

DATA SCIENCE EXPLORING COVID 19 GLOBAL HEALTH TRACKER

Bhavya Alankar and Siddhartha Sankar Biswas

Department of CSE, Jamia Hamdard, New Delhi, India

bhavya.alankar@gmail.com and ssbiswas@jamiahamdard.ac.in

ABSTRACT

This research study emphasizes the importance of data analysis through the usage of data visualization software to help you gain an understanding of data and how it can be transformed into information that can enhance the decision-making process. The appliance requires being in running mode in the slightest degree times to continue tracing individuals actively. The API of the appliance are often employed in such some ways that it enables your smartphone to exchange the tracing keys periodically. This may help to locally store the unique ID of the people that have get contact with the user.

KEYWORDS: Tableau, Data Visualization, Business Intelligence, Dashboard, Covid 19, Data Exploration, Data Visualization Techniques

1. INTRODUCTION

Numerous governments in their nations have been quickly developing COVID-19 tracking technologies or contact-tracing applications. In this study, the Covid 19 tracker, an instrument created by the Indian government, is examined. It is a smartphone application created by the Ministry of Healthcare as a part of the government's E-Governance effort to track and encourage Indian residents in an extremely cooperative fight to combat the globalization of COVID-19. The purpose of the research is to learn about the many beneficial aspects of this instrument and to provide numerous understandings of scientific ideas that have been used in the appliances as well as the significance of this tool in controlling the continuing pandemic. The app alerts a user when they are close to a COVID-19-contaminated individual using Bluetooth connectivity and the use of GPS. To examine COVID-19 distribution in India, the application makes use of a variety of Data Science principles, including categorization, mining of association rules, and segmentation. The study also demonstrates possible enhancements to the program, which uses computation and artificial intelligence (AI) to identify COVID-19 patients. Professionals in the fields of information science, smartphone technology, medicine, frontline personnel in the healthcare industry, and government officials and administration might all benefit from the findings of this investigation.

The World Health Organization has classified the dangerous coronavirus, medically known as COVID-19, as an epidemic since it has spread to every corner of the globe. In November 2019, it first appeared in China, and it has since become more prevalent in all the world's main regions. Around 187 countries and provinces documented around 16.52 million COVID-19 cases as of July 27, 2020. The first coronavirus case to be recorded in India occurred on January 30, 2020, and there are already 1.48 million infections there. Mild moderate to severe breathing difficulties is experienced by people of all ages who have COVID-19 infections. If someone tests positive for the coronavirus, it is advised that everyone who has encounter them depart for quarantine for two weeks to break the chain of transmission and prevent the sickness from spreading further. For COVID-19, there isn't yet a

particular medication or vaccination. A lot of nations are working to create contact-tracing methods that will help them find the individual who is thought to be infected. For instance, the government of South Korea has a directory that contains patients who have been identified with additional information about them, such as their gender, year of birth, and employment. In Israel, the government is permitted to track the cell phone information of those who are allegedly infected. A smartphone app called TraceTogether was created in Singapore and uses Bluetooth to monitor users anytime they are close to one another. An individual's contact mapper can decide on further steps if their report identifies them as having COVID-19 [1,2].

2. MOTIVATION AND OBJECTIVE OF STUDY

Employing the database's information and algorithms, the planned effort will find nearby cell phones with the application loaded and will notify users if they are linked to infected individuals. Additionally, the app offers crucial guidance on how to isolate oneself and what items should be thrown away if someone exhibits symptoms. The app suggests keeping Bluetooth turned on as little as possible, and with it, location services. To use the application, one needs to sign up using a cell phone number and validate it with an OTP. Users are asked to fill out an application on the app that requests specific data about them, including their first and last names, age, work, and any recent overseas travel. Additionally, the Covid 19 tracking Implementation has a "Self-Assessment Test" option that asks certain questions about the health and signs after asking the user a series of questions and recording their responses, the device displays the user's hazard level using a variety of colour codes. If it reads "You are safe" in the text and the colour code is green, there is certainly no risk. If it displays the yellow signal and the phrase "You are at moderate risk" simultaneously, the danger is low. If it displays an orange colour code, there is a high likelihood of risk; in that case, the user should call the phone number for assistance provided in the app. The device also recommends measures that will be done to stop the transmission of the coronavirus, such as maintaining a distance from others, maintaining good cleanliness, and staying in one place as a precaution. The information collected through the form is shared solely with the Indian government and is not permitted to be divulged to any other parties, as stated in the statement of privacy for the application that was submitted. If required, the material will support the government's decision to start the process of isolating itself and take the necessary actions. It will not be made public to reveal the user's identity or phone number. The COVID-19 tracker is divided into four categories: Your Status, Self-Assess, COVID-19 Updates, and E-Pass. The user's risk of contracting COVID-19 is disclosed in the "Your Status" section. Users are warned about the risk of contracting an infection in the "Self-Assess" section. A mobility pass may be obtained through E-Pass merely in case of emergency, and the COVID-19 Update section provides information on both local and national COVID-19 cases. The App has been developed on a platform that may offer an API, or Application Programming Interface, to enable various application types so that additional computer programs, such as web-based offerings and mobile apps, can utilize the data accessible and other aspects of the Covid-19 tracker [3].

