Received: 11 April 2022 | Revised: 2 August 2022 | Accepted: 8 August 2022

DOI: 10.1002/spe.3144

RESEARCH ARTICLE

WILEY

STOPPAGE: Spatio-temporal data driven cloud-fog-edge computing framework for pandemic monitoring and management

Shreya Ghosh[1] | **Anwesha Mukherjee**[2] | **Soumya K. Ghosh**[3] | **Rajkumar Buyya**[4]

[1]College of Information Sciences and Technology, The Pennsylvania State University, State College, Pennsylvania, USA

[2]Department of Computer Science, Mahishadal Raj College, Mahishadal, West Bengal, India

[3]Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, India

[4]School of Computing and Information Systems, The University of Melbourne, Parkville, Victoria, Australia

Correspondence
Shreya Ghosh, College of Information Sciences and Technology, The Pennsylvania State University, State College, PA 16802, USA.
Email: spg5897@psu.edu

Funding information
Department of Science and Technology, Government of India

Abstract

Several global health incidents and evidences show the increasing likelihood of pandemics (large-scale outbreaks of infectious disease), which has adversely affected all aspects of human lives. It is essential to develop an analytics framework by extracting and incorporating the knowledge of heterogeneous data-sources to deliver insights for enhancing preparedness to combat the pandemic. Specifically, *human mobility, travel history*, and other *transport statistics* have significantly impact on the spread of any infectious disease. This article proposes a spatio-temporal knowledge mining framework, named ***STOPPAGE***, to model the impact of human mobility and other contextual information over the large geographic areas in different temporal scales. The framework has two key modules: (i) *spatio-temporal data and computing infrastructure* using fog/edge based architecture; and (ii) *spatio-temporal data analytics* module to efficiently extract knowledge from heterogeneous data sources. We created a *pandemic-knowledge graph* to discover correlations among mobility information and disease spread, a deep learning architecture to predict the next hotspot zones. Further, we provide necessary support in home-health monitoring utilizing Femtolet and fog/edge based solutions. The experimental evaluations on real-life datasets related to COVID-19 in India illustrate the efficacy of the proposed methods. STOPPAGE outperforms the existing works and baseline methods in terms of accuracy by ≈(18–21)% in predicting hotspots and reduces the power consumption of the smartphone significantly. The scalability study yields that the STOPPAGE framework is flexible enough to analyze a huge amount of spatio-temporal datasets and reduces the delay in predicting health status compared to the existing studies.

KEYWORDS
COVID-19, deep learning, healthcare, Internet of Spatial Things (IoST), knowledge graph, pandemic, spatio-temporal data

Abbreviations: IoST, Internet of Spatial Things; IoT, Internet of Things; PKG, pandemic-knowledge graph; QoS, quality of service.

1 | INTRODUCTION

The significant growth of global travel, improved communication, urbanization, and greater exploitation of the natural resources have escalated the likelihood of outbreak of infectious diseases in a larger geographical scale.[1] These fatal infectious diseases pose real threats for public health. The government agencies require effective health measures in such pandemic situations. In this context, the global outbreak of infectious disease COVID-19 (coronavirus disease 2019), caused by SARS-Cov-2, has swept 200+ countries or territories and affected more than 490+ million people, and over 6.39 millions have died (as in the last week of July 2022). This human-to-human disease transmission is highly contagious, and it is already observed that traditional infection-control or public-health measures to combat COVID-19 are inadequate. In the pandemic situation, the governments had already taken various measures and policies such as travel-restriction, lockdown* of several regions, self-quarantine to control the rapid growth/spread of the pandemic. However, the threat of second and third waves of this pandemic still persists in several countries of the world.

