The New York Academy of Sciences

A Novel Approach To Combat A Novel Virus – Global Solution For Contact Tracing Using Bluetooth And Intelligent Analytics

Combating COVID-19

The New York Academy of Sciences

TEAM: Ishika Gupta, Krish Gangaraju, Sana Mohammed, Sreshta Pothula, Ved Srivathsa

(TISB, Bangalore, India)

Short Video Demo of Our End-to-End Solution

Mouse over on to Video Object to see play button

If any issues with the video above, please click the link below:
https://youtu.be/jWSwdXZg17U

Introduction & Literature Survey

- From past pandemics [19][20] we can see that social distancing was an effective strategy. Previously, it was easier to implement social distancing because the movement of people was geographically restricted due to poor transportation. However in the 21st century, due to advancements in transportation, and for socio-economic reasons, travel across geographic boundaries has become a norm. The entire globe can be circled in 24 hours and thus COVID-19 has been able to spread to various corners of the world within 60 days.
- In the absence of a vaccination for COVID-19, enforcing social distancing is the only option governments have.
- Most countries have enforced social distancing and shut borders to relieve stress on the healthcare system. This helps contain the virus and buys time to understand its spreading patterns. Therefore, most countries are rushing to come up with a digital solution, primarily focusing on contact tracing to Identify, Isolate, and Treat.
- Few Asian & Middle Eastern countries [9][10][11][14][16] gave less significance to privacy while enforcing contact tracing. However, the USA and Europe have come up with their own solutions keeping 'privacy' in mind. Here is the summary of 'approaches' taken by various countries (as of 30-APR-2020)

Country	Date of News	Approach	Mechanism for Contact Exchange	Mechanism to report Infected Patient	Mechanism for Sending Alerts	Command Centre Involved	Technology
UK [16]	March 19, 2020	Centralized	IDs/Keys exchanged when two mobile phones with Bluetooth enabled spend some time together	Command Centre's Responsibility	Unknown – details are still being worked out.	Yes, NHS	Mobile Devices with Bluetooth
Singapore [7]	March 21, 2020	Centralized	IDs/Keys exchanged when two mobile phones with Bluetooth enabled cross each other	Command Centre	Command centre looks at app **logs** and decides whom to send alerts to	Yes	Mobile Devices with Bluetooth
US (Apple & Google) [3][4]	April 10, 2020	Decentralized	IDs/Keys exchanged when two mobile phones with Bluetooth enabled cross each other	Unknown	IDs/Keys exchanges are stored locally on phone, based on match of published infected key with keys stored **locally** alerts are generated	Unknown	Mobile Devices with Bluetooth
Europe & Australia (PEPP-PT) [5][13]	April 17, 2020	Centralized	DP-3T Protocol; IDs/Keys exchanged when two mobile phones with Bluetooth enabled spend 15 minutes together	Command Centre's Responsibility	IDs/Keys exchanges are stored locally on phone, based on match of published infected key with keys stored **centrally** alerts are generated	Yes	Mobile Devices with Bluetooth
Switzerland [6]	April 21, 2020	Decentralized	DP-3T Protocol; IDs/Keys exchanged when two mobile phones with Bluetooth enabled spend some time together	Unknown	IDs/Keys exchanges are stored locally on phone, based on match of published infected key with keys stored **locally** alerts are generated	Unknown	Mobile Devices with Bluetooth

Issues with Current Solutions

Issues with Solutions Proposed by Several Countries

The proposed solutions to date are available in the form of top-level concepts or ideas with lots of unknowns and many unanswered questions. We don't have any insight or access to their implementations and/or any other technical details. In addition to this, some proposed solutions heavily focus on privacy while others don't. Some of the issues with proposed solutions are listed below:

1. **Non-Mobile Devices:** The above solutions work only with mobile phones but don't address the following scenarios:[3][4][5][6][12][15]

1. - If a person jogging in the park comes in contact with, or is near an infected person, and doesn't have a phone, but is wearing a smartwatch

2. - If a person visits an infected local newspaper agent without his/her cell phone, but with car/house keys which have a key finder Bluetooth device attached

3. - If a student in his/her local library has Bluetooth headphones and comes in contact with an infected person

4. - If an individual is in a plane or museum where phones are required to be switched off or in flight mode but people have wearable devices (FitBit/pedometers/health monitors/headsets etc.) on

