



Corona Virus Tracking System for Infected Patients

Shruti¹, Sahil², Dishita³, Khushi⁴, Harkesh kumar⁵

Department of Computer Science & Engineering (AI & DS), Panipat Institute of Engineering and Technology, Panipat, India ^{1,2,3,4,5}

shrutisharma660660@gmail.com¹, rawalsahil43@gmail.com², dishita2727@gmail.com³,
khushiraman00043@gmail.com⁴, Harkesh.cse.ai-ds@piet.co.in⁵

Abstract. *The recent coronavirus disease outbreak (COVID-19) has been recognized as one of the major health crises in the 21st century. As India is already an overcrowded nation with diverse demography and socio-economic status, hence, making the management of the disease outbreak challenging. Therefore, in this research paper, we have used the multi-dimensional data analytical approach to explore the impacts and control measures for the pandemic of COVID-19 in India. Multiple data analytic approaches have been employed in this study using quantitative data including infections rate, mortality rate, recovery rate, reproduction number (R_0) and doubling time, among others. Moreover, an analysis of lockdown and non-pharmaceutical interventions in controlling the disease transmission has also been conducted. The findings show that certain factors including urbanization, density of population and economic activities were some of the critical determinants of virus transmission. The lockdown was effective in the initial stage; however, its impact could not be sustained owing to improper public cooperation.*

Keywords: COVID-19, Data Analysis, Pandemic, India, Machine Learning, Visualization.

Introduction

Since the COVID-19 outbreak was characterized by rapid proliferation and numerous people infected, it became a serious issue that faced medical institutions. In order to manage the situation effectively, the precise and accurate tracing of the ill patients was required, which could be only achieved using computer software. Thus, the need for an automated computerized system appeared. The goal of developing such a system was related to collecting data from different sources, such as hospitals, laboratories, and government organizations. Besides, the system would be helpful in tracking the patients' condition, namely those sick, cured, and dead due to coronavirus. In addition, the system supports contact tracing and uses data visual display techniques like charts and dashboards to present information directly. This enables officials and healthcare professionals to make quick and well-informed choices. Overall, the tracking system plays major role in upgrading pandemic management by enhancing monitoring, analysis, and response to COVID-19.

II. Literature Review

Table 1: Literature Review Table

Method/Approach	Author(s)	Year	Key contributions
Contact tracing analysis using network models	Li et al.	2022	Provided insights into tracing efficiency and disease spread
GPS-based digital tracking system	Wang et al.	2020	Enabled real-time location-based



Method/Approach	Author(s)	Year	Key contributions
			monitoring
Digital contact tracing applications	Kleinman & Merkel	2020	Highlighted importance of mobile-based tracking
Review of tracking technologies	Shahroz et al.	2021	Compared different digital tracing techniques
Mathematical modeling of transmission	Ferretti et al.	2020	Showed effectiveness of fast contact tracing
Analysis of digital tracing systems	Hegde & Masthi	2020	Identified challenges in real-world implementation
Privacy-preserving tracking system	Peng et al.	2021	Proposed secure data handling approach
Adoption analysis of tracking systems	Chen & Thio	2021	Identified barriers to user acceptance
Use of smartphones and wearables	Ng et al.	2021	Introduced device-based tracking solutions
Security-focused tracking system	Hussein et al.	2020	Addressed privacy and cyber security concerns

III. Proposed System Architecture

COVID-19 tracking system for infected patients created by the researcher includes several data layers related to collection, processing, analyzing, and visualizing the collected data. In general, data gathered by the system can be described as integrated since the data gathered from various sources is processed into significant results. All processes performed within the framework of the system operate on the basis of the following rule when each part performs certain tasks in order to provide flexibility, stability, and maintainability. Generally speaking, the system architecture consists of a series of data layers. These include data gathering, data preparation, data storing, data analysis, visualization, and decision-making.

A. Data Gathering Layer

Data gathering layer involves data gathering process performed using different sources of information such as hospitals, diagnosis laboratories, testing centers, health departments, and others. Thus, patient's data regarding demography, disease condition, recovery process, geolocation, and others are gathered as raw data. Data gathering process is also conducted via using public API and publicly available datasets. The gathered data is considered continuous since it shows that information is updated and active. Moreover, gathered data can have diverse structures.

B. Data Preprocessing Layer

The data cleaning layer is applied to enhance the quality of the data. Data obtained from real-world cases may contain some null values, duplicates, and inconsistencies that might interfere with the analysis. Therefore, data cleaning, filtering, scaling, and data transformation processes are conducted at this layer. This ensures that the data used is accurate and credible since it is fundamental in conducting any meaningful analysis.