To keep the data associated with known instances and suspected cases, the program leverages a secure database maintained by the Indian government. The government has not released its database or any information-related program freely available to prevent the device and the accompanying information from hostile hacking efforts. The technical features of the App are not given much care because of its security. The App's policies and conditions, front-end evaluation, and press announcements are the only sources of further information. To inform users if they are close to someone who has COVID-19 or who may be at high risk of contracting it, the application leverages both Bluetooth connectivity and GPS-generated information to perform contact tracking. For the App to operate properly, the user's device's GPS and Bluetooth must be switched on. For the appliance to determine if two people are close enough to each other, Bluetooth access might be a crucial factor. When two mobile devices with the App downloaded are within Bluetooth range of each other, the device gathers data.

We may better define or understand how an application functions and how it analyses its data by doing a thorough examination of similar apps, such as the Covid-19 tracker. A program called TraceTogether aids Singapore's attempts to reduce the effects of the ongoing epidemic. Like Covid-19

Tracker, TraceTogether analyses Bluetooth signals to determine when a user's phone is close to another phone that also has the device installed. Anonymized IDs are sent between the two phones, and only the user's smartphone uses cryptographic techniques to store these IDs in a form that is encrypted. The anonymized IDs disappear following a quarter-hour of creation, which is a benefit of the application since it prevents hackers from seeming to be able to change the data. The Massachusetts Institute of Technology has created another program called Safe Paths, which is in the Private Kit. The user is also prompted to indicate whether they are employed in the provision of necessities and whether they could be available to help in an emergency. To track down possible instances and identify volunteers who are eager to help during a crisis, the government may then download this data. Since proximity isn't the only element contributing to contagion, other separate variables, such as being in an enclosed or unstructured environment, have also been linked, further research on the usefulness of closeness mapping techniques like Bluetooth connectivity and satellite navigation has been done.

The application alerts the opposing user anytime a user is tested positively, helping the authorities keep track of potential COVID-19 occurrences. The program also guides any local virus illnesses as a supplement to these cautions. Along with connecting via Bluetooth and data from the website, the software also collects demographic data about users, including their gender, age, entire name, and history of overseas travel.

The usefulness of the App can be greatly diminished by a small user base caused by a slow rate of adoption. Apps that use Bluetooth are dependent on the handsets' Neighbour Discovery protocol. As a result, many cell phones have wildly different accuracy levels for signal detection and emission. The protocol establishes the quantity of electrical power used (i.e., battery life), the dependability of operations when many phones are close to the cell's receptor order, and the speed at which two phones can identify each other once they've come into proximity. The number of erroneous negatives and erroneous positives produced by proximity-based apps is rather high. Since the wireless signal weakens in open space, distance is calculated. But even human beings can significantly dampen the signal. The signal could become weaker if two people are facing each other from behind, misunderstanding their distance from each other and producing an unintentional negative as a result. The location of smartphones concerning one another (such as whether the mobile device is in a person's purse or hands) determines how accurate the App will be. For instance, if there is a thin wall separating two flats, the signal may not be attenuated enough and result in a proximity event even though it is not. False positives might cause unaffected individuals to be isolated, while erroneous negatives can undermine the application's most important goals [4].