2. **Indirect Contact Chain:** The above solutions send alerts if a person is in direct contact with an infected person (i.e. direct contact chain), but don't consider the indirect contact chain (i.e. person A is in contact with person B and B with C). [3][4][5][6][12][15]
3. **Country-Specific:** The solutions are country-specific instead of taking a global perspective. Some solutions are centralized while others are decentralized. Moreover, users would have to download country-specific apps to receive an infected person's IDs/Keys for the app to work globally.
4. **RSSI Unreliability:** None of the above solutions take into account details regarding the unreliability of RSSI signals for distance measurements.[1][2]
5. **Global Issues & Softer Local Solutions:** With local or country-specific solutions:

(a) expertise is spread locally, each looking at country-level data.
(b) There is potential to miss global disease spreading patterns.
(c) Weaknesses begin appearing in local solutions if countries start cutting corners for financial reasons.

Our solution:

Addressing the above issues, we propose a complete end-to-end solution that includes: conceptual architecture, a mobile app, hardware simulations, analytical models, and data visualizations for effective contact network tracing.

- A centralized solution that works with all Bluetooth devices
- Utilizes GPS and Bluetooth technology
- Proposing an analytical solution addressing the RSSI unreliability with potential to enhance the model further by considering other sensors available in phones (gyroscope, accelerometer, compass, pressure, etc.)
- With the power of analytics, we show the effectiveness of our solution and how false alerts can be minimized.
- With a global solution:
  (i) all expertise is in one place,
  (ii) researchers would have access to global data to study, both for current, and all future pandemics
  (iii) there is no risk of countries taking relaxed approaches and undermining the global effort in containing the disease,
  (iv) if new disease spreading patterns emerge, all tools are available for a swift response.

Our philosophy is – viruses do not have boundaries, so solutions must be global and can be managed centrally by trusted organizations like WHO, UN, etc.

Key Takeaway from this slide

1. Weaknesses in the solutions proposed by various countries.
2. Highlights why pandemics require global solutions.
3. A centralized solution is proposed while addressing all issues.
4. The potential to apply and learn local and global disease spreading patterns with centralized data.
5. Since it is a global solution with centralized data, the world would be ready to face any future health crises. Pandemics cannot be fought locally

Our Solution – Concept & Architecture

- Download App and get Unique ID
- Option to register additional personal devices using the same Unique ID

STEP 1

At regular frequency, the app (aka Source) scans for nearby Bluetooth devices (aka Sink). If any Sink devices are registered (based on match of device Hash token) then it writes all Sink devices it found to the Central Database. The attributes are: "Datetime/Source ID/Sink ID/GPS/RSSI"

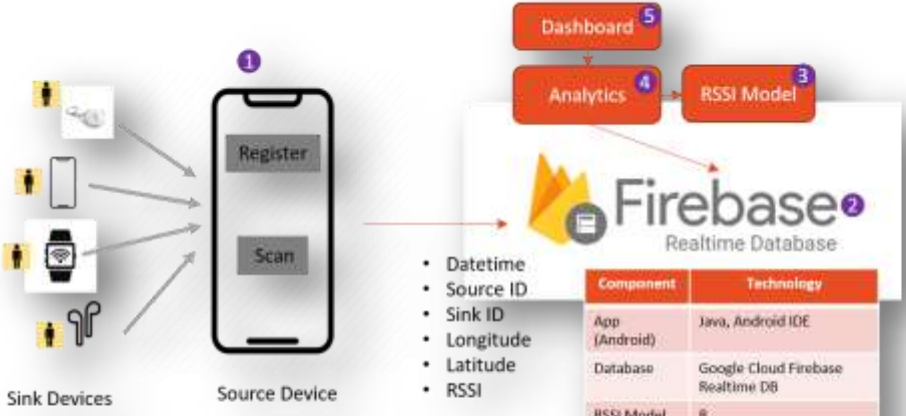
STEP 2

Hospital or Command Centre Declares Person X as Covid-19 on Day-Y

STEP 3

The Command Centre equipped with Analytics built on proximity data and Network Graph Visualisation, it can alert all people who were in the Contact Network Chain. The Analytics provide rich and deeper contextual information such as:
(a) Was the Source in crowded or non-crowded places?
(b) What was the Source state (walking or driving or public transportation) and
(c) much more..