C. Data Storage Layer

The processed information will be stored in a point database that provides effective means of managing and accessing the data. This will facilitate monitoring of trends since there will be real-time as well as historic data. The databank will be able to manage large amounts of information and provide safe storage for confidential health information. It will also ensure that the data can be accessed rapidly.

D. Data Analysis Layer

All the significant deductions related to the data saved in the database will be made at this stage of data analysis. Various types of analysis methods are employed in order to determine significant parameters such as recovery rate, and mortality rate. Analysis of patterns is done in order to determine the spread of the virus. Pattern recognition, and critical regions recognition are some of the other aspects that must be considered here.

E. Contact Tracing and Risk Assessment Module

This module plays a crucial role in detect individuals who may have been at risk to infected patients. By scanning contact and location data, the system detects potential spread chains and helps in Segregating individuals. Risk assessment is performed by grouping regions into different risk levels based on case density and growth rate. This helps authorities to take preventive measures in high prone areas.

F. Visualization and Decision Support Layer

The final stage provides a hint on how the data gathered has been analyzed. The presentation of the data in graphs and charts makes it easy for the user analyzing the data to understand. He/she is able to get instant updates and also analyze data from various regions.

One can also prepare for future events by properly analyzing the data available. This could depend on how the data has been presented.

G. System Integration and Implementation

This technology utilizes up-to-date technological advancements to ensure efficiency and effectiveness. Data storage and processing is carried out by the backend technology, while the frontend provides an easy-to-use interface. It has been designed to be able to update in real time and handle large volumes of data effortlessly.

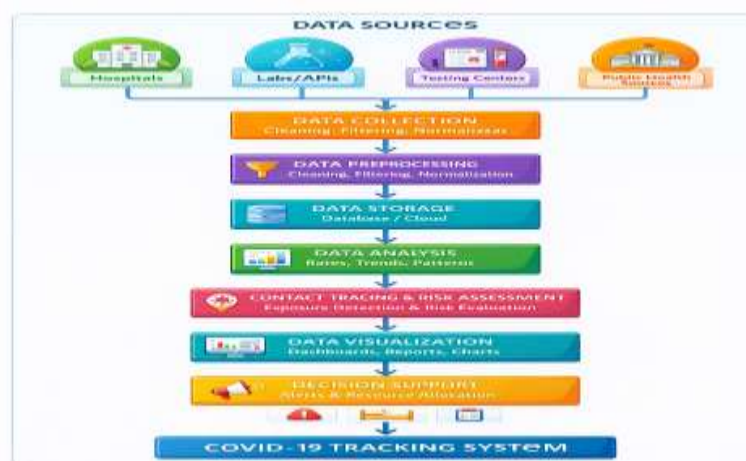


Fig. 1: Proposed System Architecture



IV. Methodology

The above system employs a step-by-step implementation process and works smoothly through the processing of visual data and the identification of missing persons as illustrated in Fig. 2. Every step of the process has played major role in the conversion of raw input data into meaningful information that can be used in making decisions [2], [10].

A. Data Collection and Preprocessing

The procedure starts by collecting the coronavirus disease data from various sources such as hospitals, testing centers, and health portals of government organizations. The collected data includes patient personal details, test results, recovered cases, and geographical locations, among others. In light of the possibility that there might be some problems with some of the data, such as redundancies and inconsistencies, it is critical to preprocess the data for better quality.

The data preprocessing stage involves data cleaning such as eliminating duplicate entries, handling missing values, and normalizing data, among others.

B. Data Integration and Storage

After processing the data, the clean data will then be placed in a central database system. In doing so, it ensures that all the data gathered from diverse sources will be consolidated into one single location. It includes live data as well as historical data, making it easier to trace.

An efficient way of storing the data will be used to ensure that the data retrieval will be done effectively.

C. Data Analysis and Processing

During this stage, the analysis of the gathered information is done in order to extract useful knowledge from it. Various methods are employed to ascertain the important factors, namely the infection rate, rate of recovery, and mortality rate. Moreover, the system observes the trend in relation to the increase or decrease in the number of patients.

The system also examines the trends regarding the transmission of the disease, which involves locations that are at risk for infection.

D. Contact Tracing and Risk Identification

There is also an option within the program to trace close contacts who have been infected by tracking their locations. This assists in tracing the route of transmission by analyzing the locations.

Risk assessment is carried out by dividing the areas based on the number of active cases and the possibility of the spread. Areas with higher risks are marked for immediate response to take preventive measures such as Quarantine.

E. Data Visualization and Reporting

Visualization techniques such as graphs, charts, and dashboards are used to present the analyzed information. The reason behind using such methods is to reduce the complexity of information to a comprehensible level.