The rapid development and emergence of the Internet of Things (IoT) have significantly improved the living quality of people by connecting billions of devices. The IoT devices need to send data to cloud servers frequently for processing and analyzing the accumulated data. However, this increases the delay, therefore affects the quality of service (QoS). Here, edge and fog nodes extend the cloud computing functionality by processing, analyzing, and storing the information near the end-user.[2,3] The combination of these emerging technologies facilitate several time-critical applications, namely, continuous patient monitoring in home, ambulance, and hospital,[4,5] as well as assisting users in emergency situations such as disasters, pandemics.[6,7] In recent times, there is a growing need for analyzing spatio-temporal datasets for extracting meaningful information and providing location-aware services,[8] such as trip-planning, next location prediction from sparse datapoints[9] weather forecasting, and even health management. Notably, one of the significant aspects of epidemiological analysis is retrieving the correlations among the people and disease spread in spatial and temporal dimensions to adapt the countermeasures effectively. For instance, it was found in spatio-temporal data analytics[10] that the source of the Cholera outbreak (in London, 1854) was contaminated through bore wells. The finding was beneficial to combat the spread of Cholera. To this end, the Internet of Spatial Things (IoST) combines IoT with spatial context,[11] where the location information of the objects plays an important role. To fight against the pandemic, spatio-temporal information and health data need to be integrated and analyzed to predict the spread of the disease. This can be beneficial for assisting users and stakeholders, such as imposing region-based lockdown. Here, we propose the term *Internet of Spatio-Health Things (IoSHT)* as integration of IoST with IoHT (Internet of Health Things), where all the health-related information is combined with the location data. This work mainly focuses on how novel *spatio-temporal data analytics method* is beneficial in deciding strategical administrative planning to enhance the preparedness to curb the pandemic. Undoubtedly, the entire world is suffering an enormous amount of challenges from the pandemic situation caused by COVID-19. The two key challenges related to pandemic are:

- (Q1) How to share information about available medical facilities and disease spread patterns in varied locations?
- (Q2) What is the most effective measure to analyze available data sources and predict next probable hotspots to further enhance the preparation for a pathological disease outbreak?

To resolve $Q1$, efficient data infrastructure is required, which can store, manage, and share authentic information amongst stakeholders. Further, a novel spatio-temporal data analytics module is required for understanding the impact of human movement in the disease spread ($Q2$). To this end, we propose a *Pandemic-knowledge graph* (PKG) to capture the interdependence and connections among several entities and users' movement log and subsequently identifying the hotspot/containment zones. STOPPAGE is specifically designed and developed to mitigate the challenges during pandemic or outbreak of infectious diseases in a large scale. For instance, it helps to model human mobility and interaction patterns and identifies hotspots ahead of time effectively. The knowledge graph is represented as a multi-relational graph consisting of entities as nodes and relations as edges, and support several artificial intelligence (AI) related applications. However, it is difficult to represent mobility and other contextual information utilizing the conventional knowledge-graph

*Lockdown: International/National border closures, restrictions of public activity (school and business closures), and overall movement restriction whenever possible.

that changes in temporal scale. On the other side, *deep learning* is the most feasible solution to extract the correlations among several factors such as population density, mobility information and other data-sources. The work is carried out considering two major challenges (Q1 and Q2). Our major contributions consist of novel spatio-temporal data mining (pandemic knowledge graph and deep learning method) and designing the overall architecture to support services provisioned to the users in the exigency time (pandemic).

The full lockdown measure (complete restriction of the movements) has a drastic impact on both local and global economies.[12] This article analyses the impact of the movement in two phases in the context of India: (a) In the pre-lock down phase when International travel was allowed. It analyses the in-flow of International flights in several regions of the country, and the cases reported within a temporal buffer of the visits. The approximate count of the travellers and air-crafts in the time span of March 10 to March 21, 2020[†] and the reported cases[‡] have been analyzed; (b) In the next phase, we develop a disease spread module based on the factors such as population density, changes of mobility patterns[§], POI (point-of-interest) information and so forth. Here, we have proposed a variant of co-occurrence pattern where the mobility information has been augmented.