STEP 4



Component	Technology
App (Android)	Java, Android IDE
Database	Google Cloud Firebase Realtime DB
RSSI Model	R
Analytics	Python & R
Dashboard	R Shiny App

Terminology

Source: Source is a device that scans for other Bluetooth devices.
Sink: the Bluetooth device which was found by the Source Bluetooth device during the proximity scan
Devices: People, person are used interchangeably with devices
Patient: Patient and Source are used interchangeably

Key Takeaway from this slide

1. A centralized architecture is proposed.
2. Effectiveness is dependent on the user registering all of his/her Bluetooth devices and giving permission to store the device and the data required, centrally.
3. Visualization dashboard is built to demonstrate the deep insights we can derive from the data.

The architecture comprises of the following 5 components:

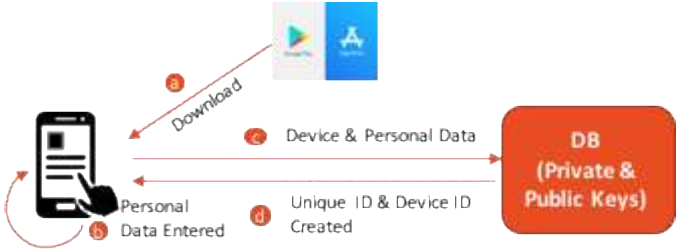
1. Mobile App: A Java-based android app for Registration and Gathering Proximity Data
2. Database: We used Google Cloud Firebase Realtime Database to store the registration and proximity data
3. RSSI Model: We built our own RSSI model to calculate the distance from RSSI Signal values
4. Analytics: The analytics creates new derived measures from the raw proximity data.
- GPS Distance Travelled: between two consecutive proximity scans
- Time Spent Duration: between source/sink devices
- Distance: between two source/sink devices
- Number of Devices: around the source at each proximity scan

5. Dashboard: Visualization platform to derive contextual insights and initiate

Step 1– Registration and Privacy

1. New User Registration (one-off)

A



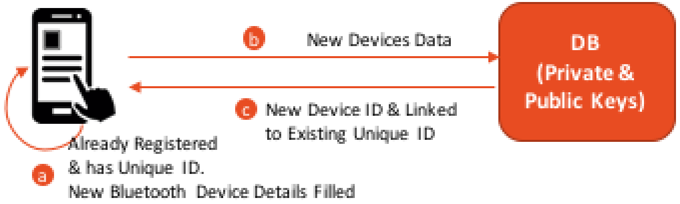
New User Registration Process Consists of Entering:

B

- a User downloads app
- b Name & contact details are provided
- c The data is encrypted and persisted centrally, with access to the command centre
- d Unique ID is created based on a message digest of the device and personal data

2. Option to Register Additional Devices using the Same ID

C



Data Privacy

E

- If one looks at the statistics of coronavirus [20], it is the West who suffered more compared to the East in terms of deaths, economy, loss of jobs, and now several businesses are on brink of collapse. It is not because lockdown was more efficient in the East, but because they resorted to digital tracking which cannot be implemented in the West. This raises a fundamental question: is one's privacy more important than other people's lives?
- In addition, due to liberty laws and privacy rules, western countries took a softer approach [3][4][5][6][13][15], (voluntary apps) pinning the hope that people would come forward out of civic duty to share their location information by installing the digital tracing apps.
- With COVID-19, the culture of the people is certainly changing. For example, in early stages of coronavirus spread, experts from France [16] looked at the digital tracking tools but the interior ministers initially rejected such moves citing threats to "individual liberties" But three weeks later and after a huge spike in deaths, the French government is now considering using a smartphone tracking app called StopCovid. Such a change in Europe was unimaginable which has very tough privacy rules.
- The softer and local solutions are not an answer to global pandemic issues. No matter how well a country is prepared to fight pandemics, if other countries are weaker, then viruses would find a way to spread across the globe.
- The COVID-19 is not going to be the last virus we will encounter. History shows [18], new viruses have broken out in the past and it will in the future. To fight the global pandemic issues the world needs to work together and find global solutions.
- The questions of ethics and data privacy are not new questions [14][17], these are age-old questions with no easy answers to it. What people would like to hear from the policymakers is that the data collected is purely used for disease control and not for anything else.