GPS-based approaches are less accurate than Bluetooth-based ones. GPS-based tracking applications are far less widely adopted than Bluetooth-based tracking apps in terms of acceptance standards, battery consumption, and false-positive rates. GPS data is incorrect and inaccessible within buildings, in addition to privacy problems with GPS tracking. In comparison to 43% of the applications that utilize GPS, 57% of contact-tracing apps use Bluetooth. In GPS-based applications, there is a significant trade-off between information and privacy. While the privacy of users may be a major worry, Bluetooth-only applications are unable to give contact-tracing data by tracking users' geographic movements. Applications for contact tracing that utilize Bluetooth and GPS together are significantly more beneficial than those that solely use Bluetooth. Health authorities can detect exposure trails more precisely with the use of location data, let alone proximity history. Therefore, it appears that the hardware and software on which Covid 19 tracking App is based are generally in good shape. Additionally, the user is informed by this program anytime they encounter a COVID positive individual. The goal of the app is to make the contact tracking process simpler. To stop the spread of contagious illnesses, contact tracking is a crucial technique. This program makes use of GPS-based configuration. However, since anonymizing GPS data is more challenging, the equipment's developers are looking at specialized technological solutions including more robust cryptography. The program is set up so that it notifies subscribers through notifications whenever they meet a COVID-positive person. The tracking is made possible therewith the implementation of Bluetooth, which is used for location-generated social media platforms, or GPS, and it notifications

anybody who has interacted with a user who was identified as coronavirus positive after being tested. Using the integration of Bluetooth as well as GPS detectors, it recognizes and monitors the user's movement. The research paper's primary goal is to learn about Covid 19 Statistics and put the strategies into practice.

3. DETAILED DESCRIPTION OF STUDY

Employing the database's information and algorithms, the planned effort will find nearby cell phones with the application loaded and will notify users if they are linked to infected individuals. Additionally, the app offers crucial guidance on how to isolate oneself and what items should be thrown away if someone exhibits symptoms. The app suggests keeping Bluetooth turned on as little as possible, and with it, location services. To use the application, one needs to sign up using a cell phone number and validate it with an OTP. Users are asked to fill out an application on the app that requests specific data about them, including their first and last names, age, work, and any recent overseas travel. Additionally, the Covid 19 tracking Implementation has a "Self-Assessment Test" option that asks certain questions about the health and signs after asking the user a series of questions and recording their responses, the device displays the user's hazard level using a variety of colour codes. If it reads "You are safe" in the text and the colour code is green, there is certainly no risk. If it displays the yellow signal and the phrase "You are at moderate risk" simultaneously, the danger is low. If it displays an orange colour code, there is a high likelihood of risk; in that case, the user should call the phone number for assistance provided in the app. The device also recommends measures that will be done to stop the transmission of the coronavirus, such as maintaining a distance from others, maintaining good cleanliness, and staying in one place as a precaution. The information collected through the form is shared solely with the Indian government and is not permitted to be divulged to any other parties, as stated in the statement of privacy for the application that was submitted. If required, the material will support the government's decision to start the process of isolating itself and take the necessary actions. It will not be made public to reveal the user's identity or phone number. The COVID-19 tracker is divided into four categories: Your Status, Self-Assess, COVID-19 Updates, and E-Pass. The user's risk of contracting COVID-19 is disclosed in the "Your Status" section. Users are warned about the risk of contracting an infection in the "Self-Assess" section. A mobility pass may be obtained through E-Pass merely in case of emergency, and the COVID-19 Update section provides information on both local and national COVID-19 cases. The App has been developed on a platform that may offer an API, or Application Programming Interface, to enable various application types so that additional computer programs, such as web-based offerings and mobile apps, can utilize the data accessible and other aspects of the Covid-19 tracker [3].

To keep the data associated with known instances and suspected cases, the program leverages a secure database maintained by the Indian government. The government has not released its database or any information-related program freely available to prevent the device and the accompanying information from hostile hacking efforts. The technical features of the App are not given much care because of its security. The App's policies and conditions, front-end evaluation, and press announcements are the only sources of further information. To inform users if they are close to someone who has COVID-19 or who may be at high risk of contracting it, the application leverages both Bluetooth connectivity and GPS-generated information to perform contact tracking. For the App to operate properly, the user's device's GPS and Bluetooth must be switched on. For the appliance to determine if two people are close enough to each other, Bluetooth access might be a crucial factor. When two mobile devices with the App downloaded are within Bluetooth range of each other, the device gathers data.