A report is generated to provide an extensive overview of the existing situation. The report contains the number of cases recorded each day along with other research findings

F. System Implementation

The system is implemented using new technologies to ensure effectiveness and productivity. Python is used as the primary programming language due to its strong support for data analysis and visualization libraries. Tools such as Pandas and NumPy are used for data processing, while visualization is achieved using libraries like Matplotlib or Tableau.



The backend system manages data processing and storage, while the frontend provides an interactive interface for users. The system is designed to handle real-time data updates and provide accurate information to users and authorities.

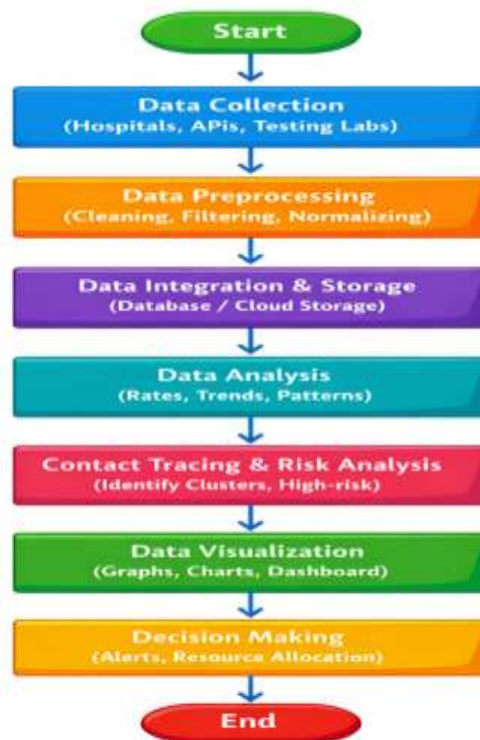


Fig. 2.: Methodology Flowchart

V. Results and Discussion

The COVID-19 tracing system is an effective tool for collecting data that allows immediate tracking of individuals infected with the disease. This system reveals the increase in the number of cases in various regions where a higher number of infected individuals live in crowded areas. The system also reveals the increase in the recovery rate of cases.

Data visualization using graphics and dashboard facilitates the identification of trends such as increases in case numbers, recovery rates, and fatality rate. This tool is important in helping to identify hot spots and contact tracing in managing the virus.

Despite being efficient in many aspects, the system might experience some challenges such as delays and inconsistency that may reduce its accuracy

A. Performance Analysis

The proposed COVID-19 tracking system has been evaluated based on its performance in data processing, response time, flexibility, and the accuracy of results which were likely to be generated.. The performance of



this system has to be a strong one. The data has to be structured in a way to allow for real-time tracking of infected patients. If this requirement is not satisfied, the users would not bother to use this service. The tracking has to give them quick access to updated information and visualizations. And as it has been mentioned before, timely decision making can only be made if the real-time tracking works well. In order to meet this requirement, an effective data processing layer has been implemented that cleans and organizes the data. The improvements would be reduced errors resulting in improved quality of analysis presented before the users in the form of charts and dashboards that represent the data clearly for better understanding and to aid in the interpretation of trends as well. Thus to process very large datasets, various proper optimization techniques have to be used and also need to be figured out depending on the requirement. As we are making real-time data processing, it also becomes very necessary to deal with the live data. It has been already pointed out that the performance of the system is based on various other criteria. But one large factor that the performance greatly depends on is the accuracy of the data that is being provided at every stage and that mainly depends on the quality and consistency of the input data which will be taken from the external sources. If the data is not accurate, then it will result in an inaccurate, ineffective, and inefficient system as a whole.

Table 2: Performance Evaluation Parameters

Parameter	Description
Response Time	Time taken to display updated data.
Data Accuracy	Correctness of processed data.
Efficiency	Speed of processing data.
Scalability	Ability to handle large datasets.
Reliability	Consistency of system performance.

B. Practical Advantage

The COVID-19 tracking system for COVID-19 patients is very useful in real life healthcare and administrative work. One of the things about the COVID-19 tracking system is that it allows us to watch COVID-19 infected cases in real time which helps the people in charge to act fast when things change. The COVID-19 tracking system always updates the information about COVID-19 patients so the people in charge can make decisions based on the information about COVID-19.

Another good thing about the COVID-19 tracking system is that it keeps all the information in one place. The COVID-19 tracking system brings together information from hospitals, testing centers and public health departments making it easy to find and manage information about COVID-19. This means we do not have to enter the information about COVID-19 more than once and it helps different groups of people work together on COVID-19.

The COVID-19 tracking system also helps us make decisions about COVID-19 by showing us information about COVID-19 in the form of pictures and charts. These pictures and charts help us see what is happening with COVID-19, where the danger zones are how COVID-19 is spreading so we can



take action to prevent COVID-19 from spreading. The COVID-19 tracking system also helps us find people who have been in contact with COVID-19 patients, which is very important, for stopping COVID-19 from spreading.