Registration of Additional Devices:

D

- a Inputs new additional Bluetooth device details
- b Device details are persisted under the same unique ID
- c Acknowledgement returned to the user

Key Takeaway from this slide

1. How to Register
2. Privacy Issues
3. Until policymakers make a commitment that any collected location data is purely used for fighting the spread of disease, no solution would work.

Step 2 - Bluetooth Proximity Scan – Experimental Setup 1

How was the data generated?:

- Using the mobile app we created, we collected the proximity data by walking for an ho A for a few days. However, due to strict lockdowns, we were unable to generate a lot of data.
- The proximity scan was performed at regular and differing frequencies, to have variation in collected data
- During this simulated walk:

(a) we went into crowded places like coffee shops, supermarkets
(b) part of the journey was by walk and part of the journey was by Bus or Tram.

Visualizations of Raw Data: The schematic diagram of the process describing the data along with the GPS path is shown here.

B

In the Contact Network Graph shown on the right, each node represents a Bluetooth device/person.
This Network Graph highlights the source that has been in proximity to several Bluetooth C devices/ people in a short amount of time.

These raw visualization plots highlight one issue: if a person in the contact network gets tested positive, **would you send alerts to every single person who was in the contact network?**
If we send alerts,

E

- The person who is in the proximity of 1-2 meters and a person who is far away (10-100m) also receives the alert. **Is this a true alert?**
- The passer-by who is cycling would also receive an alert even though he/she was far away and did not spend more than a few seconds with the infected person. **Is this a true alert?**

Therefore, we need a filter to narrow down the contact network to people who were in very close proximity for enough duration with the infected person (distance/duration as suggested by health authorities).

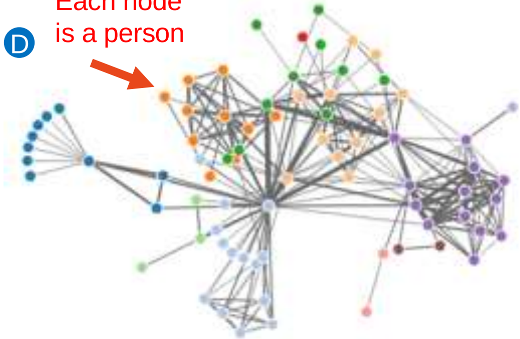
The answer to this filter is our analytical solution: RSSI Model for Proximity.

F

The GPS maps are useful because they give context as to where the COVID-19 patient has been in the past.

D

Each node
is a person



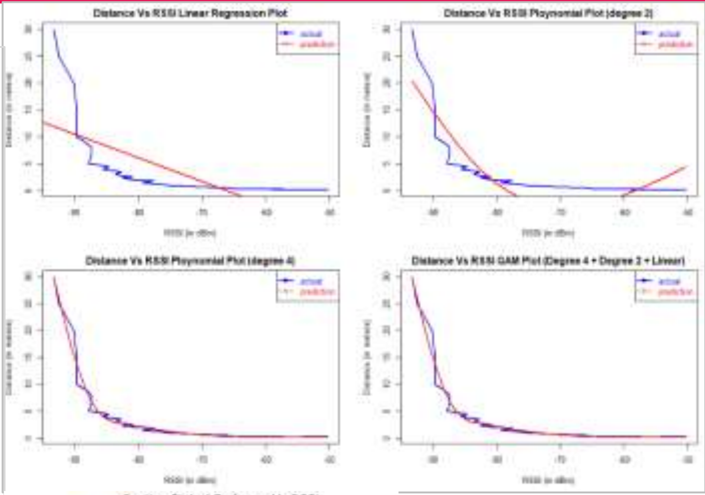
Key Takeaway from this slide

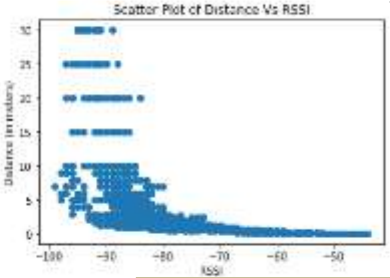
1. Raw proximity data gives a good overview but not useful enough as there is a lot of room for sending false alerts.