We may better define or understand how an application functions and how it analyses its data by doing a thorough examination of similar apps, such as the Covid-19 tracker. A program called TraceTogether aids Singapore's attempts to reduce the effects of the ongoing epidemic. Like Covid-19 Tracker, TraceTogether analyses Bluetooth signals to determine when a user's phone is close to another phone that also has the device installed. Anonymized IDs are sent between the two phones, and only the user's smartphone uses cryptographic techniques to store these IDs in a form that is encrypted. The anonymized IDs disappear following a quarter-hour of creation, which is a benefit of

the application since it prevents hackers from seeming to be able to change the data. The Massachusetts Institute of Technology has created another program called Safe Paths, which is in the Private Kit.

The usefulness of the App can be greatly diminished by a small user base caused by a slow rate of adoption. Apps that use Bluetooth are dependent on the handsets' Neighbour Discovery protocol. As a result, many cell phones have wildly different accuracy levels for signal detection and emission. The protocol establishes the quantity of electrical power used (i.e., battery life), the dependability of operations when many phones are close to the cell's receptor order, and the speed at which two phones can identify each other once they've come into proximity. The number of erroneous negatives and erroneous positives produced by proximity-based apps is rather high. Since the wireless signal weakens in open space, distance is calculated. But even human beings can significantly dampen the signal. The signal could become weaker if two people are facing each other from behind, misunderstanding their distance from each other and producing an unintentional negative as a result. The location of smartphones concerning one another (such as whether the mobile device is in a person's purse or hands) determines how accurate the App will be. For instance, if there is a thin wall separating two flats, the signal may not be attenuated enough and result in a proximity event even though it is not. False positives might cause unaffected individuals to be isolated, while erroneous negatives can undermine the application's most important goals [4].

GPS-based approaches are less accurate than Bluetooth-based ones. GPS-based tracking applications are far less widely adopted than Bluetooth-based tracking apps in terms of acceptance standards, battery consumption, and false-positive rates. GPS data is incorrect and inaccessible within buildings, in addition to privacy problems with GPS tracking. In comparison to 43% of the applications that utilize GPS, 57% of contact-tracing apps use Bluetooth. In GPS-based applications, there is a significant trade-off between information and privacy. While the privacy of users may be a major worry, Bluetooth-only applications are unable to give contact-tracing data by tracking users' geographic movements. Applications for contact tracing that utilize Bluetooth and GPS together are significantly more beneficial than those that solely use Bluetooth. Health authorities can detect exposure trails more precisely with the use of location data, let alone proximity history. Therefore, it appears that the hardware and software on which Covid 19 tracking App is based are generally in good shape. Additionally, the user is informed by this program anytime they come in contact with a COVID positive individual. The goal of the app is to make the contact tracking process simpler. To stop the spread of contagious illnesses, contact tracking is a crucial technique. This program makes use of GPS-based configuration. However, since anonymizing GPS data is more challenging, the equipment's developers are looking at specialized technological solutions including more robust cryptography.

4. DATA SCIENCE TECHNIQUES

Find significant "hotspots" of COVID-19 patients and impose the necessary controls on these hotspots to stop further virus transmission is the main objective of contact tracing.

With the use of various fields of computer science or machine learning approaches, clustering is frequently generated using the coordinates of the longitude and latitude of COVID-19 patients as well as at-risk individuals. Thus, hotspots may be thought of as these groupings. Unsupervised algorithms for machine learning, such as K-means clustering, can be used to create groups of people with infections and find groupings. There is a chance the use of the density-based function HDBSCAN (which is based entirely on the DBScan technique) will be used in the equipment or may be employed in identical manners. Furthermore, the self-evaluation test, backed by the ICMR criteria, assesses the probability of COVID-19 infection backed by the individual's self-reported symptoms along with additional pertinent data including recent departure, chronological age, and ethnicity. Users might be divided into categories like green (safe), yellow (ranging from slight risk), orange (great risk), and red (validated patients) using an algorithm for categorization like an assessment tree or randomly generated forest. The self-evaluate exam requires users to respond to the questions. Green, yellow, or