The major significance and contributions of **STOPPAGE** are as follows:

(1) Spatio-temporal data analytics

- **PKG:** We propose a novel time-dependent *PKG* based on the active cases, POI, road, air-travel connectivity index, and users' movement history. *PKG* captures correlations and impacts among locations, temporal information, movement semantics, and infectious disease spread.
- The article proposes a **deep learning network** which incorporates the spatio-temporal data instances and finds out the probable next hotspot zones by mining the inherent knowledge. To be specific, the deep learning architecture is capable of incorporating the impacts of human mobility and other spatio-temporal features of PKG, and finds out the hotspots effectively.

(2) System architecture

- **IoSHT**: An end-to-end framework named, IoSHT is designed and developed. The proposed framework stores, manages and analyses the pandemic related data effectively and assists in taking effective decisions. IoSHT is capable of collecting BAN (body area network) information and other contextual data such as environment-temperature, location data, mobility information and so forth.
- **Spatial data infrastructure**: The article develops backbone spatial data infrastructure for data sharing, service orchestration and spatial service search and discovery in minimum delay and less manual intervention. To the best of our knowledge, this is the first work to design and develop the underlying spatial data infrastructure in a systematic way[¶]. This can be beneficial for the research community to develop other large-scale spatial services.
- **Delay-aware service provision**: Femtolet[13] is used for faster reporting of patient information of hospitals. Femtolet is a small cell base station that has storage and computation ability. The users registered under a Femtolet can make and receive voice calls as well as can offload their data and computation inside the Femtolet.[13] A Femtolet provides the facility of communication and computation both to the users registered under the Femtolet. In the proposed framework we have used Femtolet inside hospital and home for reporting patient information. The delay and power consumption of the user device for health status monitoring have been determined and compared with the cloud only system to show the efficacy of the proposed framework.

(3) Experimentation with real-life dataset (COVID-19)

- The experimentation have been carried out using the real-life datasets available in India. The road-network, POI information, and air-traffic datasets have been collected from OpenStreet Map (OSM), Google Map

[†]The international flights were suspended in India from March 22, 2020 onwards.

[‡]https://api.covid19india.org/

[§]Google mobility data: https://www.google.com/covid19/mobility/

[¶]We have adapted the underlying framework from http://ggim.un.org/IGIF/part2.cshtml. This provides the theoretical concept, and ours provides proper development and implementation.

Services and crawling web-pages of Airports Authority of India (AAI). The information about the testing facilities and hospitals have been collected from Indian Council of Medical Research#. The changes of mobility information have been analyzed from Google Mobility Report. Finally, we have also accumulated individuals' mobility history from Google Map Timeline from specific regions of India to evaluate the efficacy of *STOPPAGE* framework.

- The performance of STOPPAGE using real-life data depicts the credibility of our proposed system architecture and data analytics techniques.

The rest of the article is organized as follows. Section 2 summarizes the related works in this domain. Then, the system architecture and the deep learning module have been presented in Section 3. Then we present a real-life example of home-health monitoring system and usages of Femtolet to reduce the reporting time in Section 4 followed by the experimental evaluations in Section 5. Finally, we conclude the article with research challenges and opportunities in this domain in Section 6.

2 | BACKGROUND AND RELATED WORK

2.1 | Background

Internet of Spatio-Health Things (IoSHT) is an emerging field of IoT, where integration of IoHT and IoST is carried out to improve the health care system. In IoHT the IoT devices are integrated with mobile technologies to process and exchange data for monitoring health condition of individuals.[29,30] While several IoT-enabled medical devices, such as, BAN accumulates health data and provides insights about symptoms and trends. The information-exchange, communication, and interoperability of IoT devices make healthcare services much more effective. Inside the healthcare centers, IoHT can monitor vital signs of a patient.[29] However, the COVID-19 pandemic situation cannot be managed by only IoHT solutions, it also requires the monitoring of human-to-human transmission of the disease, and understanding disease spread patterns over time and space. To this end, IoHT needs to be integrated with IoST to monitor the overall spread of the disease and reduce the transmission rate by early identification and taking preventive measures. For instance, identifying hotspot zones and disease spread patterns over geographical location and temporal scale may help in taking early administrative measures. Further, IoT devices associated with real-time location tracking of medical equipment such as oxygen pumps, wheelchairs, defibrillators make medical resource management efficient in this pandemic. It is evident that location-information has a huge significance to curb the pandemic. Obvious that mobility is also an important factor for spreading infectious diseases. *Internet of Spatial Things (IoST)* is a new domain and only a very few works[2,11] have been carried out in this direction. In this article, we introduce *Internet of Spatio-Health Things (IoSHT)*, where individual health monitoring and assistance can be anytime and anywhere, as well as effective information-sharing can be provisioned. Next, pandemic/disease spread pattern are analyzed to enhance the preparedness (e.g., prediction of hotspot zone can be helpful in acquiring medical resources a priori, restricting movement of specific region or make residents aware).