Step 2– RSSI Model for Proximity – Experimental Setup 2

RSSI Mathematics and Unreliability Issues

- The RSSI value represents the Bluetooth signal strength between two devices. A strong signal shows close proximity.
- The mathematical relationship between RSSI and distance ‘d’ between the two Bluetooth devices is: “$RSSI = -(10*n) * \log_{10}(d) - A$”
  In this equation: ‘n’ is the signal propagation constant, ‘A’ is a reference received signal strength in dBm, ‘d’ is the distance between two Bluetooth devices and RSSI is the received signal strength indicator in dBm.
- The challenge is that, in the above equation, there are three unknowns and we only know one value.
- Another big challenge is that RSSI is **highly unreliable [1][2]** in accurately measuring the distance between two people/devices resulting in the risk of false proximity alerts. It is **device-dependent** (Android & Apple phones give different measurements for the same distance). **The way the phone is held, the amount of open space around, the presence of walls/blockages between people, signals going through the body, etc. are all disregarded.**
- Addressing these challenges requires more data. In this paper, as a first step, we are proposing an **analytical solution** addressing the **device/manufacturer dependency** issues. In the future, we would like to extend this concept by leveraging **other mobile sensors**’ (compass, gyroscope, accelerometer) data to further improve proximity calculations.
- To **solve the above conundrum of an equation with multiple unknowns**, we decided to apply **functional mathematics** assuming that there is a relationship exists between RSSI and d, which can be expressed as follows: $d \Rightarrow f(RSSI)$
- The function “f” is modeled through by taking a series of RSSI signal values **(1000+)** for various distances from 10cm to 30m using the **heterogeneous Bluetooth devices (listed in the table on the left below)**.
- Once we have series of measurement values for RSSI & ‘d’ for various distances (10cm to 30m), we looked into various **regression models** and finally decided to use a Polynomial regression model to map the values from RSSI domain to distance domain.





Bluetooth Device (Node 1)	Bluetooth Device (Node 2)
HC05	Arduino
BLE Device	Raspberry Pi 3B+
Air Pod	Raspberry Pi
Headset	Android Phone
Tablet	Android Phone
Smart Watch	Android Phone
Fit Bit	iPhone

Model	Mean Square Error (MSE)
Linear Regression	33.5476
Polynomial (Degree 2) Regression	12.6430
Polynomial (Degree 4) Regression	1.7363
Generalized Additive Model (GAM of Degree 4+2+1)	1.7363

Key Takeaway from this slide

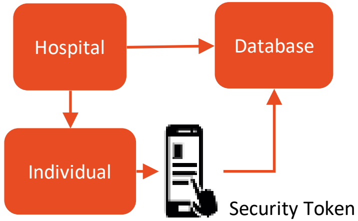
Addressing the unreliability and device dependency issues with RSSI values, we conducted experiments to derive a regression model that transforms RSSI values to distance in meters which can be used as a filter to narrow down the contact network.

Step 3 – Command Centre Identifies Infected Person

1. Patient "X" Tested Positive on Date "Y"

A

Following the footsteps of the DP-3T protocol, either a local healthcare system or a patient with a security token issued by a hospital can declare themselves as a COVID-19 patient on a specific date.

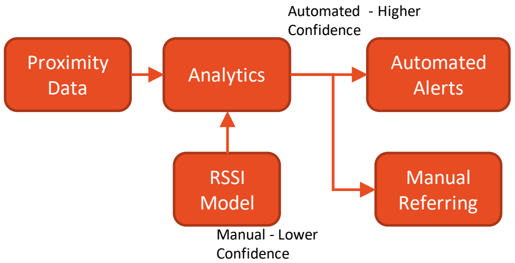


2. Analytics and Power of Visualization

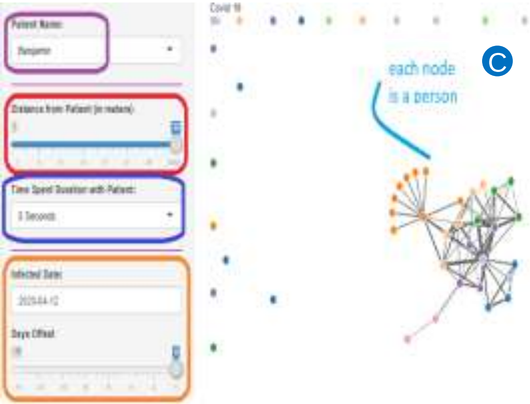
B

As soon as a person is declared COVID-19 patient:

