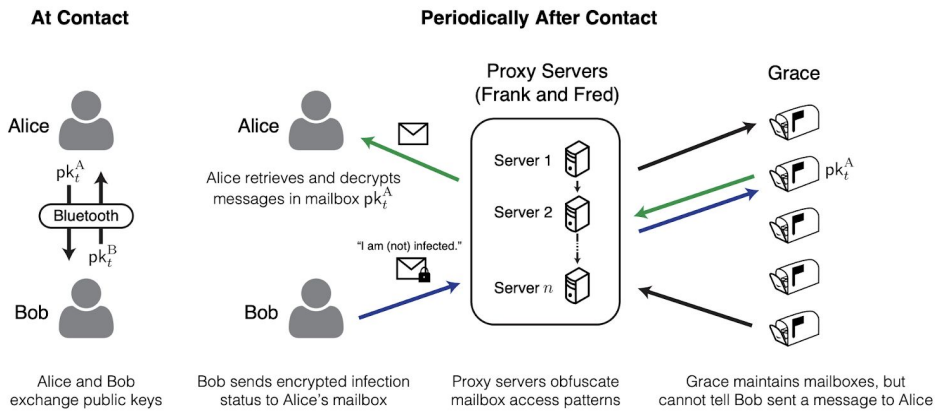
   The conclusion of the research is a couple of ways to enhance the privacy of the TraceTogether app. Here we compare if the approaches are better than or equivalent to what our project's security measures.

   a. Partial anonymising using polling: The authors suggest using polling the data of the central server by non-diagnosed users to see if they have come in contact with infected users. This system does not ask for mobile numbers, unlike in TraceTogether. Again, our approach does not ask for mobile numbers but instead sends out an exposure notification indicator without revealing the data of diagnosed and non-diagnosed users.

   b. Public Database of Infected Users' Tokens is Efficient but Less Private: This approach is similar to the decentralised approach we have considered for our project. Here, the central server consists of only the diagnosis keys of infected individuals. It guarantees secure privacy to non-diagnosed individuals as their data never leaves the phone, and all the match-making happens in the user's phone itself. As such, this approach is adopted in our project.

      The only downside to the approach is the amount of data that could be present if the pandemic reaches its peak. However, in our solution, individuals' data are deleted after 14 days, which is the average time frame for individuals to recover from COVID. Furthermore, since the RPIs does not contain any identity-related information, the infected user's identity is safe.

c. Privacy from Authorities based on Private Messaging Systems:



This approach follows something similar to the TOR approach. The data from the sender is encrypted at multiple servers decoupling the sender from its destination. As such, at each server, the outermost layer is peeled off. Only the recipient can decrypt the message with the public key of the sender, which was exchanged at the encounter. This guarantees additional privacy to both diagnosed and non-diagnosed users. The downside of this approach is its computational feasibility and scalability.

3. How can you apply this knowledge to your project?

As mentioned amongst the various questions above, we can adopt the “Public Database of Infected Users’ Tokens is an Efficient but Less Private” approach to our project using GAEN.

Privacy Risk and Preservation For COVID-19 Contact Tracing Apps

1. What research question does the paper address?

   The paper mainly addresses the privacy risks, the technologies used, and the data collection of various apps in different geographics. Some of these differentiations include GPS, Bluetooth, and GPS + Bluetooth. Some of the approaches include centralised and decentralised approaches.

   Some of the apps like ProteGO and SafePaths use a centralised approach which involves uploading entire data to central servers for mass surveillance. COVIDSafe uses a decentralised approach which uses only data from diagnosed users, guaranteeing extra protection to non-diagnosed users.

   In our project, we do not use location data. Instead, we use Bluetooth Low Energy for proximity detection and a decentralised approach as well. Furthermore, there is more accuracy in our approach from the perspective that false positives are mostly avoided. This is on the basis that Bluetooth signals do not pass through walls. As such, an event won't be concluded as a contact event when the users are actually separated by a wall. In other words, false positives are avoided. GPS-based approaches tend to produce more numbers of false positives.

   Furthermore, the privacy of the users is enhanced through the random refresh of unique identifiers which could prevent attacks like linkage attacks. The paper addresses the advantage offered by GAEN of the fact that they offer cryptographic solutions. Since we are adopting GAEN, we are able to make use of the advantage.