The onset of COVID-19 resulted in large volume of research papers proposing several methods to help in governments efforts to curb the pandemic. While there are numerous articles dealing with the pharmaceutical methods, we emphasize more on analysis of contextual information collected from mobile-phones and other sources, such as movement patterns, cumulative count and so forth to predict the future spread of the disease.

2.2 | Related work

Table 1 shows four major features and the existing approaches in the context of COVID-19. Data management and information retrieval is one of the major step in combating the pandemic. In recent decade, deep learning models have gained remarkable achievement in modeling, clustering, anomaly detection, and predicting trends of spatio-temporal data instances.

#ICMR:https://www.icmr.gov.in/

TABLE 1 Comparison of significant existing works and **STOPPAGE** framework in the context of COVID-19.

SI	Publication reference	System	Feature A	Feature B	Feature C	Feature D
[I]	Wang et al.[14]	Comprehensive knowledge discovery framework to extract fine grained knowledge elements from scientific literature, and a case study on *Drug Repurposing Report Generation* has been presented.	✓	✗	✗	✗
	Esteva et al.[15]	Retriever-ranker semantic search engine to resolve complex queries to find scientific answers.	✓	✗	✗	✗
	Fernandez et al.[16]	Cause-and-effect network generated from scientific literature and presented a knowledge graph. The article analyzes comorbidities, symptoms, and discovered over 300 candidate drugs for COVID-19.	✓	✗	✗	✗
	Xu et al.[17]	Formalization and extraction of meaningful insights from the PubMed dataset and generation of knowledge graph using BioBERT deep learning method.	✓	✗	✗	✗
	Reese et al.[18]	Present flexible framework to integrate biomedical data for producing knowledge graphs (KGs).	✓	✗	✗	✗
[1] Most of the works deal with knowledge mining related to pharmaceutical data, drug discovery, or extracting information from scientific literature. No existing works deal with information storage and knowledge extraction from heterogeneous data sources, like travel-logs, aggregate movement information, COVID-19 statistics in different spatio-temporal granularity and find the correlations amongst them.						
[II]	Kraemer et al.[19]	The spatial distribution of COVID-19 cases in China and the mobility traces are analyzed.	✗	✓	✗	✗
	Pepe et al.[20]	The daily time-series of three different aggregated mobility metrics have been presented. The authors use 170,000 smartphone users' data to monitor the impact and assisting in making decisions.	✗	✓	✗	✗
	Huajun et al.[21]	Identification of suspected infected crowds based on the human movement trajectories.	✗	✓	✗	✗
	Hamda et al.[22]	A mathematical modeling with the movement dynamics and disease outbreaks have been proposed. The work helps in taking measures of social distancing and other policies in the 25 counties in the USA.	✗	✓	✗	✗
[III]	Rahman et al.[23]	Distributed deep learning framework for COVID-19 diagnosis in 5G network at the edge.	✗	✗	✓	✗
	Kapoor et al.[24]	Forecasting approach using graph neural networks and mobility data for COVID-19 case prediction.	✗	✗	✓	✗
	Luz et al.[25]	An efficient method of COVID-19 screening in chest x-rays in less time and computational cost.	✗	✗	✓	✗
[II]-[III] To the best of our knowledge, there is no deep learning module which analyzes the movement patterns and other contextual facts to understand the impact of the disease spread and predicts hotspots efficiently.						
[IV]	Tuli et al.[26]	Cloud computing based framework to predict growth of the epidemic and appropriate strategies have been presented. The authors present a prediction framework.	✗	✗	✗	✓
	Md et al.[27]	Utilize mobile and fog computing to trace and prevent COVID-19 community transmission	✗	✗	✗	✓
	Adarsh et al.[28]	Multi-layered architecture to collect real-time information from drones and utilize in disease monitoring, control, thermal imaging, social distancing, and statistics generation.	✗	✗	✗	✓
	STOPPAGE (proposed framework)	Cloud-fog-edge enabled collaborative framework conducive to construct a novel time-dependent Pandemic-knowledge graph (PKG) based on the spatio-temporal data. Presenting deep learning based analytics module to find out hotspots by mining contextual data.	✓	✓	✓	✓