- **Automated:** The analytics scan through raw proximity data. If a person is in close proximity (with a higher confidence limit) then an alert is triggered automatically. (this is a Future work using Probabilistic & Decision Based Models)
- **Manual:** In case of a lower confidence on proximity, the command centre assesses the contact context using the visualization to make the decision to trigger the alerts



C

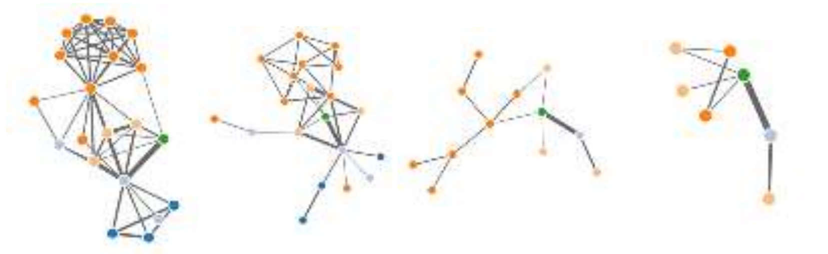


Filter 1: Distance Between the Patient and Others

D

Distance between a patient and other people derived from the RSSI model can be used as a filter to narrow down the contact network and make appropriate judgement calls by combining deeper context (explained in the next slide) to decide whether or not an alert should be sent.

Contact Graph (Distance 30m)	Contact Graph (Distance 15m)	Contact Graph (Distance 5m)	Contact Graph (Distance 3m)



Key Takeaway from this slide

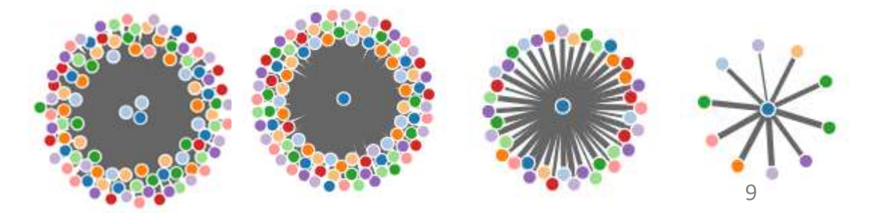
By applying Analytics using the Time Spent Duration (with the sink device) and Distance Derived from the RSSI Model (between source and sink), the command centre can narrow down the contact network and minimize the potential for false alerts.

Filter 2. Time Duration Spent by Patient with Others

E

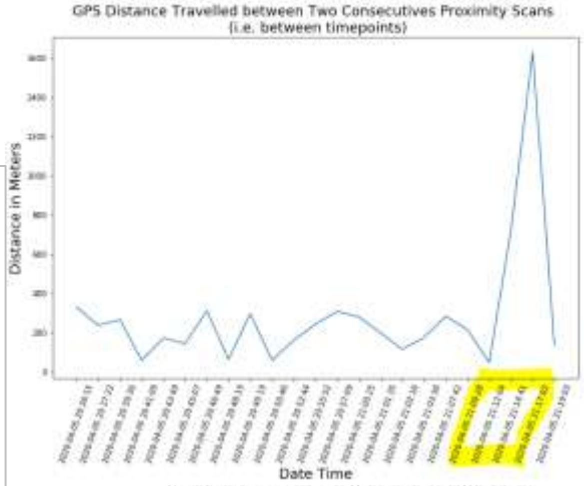
The network graph shows how the amount of time a patient has spent with others which can be used to narrow down the network and make judgement calls as shown below.

Time Duration with Patient (10 Seconds)	Time Duration with Patient (30 Seconds)	Time Duration with Patient (2 Minutes)	Time Duration with Patient (10 Minutes)



Step 4 - Power of Visualization – Deriving Deeper Insights

The GPS maps are useful because they give context as to where the COVID-19 patient has been in the past. But when we combine GPS Information with other analytics much deeper context can be derived. Two examples are provided below:



Example 1: Interpreting Source State

The two plots on the right are:

(a) Distance travelled by the Source between two consecutive proximity scans against time.

