Security & Privacy Analysis of Apple&Google's Contact Tracing

CS161: Computer Security

Prof. Raluca Ada Popa

April 24, 2020

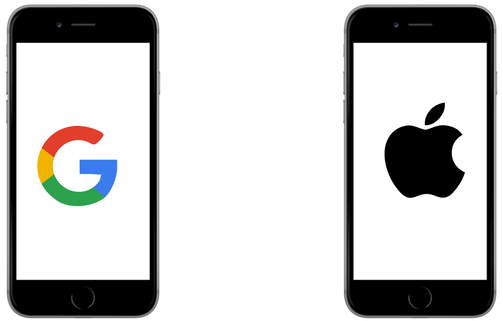
Some content taken from: https://www.blog.google/documents/57/Overview_of_COVID-19_Contact_Tracing_Using_BLE.pdf

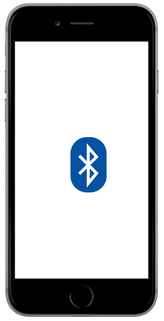
Announcements

- Starting recording
- Project 3 part 2 is released and will be due on **May 3 at 11:59 pm**

Apple & Google’s contact tracing protocol

- The two companies teamed to create a decentralized contact tracing tool using which users can determine if they were exposed to COVID-19 with privacy and security considerations at its core
- Why the two companies in particular?





Uses Bluetooth technology

- Because COVID-19 can be transmitted through close proximity
- Bluetooth range: ~33 feet/10 meters
- If someone is in your Bluetooth range, there could have been a contact
- This algorithms aims to trace contact

Privacy and security are core to the algorithm

Why?

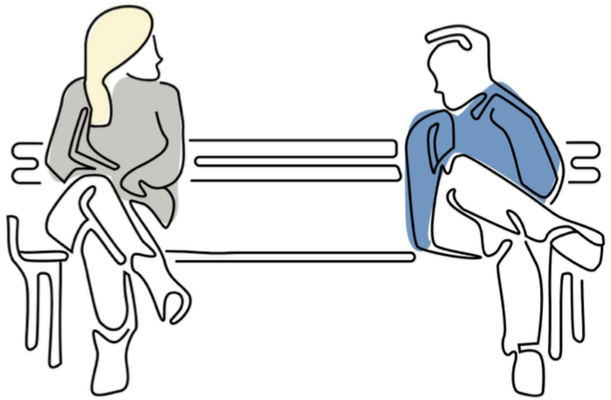
- very private information such as location/time/people met
- mass privacy invasion has affected company's reputation (e.g., Facebook)
- malicious users could try to affect the tracing

User consent

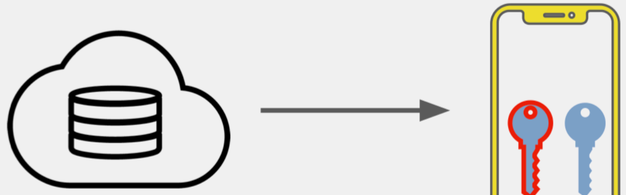
- Opt-in to install app
- Opt-in to declare if diagnosed with COVID-19

Workflow

Alice and Bob meet each other for the first time and have a 10-minute conversation.



Alice's phone periodically downloads the broadcast beacon keys of everyone who has tested positive for COVID-19 in her region. A match is found with the Bob's anonymous identifier beacons.



Anonymous identifier keys are downloaded periodically

A match is found

Alice sees a notification on her phone.

Alice's phone receives a notification with information about what to do next.

Additional information is provided by the health authority app or website

Google

The cryptographic protocol
running on each user's phone

Setup (once):
- Generate a random Tracing Key **tk**
  - Never leaves device

Every 24h period D_i:
- Generate Daily Tracing Key **dtk_i**

  dtk_i ← HKDF(**tk**, "CT-DTK" || D_i)

(HKDF = key derivation function based on HMAC)

Every 10 minute epoch TIN_j:
- Generate a rolling identifier

  $RPI_{i,j}$ ← HMAC(**dtk_i**, "CT-RPI" || TIN_j)

- Transmit **$RPI_{i,j}$** via Bluetooth to all phones nearby

Receive:
- For every advertisement reception, store (**$RPI_{i,j}$,** D_i) pairs locally.

If user is diagnosed:
- Release (D_i ,**dtk_i**) of this user for the last some-number of days (called diagnosis keys)

Diagnosis server

- Aggregates all diagnosis keys for the past N days
- Serves them to each user downloading periodically
- User identity and contact information is not uploaded to the server operator -- contact tracing is performed entirely locally

The cryptographic protocol
running on each user's phone

Setup (once):
- Generate a random Tracing Key **tk**
  - Never leaves device

Every 24h period D_i:
- Generate Daily Tracing Key $\mathbf{dtk_i}$

$\mathbf{dtk_i} \leftarrow$ HKDF(**tk**, "CT-DTK" || D_i)

(HKDF = key derivation function based on HMAC)

Every 10 minute epoch TIN_j:
- Generate a rolling identifier

$\mathbf{RPI_{i,j}} \leftarrow$ HMAC($\mathbf{dtk_i}$, "CT-RPI" || TIN_j)

- Transmit $\mathbf{RPI_{i,j}}$ via Bluetooth to all phones nearby

(Cannot be linked with each other)

Receive:
- For every advertisement reception, store ($\mathbf{RPI_{i,j}}$, D_i) pairs locally.

If user is diagnosed:
- Send (D_i ,$\mathbf{dtk_i}$) for the last some-number of days (called diagnosis keys) to the server

Periodically:
- Download all new diagnosis keys (D_i ,$\mathbf{dtk_i}$)
- Generate every related rolling identifier and check against stored advertised pairs.
- If a match is found, you've been in contact with an COVID-19 patient.

Security analysis

Privacy analysis

What sensitive information should we worry about in this app?

- Time-Location samples of each user
- Can identify who the user is and where-when they have been and with whom they came in contact

Privacy analysis

What private information do users see?

- Distinction between what a client app can see and what the user sees on the screen from an honest app
- For user Bob who declared COVID: Alice's client could figure out who the user was and where she met him. If a few users come together who were around Bob, they could reconstruct all the time-location path of Bob. Basically, you should assume that you have no privacy if you declare you have COVID.
- For users who did not declare COVID, you could track the user for 10 minutes using the RPI, or more if more clients collude or if more contact

Privacy analysis

What private information does the server see?

- For users who declared COVID: their rolling identifiers. Put together with location data (e.g., from some users) it can identify the individuals.

- Less for non-diagnosed users: number of users, when they check for updates. Any information received from users colluding with server.

Privacy analysis

What other attacks could there be?

- Attackers install recording devices in many places. Reconstruct identity and path of users who declared COVID.
- others

Consequences of no privacy for opt-in diagnosed users

Users might be afraid to declare they have COVID because:

- people might mistreat them (including violence cases)
- someone who contracted from this user could forever hold grudge
- others

Integrity analysis

Can a malicious server affect the correctness of the tracing?

- Yes, entirely. This protocol trusts the server for integrity

What can malicious users do?

- Create false positives: upload fake “COVID” diagnosis and create panic; broadcast their RPI ids in many places in the world by replaying it there to create a lot of contact;
- Cannot prevent honest user with COVID to upload their own diagnosis unless the attacker can jam the network for that user or receiving users

In summary

- Google and Apple's contact tracing protocol via Bluetooth aims to inform users if they were exposed to other users
- Lots of privacy and security concerns:
  - for privacy, in a nutshell, users without COVID-19 have some degree of privacy, those with COVID-19 do not have much privacy all
- The privacy and security concerns are likely to hamper adoption