Abbreviations: A, efficient COVID-19 related knowledge retrieval and management; B, movement data analysis to find out the impact; C, deep learning module to analyze pandemic data; D, cloud-fog-edge enabled framework deployment for faster response.

A comprehensive insight on how IoMT (Internet of Medical Things) systems can be beneficial in the context of COVID-19 along with the architecture, tools and technologies is discussed in Reference 31. An energy efficient dengue fever diagnostic architectural paradigm leveraging cloud and fog technology is presented in Reference 32. A crowdsourcing approach for green healthcare service is provisioned by *RESCUE* framework.[33] Vedam et al.[34] present an agent-based framework along with population mixing algorithm to demonstrate the treand of COVID-19 considering several factors such as social distancing, vaccination isolation, containment zones and so forth. Marcello et al.[35] have stated that large scale collection of data and maintaining the privacy and public trust is one of the prime steps. In this regard, *crowdsourcing* is another crucial factor where a task is performed by a designated agent followed by outsourcing the same to a larger audience. A novel algorithm, named *CovidCrowd* is proposed in Reference 36 to collect contact tracing data from crowdsourcing perspective. Song et al.[37] propose a novel co-evolutionary transfer learning approach for predicting the demands of medical materials to prevent and control COVID-19. The paper authors evaluated their method using two stages of COVID-19 spreading in China and their method outperforms the state-of-the-art transfer learning and multi-task modules. Yang et al.[38] have formulated the research problem from the context of visual data, and proposed a novel architecture namely *V-HIoT* for IoT assisted healthcare system.

There are several works on summarizing the knowledge from drug discovery or other pharmaceutical methods.[16,17] Also, researchers[14] have put significant efforts to summarize information from scientific literature by proposing knowledge graph. There are works on aggregate movement dynamics and understanding the impact in disease spread.[19,20] Huajun et al.[21] have presented a novel trajectory pruning method to find out suspected crowds in a region. An interesting study[39] explores the causal impact on mobility and COVID-19 incidence based on public policy decision of closing all bars and restaurants down in Cataluña, Spain. The authors have demonstrated interesting insights from economical standpoint, public reaction, and incidence rate of COVID-19. There are also few works leveraging cloud-fog paradigm for contact tracing, real-time drone based system and analyzing the growth of the disease.[26,27] A novel framework integrating web of things and computer vision algorithms has been proposed.[40] The system achieved better performance in terms of latency, response time, and network latency. Abid et al.[41] propose a blockchain-enabled framework, named, *NovidChain* for COVID-19 test/vaccination certificates to ensure data integrity and immutability. Ghosh, and Mukherjee[42,43] present cloud-fog-edge enabled framework named *STROVE* to provision fast healthcare services leveraging femtolet and mobility-based solutions. Mohan et al.[44] proposed a hybrid model for forecasting impact of COVID-19 in long-term both in India and global level. A lightweight framework named *FogBus* is presented in Reference 45 for integrating IoT-enabled systems, fog and cloud infrastructure using blockchain technique. An IoT-fog-cloud based architecture to develop COVID-19 monitoring system has been proposed.[46] However, our work emphasizes on spatio-temporal aspect both from system and application perspective. It may be concluded that machine learning and movement data analytics may act as major scientific tools for combating the spread of COVID-19 pandemic. However, to the best of our knowledge, no existing works has extensively studied the movement patterns and contextual information in a region for predicting hotspots effectively ahead of time. Our work explores and analyses the human movement related data proposing novel knowledge graph (PKG) and deep learning architecture. Further, STOPPAGE offers detailed discussion on how spatial data infrastructure can be developed and implemented which in turn efficiently stores and manages huge volume of spatio-temporal information (in our case, COVID-19 related data) for better decision making.