By aiding in the more effective allocation of medical materials (like hospital beds, staff and equipment) among many departments of the system as opposed to each department being allowed to allocate their own resources, the system helps officials make better decisions. The system reduces the amount of manual labour done to collect and process information by automating the process, thereby reducing human error. In addition, the system increases awareness about public health and the availability of accurate and easily accessible information about public health to help individuals adhere to health guidelines.

C. Restriction

The COVID-19 patients' tracking system outlined above will provide an efficient method of tracking and statistically assessing how quickly COVID-19 is spreading. Unfortunately, there are many limitations regarding this system that could negatively impact how effective and dependable this system will be. One major limitation of the system is the expectation on the accuracy and availability of the data being used. The data used to support this system will consist of data collected from various locations, such as hospitals, testing facilities, and public health departments, etc. Unfortunately, because these data come from different sources, variations in timing, delay, and/or incompleteness will all lead to incorrect analysis and unreliable outcomes. Another major limitation will be in relation to the protection of the data. Because this system will use confidential patient data, improper use or lack of protection of that data will potentially expose the data to theft or unauthorized use. Finding a balance between accessibility and protection of this data will continue to be a significant challenge.

User participation plays a large part in how effective a system is, especially for contact tracing, where low user participation or incorrect data entered will result in reduced effectiveness and accuracy of the system. The system can also have performance issues caused by the need to process large amounts of real-time data if not effectively managed. Technical limitations of technology such as GPS or Bluetooth as location technologies can also contribute to inaccuracies in the tracking results. In addition, integration of data across multiple platforms can create incompatible data. Areas with low digital infrastructure and/or lower-quality internet can also create less usable systems.

Vi. Conclusion and Future Work

As such, the COVID-19 patient tracking system is quite effective at monitoring the progress of patients, as well as providing useful information regarding the infection process, and helping healthcare professionals observe the dynamics. Moreover, the use of data analytics and visualizations helps identify risk regions and provide useful insights into the issue that allow authorities make necessary decisions. Besides, the proposed approach helps facilitate procedures like contact tracing and resource allocation, which become faster and more efficient because of this tool.

Nevertheless, as it has been discussed previously, the efficiency of the tool depends on numerous factors, including accurate data, correct implementation, and active participation of users. Data privacy issues have to be considered too.

Speaking about potential improvements of the system in the future, it should be noted that they are quite numerous and range from incorporating machine learning techniques for estimating spread risks to the use of mobile apps. Furthermore, technologies like GPS and IoT can also be used in this regard to enhance the



efficiency of the tool. Data security can also be significantly strengthened, while other diseases can be monitored in addition.

References:

1. H. Li, Y. Zhu, and Y. Niu, "Contact tracing research: A literature review based on scientific collaboration network," *International Journal of Environmental Research and Public Health*, vol. 19, no. 15, 2022.
2. S. Wang, S. Ding, and L. Xiong, "A new system for surveillance and digital contact tracing for COVID-19 using GPS and network data," *JMIR mHealth and uHealth*, vol. 8, no. 6, 2020.
3. R. Kleinman and C. Merkel, "Digital contact tracing for COVID-19," *Canadian Medical Association Journal*, vol. 192, no. 24, 2020.
4. M. Shahroz, F. Ahmad, and M. Younis, "COVID-19 digital contact tracing applications and techniques: A review," *Sustainable Cities and Society*, vol. 74, 2021.
5. A. Hegde and R. Masthi, "Digital contact tracing in the COVID-19 pandemic: A tool far from reality?" *Digital Health*, vol. 6, 2020.
6. M. Shahroz et al., "Analysis of digital contact tracing systems and challenges," *ScienceDirect*, 2021.
7. L. Ferretti et al., "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing," *Nature*, vol. 368, 2020.
8. Public health effectiveness of digital contact tracing in COVID-19 pandemic," *PubMed Central (PMC)*, 2021.
9. Z. Peng et al., "BU-Trace: Privacy-preserving mobile contact tracing system," *arXiv preprint*, 2021.
10. A. Chen and K. Thio, "Drivers and barriers to adoption of digital contact tracing systems," *arXiv preprint*, 2021.
11. P. C. Ng et al., "Personal devices for contact tracing: Smartphones and wearables," *arXiv preprint*, 2021.
12. M. R. Hussein et al., "Digital surveillance systems for tracing COVID-19: Privacy and security challenges," *arXiv preprint*, 2020.
13. World Health Organization, "Contact tracing in the context of COVID-19," 2020.