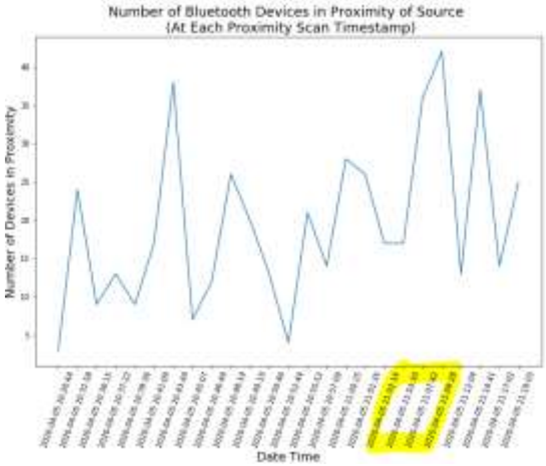
(b) GPS map along with timestamp.

The plot reveals:

The source is travelling at an average of around 100-200m

At 21:12 the distance travelled shot up to 750m, and then at 21:19 to 1.6km

- If you combine the distance travelled from the GPS location, GPS Map and Average Speed – we can derive the state of the source:
  - walking/cycling: (depending on average distance),
  - Driving: higher average distance and GPS path shows public road, **or**
  - on public transportation: higher average speed and GPS path shows tram or rail track



Example 2: Interpreting Source Surroundings

The plot on the right shows the number of Bluetooth devices nearby after each proximity scan by the source.

The plot reveals:

- On average, the number of Bluetooth devices around the source is 15-20
- At 20:41 and 21:14, the number of devices shot up to 49.

When we combine this information with GPS location, we can interpret the source surroundings such as **Crowded or Non-crowded** places

Key Takeaway from this slide

The key takeaway from these examples are:

(a) Source State (Walking/cycling/driving and public transportation)

(b) Source Surroundings context: crowded/uncrowded - helps the command centre make an appropriate judgement call as to whether or not a person should be alerted. Hence increasing the potential to minimize false alerts.

Conclusions & Next Steps

Conclusions:

- RSSI Model: addressing unreliability, we proposed a primitive regression model for distance calculations using RSSI. This model has the potential to enhance and derive deeper contact context (floor levels/orientation/obstruction, etc.) by leveraging other mobile sensors (e.g. compass, gyroscope, accelerometer, etc.).

- Visualizations for contact tracing: With the power of visualization techniques, we demonstrated that Distance & Time Duration filter can be used to precisely narrow down the contact network and hence minimizing the false alerts.

- Deeper Context: we also demonstrated that applying analytics on top of proximity data can be used to derive richer context, which further helps in minimizing the false alerts.
  A person's state: walking, cycling, driving, or public transportation.
  A person's surroundings: crowded or uncrowded place.

- **Key Conclusions: If we rely entirely on Bluetooth for contact tracing there is significant potential for false alerts due to Bluetooth unreliability. Therefore, contact tracing requires deeper context around how the contact occurred (time duration, distance, surroundings, etc.). Through experiments and visualizations we proved that deeper contact context can be achieved if we combine RSSI modelling with analytics.**

Next Steps:

- Random Forest Machine Learning Models: Our next step is to implement self-learning algorithms such as Probabilistic and Decision-based models for interpreting contact context, and making the decision to send an alert based on a set probability threshold.

- Enhance RSSI Model: We would like to extend the RSSI model to take data from a wider variety of sensors (accelerometer, compass, gyroscope, pressure, and temperature) to derive a richer context on things like mobiles in landscape/portrait, two people in the same room or separated by a thick/thin wall, people facing each other, etc.

- Diagnosing COVID-19: We would like to take our app into the disease diagnosis world, where a user can upload a short 10 seconds audio clip of a cough. A Deep Learning model can inform the user whether he/she potentially has COVID-19.

- Enhance Visualizations: We would like to enhance the network graph visualizations where operating staff can further zoom-in or zoom-out to gain further insight into contact tracing.