3 | STOPPAGE: SYSTEM ARCHITECTURE

The proposed framework *STOPPAGE* has two major modules, namely: spatial data infrastructure and a data analytics engine running over the spatial data infrastructure. Figure 1 illustrates the overall workflow and modules of *STOPPAGE*. The spatial data infrastructure consists of backbone IoT network and geospatial resource management module. The spatio-temporal data analytics engine analyses the COVID-19 data and mobility related information to find out the probable hotspot zones by deploying deep learning architecture. Furthermore, *STOPPAGE* proposes the use of Femtolet to report any health events in minimal latency. Also, the health status monitoring is provisioned in the framework by a common API-endpoint. Figure 2 illustrates the overall block-diagram of data analytics engine of STOPPAGE. It has three major components (i) pandemic knowledge graph (PKG) construction, (ii) deep learning based movement analytics, and (iii) faster reporting of COVID-19 events and data sharing using cloud-fog-edge based paradigm. Figure 3A represents the patial data infrastructure and Figure 3B illustrates the deployment of IoSHT in cloud-fog-edge structure.

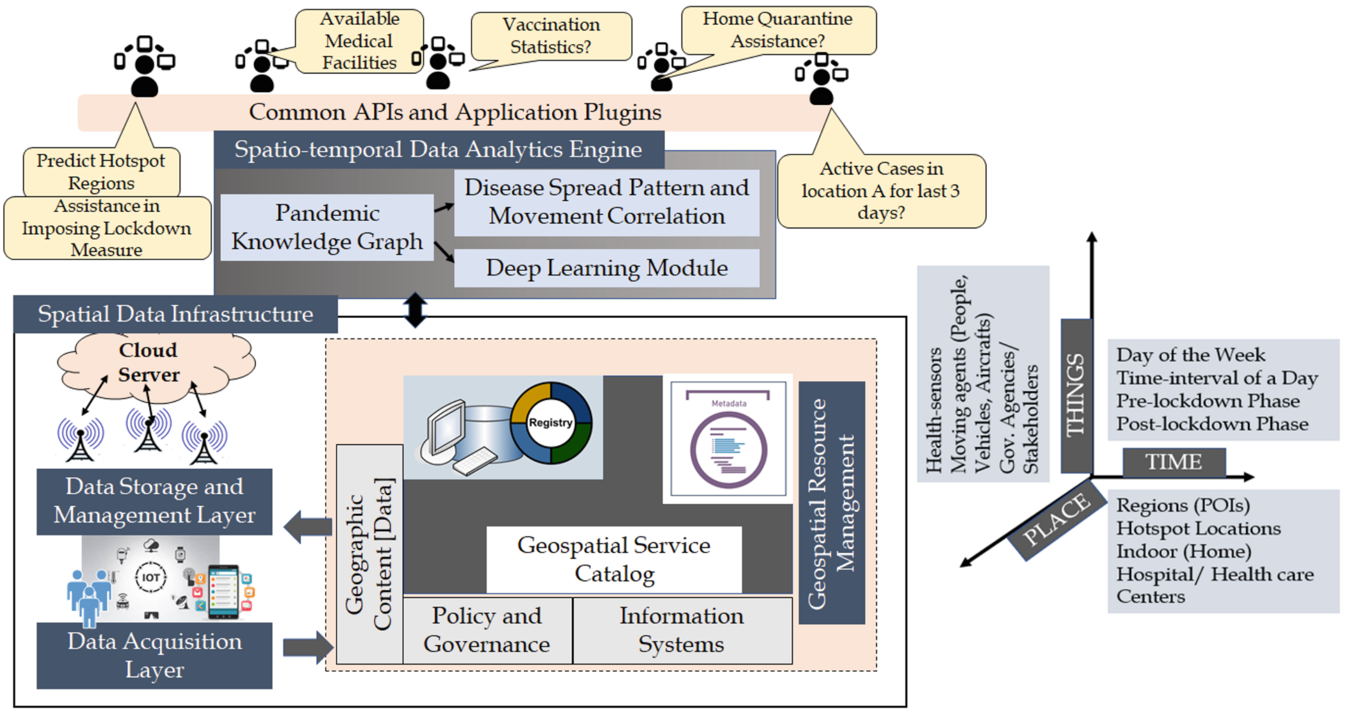


FIGURE 1 **STOPPAGE**: Overall workflow

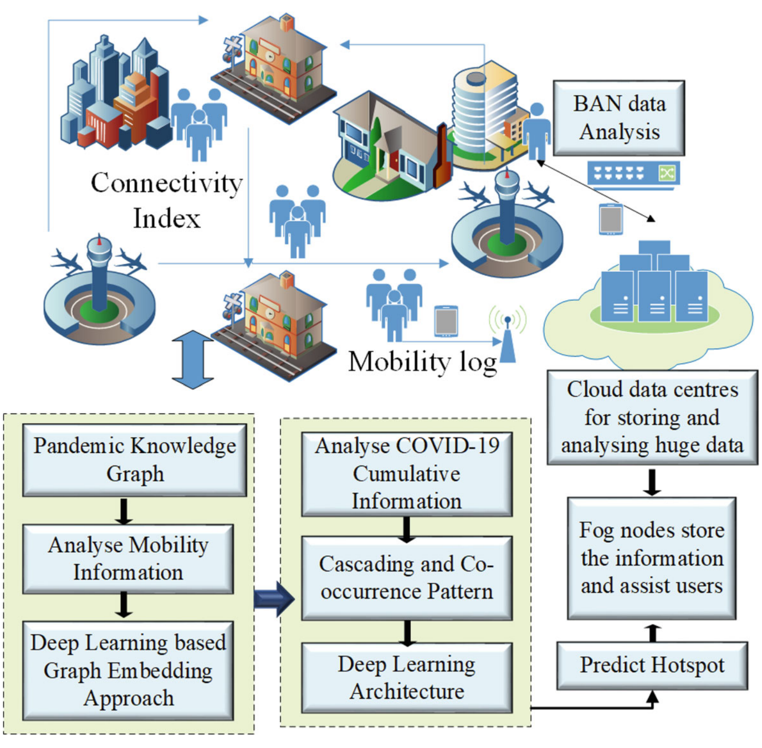


FIGURE 2 **STOPPAGE**: Spatio-temporal data analytics module

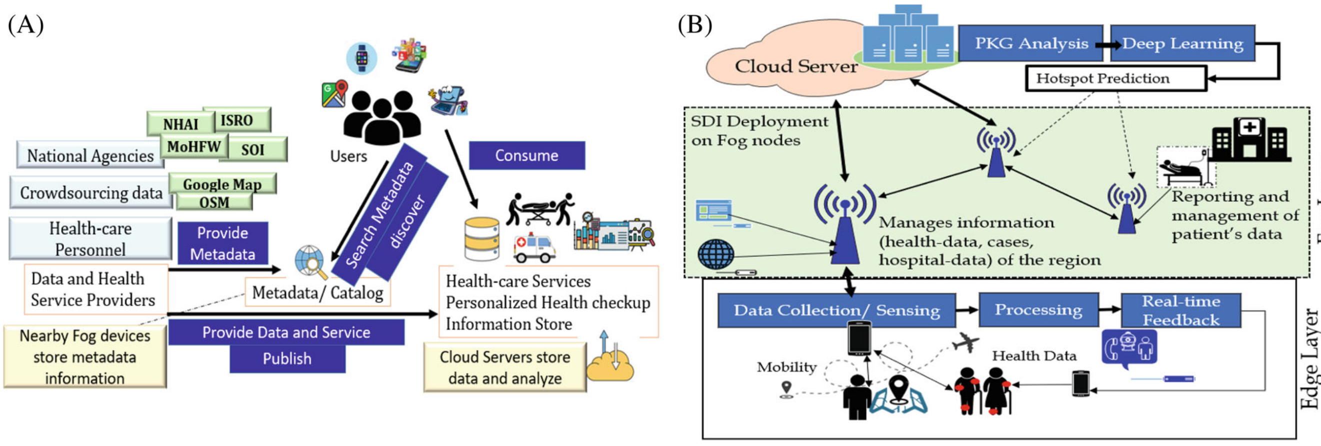


FIGURE 3 (A) SDI (spatial data infrastructure) in the context of IoSHT, (B) Cloud-fog-edge hierarchical structure

3.1 | Spatial data infrastructure (SDI): An enabler of IoSHT