2. What was the conclusion of the research?

   The primary conclusion of the research is how privacy plays an important role. Some of the phrases used in the paper such as "*Minimizing data collection and limiting access and retaining data only for the minimum amount of time that is necessary*" are well applied in our project as well. More specifically, the data collected in the diagnosis key server is deleted every 14 days with respect to the user. Data is not collected from non-diagnosed users. "*Obtaining consents is also commonly used for privacy protection".* In our project, the user has the feature to consensually upload his/her data upon diagnosis.

   Overall, we believe the decentralized architecture with GAEN approach is very feasible in terms of privacy preservation. GAEN is constantly audited and updated by Google as such, we are able to use it for making a better app.

3. How can you apply this knowledge to your project?

   The overall research revolving around the contact tracing apps has helped us in choosing to validate why a decentralized architecture for our project is best suited. Furthermore, we are able to provide validation against choosing GAEN and Bluetooth as part of the contact tracing framework.

COVID-19 Contact Tracing: Eight Privacy Questions Explored

2. What research question does the paper address?

   The research paper attempts to answer 8 privacy questions based on 3 contact tracing approaches, namely: Using location, using Bluetooth, Using Bluetooth with changing identifiers.

   In this section, we will be applying the answers to our own project to see if it satisfies the requirements expected by the questions.

   a. How do you limit the personal data gathered by the authority?

      In the location protocol, the answer to this question is not satisfied as the trajectories of the users are directly used by the authorities to draw the social graphs.

      In Bluetooth protocol, the identifiers generated by users are fixed which could lead to privacy concerns later on.

      In Bluetooth with changing identifiers protocol however, this question is satisfactory as the identities are not shared with the authorities. Furthermore, non-diagnosed user privacy is more enhanced as their data never leaves the device.

      In our project, we are using the third protocol and as such it satisfies the question.

   b. How do you protect the anonymity of every user?

      In the location protocol, the anonymity of the users is compromised as there involves direct trajectory transmission to the authority.

      In the two Bluetooth protocols, the anonymity is maintained in the form of temporary identifiers especially in the case of the protocol with changing temporary identifiers.

      In our project, we are using the third protocol which protects the anonymity of the users by changing identifiers.

c. Does your system reveal to the authority the identity of users who are at risk?

   No, the system does not reveal to the authority which users are specifically at risk. It uses a decentralized approach. In other words, users who get notified of their exposure can voluntarily go to a government agency to report their contraction of COVID-19.

d. Could your system be used by users to learn who is infected or at risk, even in their social circle?

   Since our system is using a Bluetooth based contact retracement using constantly changing identifiers, the system cannot be used by adversaries to learn about who they have been in contact. Instead, they'll only be notified that they have been exposed.

e. Does your system allow users to learn any personal information about other users?

   No, our system does not allow users to learn about other users. It merely acts as an exposure notification indicator to notify users of exposure to COVID-19 positive users

f. Could external parties exploit your system to track users or infer whether they are infected?

   Similar to question d.

g. Do you put in place additional measures to protect the personal data of infected and at-risk users?

   Some of the measures that our project has are:
   i. Decentralized
   ii. Encrypted local data
   iii. Changing temporary keys for the prevention of tracing
   iv. How can we verify that the system does what it says?

3. What is the conclusion of the approach?

The conclusion of the approach is the different questions was analysed with respect to the three technologies location, Bluetooth and Bluetooth with changing identifiers. I have, however, analysed the questions with respect to my project and answered the questions accordingly.

4. How can we apply to our project?

   As explained in the above questions, I was able to validate the various questions with respect to my project.

References

1. Cho, H., Ippolito, D., and Yu, Y.W., 2020. Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs. arXiv preprint arXiv:2003.11511.
2. Wang, D., and Liu, F., 2020. Privacy Risk and Preservation For COVID-19 Contact Tracing Apps. arXiv preprint arXiv:2006.15433.
3. Lawson-Tancred, H., Price, H.C. and Provetti, A., 2020. COVID-19 Contact Tracing: Eight Privacy Questions Explored. arXiv preprint arXiv:2005.11416.