orange colors, which represent an increasingly larger probability of infection, are used to transmit the findings of this evaluation, which is conducted on the appliance, to the consumers right away. The app does categorization following the users' responses. Further investigation may reveal that the equipment also uses the iBeacon Bluetooth proximity data collection protocol. It is a protocol created by Apple. Beacons, a category of Bluetooth Low Energies Devices manufactured by several manufacturers (which incorporates Android), are equipment broadcasters that are iBeacon compatible. While most Bluetooth-enabled gadgets require a procedure called "pairing" for the transmission of information, which means that the individual in charge of the device in question needs to authorize the transfer of information on both gadgets for the operation to proceed, this particular gadget continually communicates a Bluetooth 4.0 connection (BLE) along with its specific identification number to nearby handheld electronic gadgets. Due to the one-way transmission of iBeacons, information may be transferred to a receiver device (such as a mobile phone, laptop, etc.) without the need for pairing, greatly simplifying the process. Near an iBeacon, the device enables smartphones, tablets, and other gadgets to take on certain activities. The Covid 19 tracker, a suitable application, first recognizes the globally unique identity. The ID and several additional bytes transmitted along with it are used to track users, locate devices physically, or start location-based actions on a device, such as push notifications. Users of the app are alerted anytime they are too close to a coronavirus-infected person or someone else who is in a high-risk situation. In the future, the Indian government or the company that created the electronic device will make available Bluetooth vicinity data from the past (determining the identifiers, the strength of the signal, etc.) to a secure database from which data may be obtained. Once such data has been extracted, the government may analyse it.

The Covid 19 tracker software gathers specifically identifiable data about the person using it, including designation, quantity, gender, age, occupation, nations travelled in the previous 30 days, and then-current medical history, and keeps it in the cloud for future use. In addition, while registering, the user's location information is obtained. A record of the user's movements is also kept by the App, which uses Bluetooth and GPS technologies to keep tabs on where the individual is continually. Substantial ramifications for the security of information might arise from this. If someone has access to this data, they can infer the user's home and workplace addresses, socioeconomic class, identification (such as religion or caste), and a lot more. Concerns regarding privacy breaches are raised by the presence of private data and the resulting requirement for utilizing the App. Invasion of privacy regarding medical information may be extremely detrimental to a person's status in society and may result in discriminatory actions.

The lists of those who have been infected by the coronavirus have been shared randomly across India. Due to the possibility of information being utilized inappropriately, such a privacy risk might cause serious harm. The state-level government are revealing the user information, including name, residence, and address, of those who are under anaesthesia quarantine, which has caused knowledge to be shared illegally. In addition, it disagrees with the Indian Supreme Court's ruling on the right to privacy. The App doesn't pass the measure of proportionality required by the right to privacy, which stipulates that the act in question must be legal, have a valid purpose, be in proportion in scope to the need for it, and be protected by procedural safeguards designed to prevent abuse. Accordingly, the proportionality test is satisfied because there are no laws controlling this app, which increases the possibility of misuse. India does not have a thorough private information security framework, which is against the law. The vast quantity of confidential data that the App safeguards goes against the standards regarding knowledge reduction, purpose any limitation, openness, and responsibility. The conditions of service and confidentiality agreements, as well as the lack of any controlling law, are responsible for an information security concern [5,6]. To protect the users' privately held privacy, the app has measures built in. It produces an original gadget ID number. The user's name and phone number are not sent to the other person's device when they contact the user; instead, this ID is sent. A rolling 30-day period determines when the placement data on the App is erased. The Covid 19 tracking app has been made open source, which can help the developers understand how it functions and the way data is handled by it. The government's dedication to transparency is demonstrated by this. The written portion of the App has undergone certain changes. A tool ID is still used even if the

data is no longer hashed to a unique ID. That private data can be organized in this way is no longer asserted by the statement regarding privacy. The information, which consists of the user's precise location knowledge that gets gathered on a quarterly hour and retained on the mobile device themselves, and information of all DiD trading platforms between individuals a user got near, will only be published to the server's memory upon submitting a COVID-19 positive status or upon asking for a test. additional than they DiD, none of the consumer's confidential data will be posted to the server. If a particular individual tested positive for malaria and is successfully treated, such data is erased following 60 days; if they test negative, it gets discarded after 45- days [7,8].