References

[1] Suresh Kumar, Swarun, Pushing the limits of wireless networks, Published 2016, Ph. D., Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2016. https://dspace.mit.edu/handle/1721.1/103675
[2] Dong, Qian & Dargie, Waltenegus. (2012). Evaluation of the reliability of RSSI for indoor localization. 2012 International Conference on Wireless Communications in Underground and Confined Areas, ICWCUCA 2012. 1-6. 10.1109/ICWCUCA.2012.6402492.
[3] Jennifer Stisa Granick, Published April 18th 2020, Apple and Google Announced a Coronavirus Tracking System. How Worried Should We Be?, https://www.commondreams.org/views/2020/04/18/apple-and-google-announced-coronavirus-tracking-system-how-worried-should-we-be
[4] Leo Kelion, Published April 10, 2020, Coronavirus: Apple and Google team up to contact trace Covid-19, https://www.bbc.com/news/technology-52246319
[5] BBC, Published April 17, 2020, Coronavirus: EU and Australian tracing apps 'ready in weeks', https://www.bbc.com/news/technology-52325352
[6] Siwss Info, Published April 21, 2020, Contact tracing app could be launched in Switzerland within weeks https://www.swissinfo.ch/eng/sci-tech/digital-solution_contact-tracing-app-could-be-launched-in-switzerland-within-weeks/45706296
[7] Saheli Roy Choudhury, Published March 25, 2020, Singapore says it will make its contact tracing tech freely available to developers, https://www.cnbc.com/2020/03/25/coronavirus-singapore-to-make-contact-tracing-tech-open-source.html
[8] Paul Mozur, Raymond Zhong and Aaron Krolik, Published March 1, 2020, In Coronavirus Fight, China Gives Citizens a Color Code, With Red Flags, https://www.nytimes.com/2020/03/01/business/china-coronavirus-surveillance.html
[9] The Guardian News Paper, Published March 17, 2020, Israel to track mobile phones of suspected coronavirus cases, https://www.theguardian.com/world/2020/mar/17/israel-to-track-mobile-phones-of-suspected-coronavirus-cases
[10] Jennifer Valentino-DeVries, Published March 19, 2020, Translating a Surveillance Tool into a Virus Tracker for Democracies, https://www.nytimes.com/2020/03/19/us/coronavirus-location-tracking.html
[11] Natasha Singer and Choe Sang-Hun, Published March 23, 2020, As Coronavirus Surveillance Escalates, Personal Privacy Plummets, https://www.nytimes.com/2020/03/23/technology/coronavirus-surveillance-tracking-privacy.html
[12] Rory Cellan-Jones, Published April 28, 2020, Coronavirus: UK contact-tracing app 'ready in two to three weeks, https://www.bbc.com/news/technology-52458759.
[13] Pan-European Privacy-Preserving Proximity Tracing “Data Protection and Information Security Architecture – Illustrated on German Implementation”. https://github.com/pepp-pt/pepp-pt-documentation/blob/ master/PEPP-PT-data-protection-information-security-architecture.pdf, Accessed on: April 19, 2020
[14] **J**ulian Schuette, AISEC Fraunhofer DE, Published April 27, 2020 , PANDEMIC CONTACT TRACING APPS: DP-3T, PEPP-PT NTK, AND ROBERT FROM A PRIVACY PERSPECTIVE, https://eprint.iacr.org/2020/489.pdf
[15] Big Data Institute, University of Oxford, Published March 31 2020, https://www.bdi.ox.ac.uk/news/controlling-coronavirus-transmission-using-a-mobile-app-to-trace-close-proximity-contacts
[16] Norimitsu Onishi and Constant Méheut, Published April 17, 2020, France Weighs Its Love of Liberty in Fight Against Coronavirus, https://www.nytimes.com/2020/04/17/world/europe/coronavirus-france-digital-tracking.html
[17] Linnet Taylor, Published 2015, Environment & Planning D. Society & Space (2015), No Place to Hide? The ethics and analytics of tracking mobility using mobile phone data , University of Amsterdam
[18] Patrick R Saunders and Daniel Krewiski, Published Dec 6, 2016, Reviewing the history of pandemic Influenza: Understanding patterns of emergence and transmission, Journal of Pathogens MDPI, https://doi.org/10.3390/pathogens5040066
[19] T. Deirdre Hollingworth, Don Klinkenberg, Hans Heesterbeek, Roy M Anderson, Published 2011, RM(2011) Mitigation Strategies for Pandemic Influenza A: Balancing Conflicting Policy Objectives, https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1001076
[20] 2020 Johns Hopkins University, Coronavirus Statistics, https://coronavirus.jhu.edu/map.html