5. SIMULATION STUDY

React JS is employed in the proposed task. Making interactive user interfaces is simple using React. Create straightforward views for each stage of your application, and as the data in those views' changes, react will quickly update and render the appropriate components. Declarative views increase predictability and simplify bug-finding in your code. It creates isolated components with independent state management, which are subsequently assembled to provide intricate user interfaces. We will simply transmit rich data throughout the application while maintaining the state of the structure of the DOM since element functionality is defined in Java compared to the template. One may utilize a Chart.js CDN, download the most recent version of Chart.js from GitHub, or use npm. The accompanying installation page will provide comprehensive instructions for installation. Utilizing Chart.js, getting started is simple. The script must be present on your page and just one of the following elements is necessary:

6. DISCUSSION & CONCLUSION

Tracker for Covid 19 Application somewhat deviates from "Privacy-focused International Guidelines" limited to solely Bluetooth-based technology employing GPS Trails. Because it conceals the mobile phone's exact location, it is simple to match devices. In enclosed spaces or on public transportation, GPS tracks don't appear to be the least dependable. Such technology is susceptible to hacking. Data can be posted to a cloud server in specific circumstances, where it might then be removed and utilized improperly. Both on the mobile device and consequently on a centralized server, the application stores the data it has gathered. The encoded information about customers that is stored on its computer servers may be kept after the coronavirus has been traced. Another conclusion that may be drawn is that it's an intriguing use that might help people stay safe from coronavirus. The appliance offers practical features and an excellent design; however, it might need some enhancements like using better Deep Learning algorithms and resolving data privacy problems in an extremely safer way. User interactivity was considered while the device was being designed, as seen by the magnificent computer program (UI). Overall, the tool aids the government in tracking down those who test positive for the virus. It's also a great idea to let people know how many cases of infection have been confirmed as coronavirus positive in their neighbourhood or if they unintentionally encountered someone who was COVID-19-positive. To continue actively tracking people, the appliance must always be in operating mode. The API of the appliance is frequently used in a manner a way that allows the phone you're using to regularly interchange its tracking keys. With this, it might be possible to locally keep each contact person's distinct ID (Trace Together also uses this strategy). The safety and confidentiality of your data will still be guaranteed even if a user is subsequently found to have a coronavirus, thanks to the simple fact that the encrypted IDs are nameless and are refreshed on a quarterly hour. A map of the neighbourhood where individuals are detected as being positive for COVID-19 may be displayed together with the information on the number of users who have performed the self-assessment exam who are classified as affirmative. When a user enters a confinement zone or red zone, alerts are delivered to them through email and SMS. With the details and recommendations, other nations and regions can use them as a model for how to use contact-tracing applications in their localities and communities. The results of this study will be useful to a variety of researchers, educators, scientists who analyse data, developers of software, those making decisions, politicians, and governmental administrators.

REFERENCES

- [1] <https://reactjs.org/docs/>
- [2] <https://disease.sh/docs/>
- [3] <https://material-ui.com/getting-started/usage/>
- [4] ReactJS: An Open Source JavaScript Library for Front-end Development
- [5] https://www.theseus.fi/bitstream/handle/10024/130495/FInal_Year_Thesis.pdf?sequence=1&isAllowed=y
- [6] Case Study: Fighting Covid-19 with Tech and Innovation
- [7] https://www.researchgate.net/publication/341757281_Case_Study_Fighting_Covid-19_with_Tech_and_Innovation
- [8] <https://www.shareitsolutions.com/blog/mern-technology-stack> Lee, S.hyun. & Kim Mi Na, (2008) “This is my paper”, *ABC Transactions on ECE*, Vol. 10, No. 5, pp120-122.
- [9] Gizem, Aksahya & Ayese, Ozcan (2009) *Cooperations & Networks*, Network Books, ABCPublishers.

Authors

Dr. Bhavya Alankar and Dr. Siddhartha Sankar Biswas are faculty members of the Department of CSE, Jamia Hamdard, New Delhi. The authors have done research on various fields such as Data Mining, Data Visualizing Fuzzy Optimization, etc.

With teaching experience of over 15 years and have guided various projects and dissertations, the authors want to explore the domain of various real life problems that can be solved by using various computer science techniques.