Spatial data infrastructure (SDI)[‖] denotes a framework consisting technologies, policies, and orchestration method to create, exchange, and utilize geospatial information and services among the community.[47] The goals of this framework are:

- ease of search and discovery of geospatial services and information;
- reduce data duplication of information among the national agencies (government);
- seamless data sharing technique;
- maintaining the data integrity and privacy.

Further, SDI enables the communication between the repositories of cities, states, countries, and industries to share the data and services. In other words, the difficulty in sharing and accessing different format/structure of data is resolved by SDI. SDI refers to a framework consisting technologies, policies, and orchestration method to create, exchange, and utilize geospatial information and services among the community. Several countries maintain their own national SDI (or NSDI)[**] for enabling interoperability between different spatial datasets (collected and maintained by federal/government agencies). NSDI maintains country specific rules/policies for data sharing and helps in various decision making processes involving geo-spatial datasets. In this work, we are specifically resolving the challenge of data sharing and retrieval when two government agencies or stakeholders have the permission to share information. For instance, different departments do not follow the same data storage format, also do not have the metadata information. There is also need of integrating two or more spatial services. The problem becomes more challenging with large spatio-temporal repositories of cities, states and industries. The major objective is to automate the seamless data and service sharing, reducing the duplicate information, and maintaining data integrity. To resolve this, we have designed and developed the architecture and implemented the system which can provide all the functionalities of NSDI or IGIF (Integrated Geospatial Information Framework) in a scalable manner.

In the context of COVID-19, the stakeholders, such as health-department, transportation authority, or epidemic-control team along with the end-users need to seamlessly share data and communicate for taking decision such as distribution of medical resources, region-based lockdown measure. Without a proper underlying infrastructure of data sharing and managing mechanism, any epidemic-control measure or policy is difficult to implement. Therefore, in our proposed IoSHT framework, we have adapted SDI as the backbone infrastructure. The SDI in the context of IoSHT is represented in Figure 3A. To combat the disease spread and take countermeasures, we need to analyze

[‖]The term spatial data infrastructure (SDI) was coined in 1993 by the U.S. National Research Council.
[**]NSDI: https://www.fgdc.gov/nsdi/nsdi.html

heterogeneous data-sources. The authentic data about the country's population, demography, health care centers and so forth can be found from the government departments. In the context of India, we collect demography and population data from SOI and open-source platform[††],[‡‡], health-related information from MoHFW[§§], transport data from NHAI[¶¶] and raster information (satellite data) from ISRO[##]. All of these datasets of country's population, demography information of the people, health care center related information, traffic data are useful to analyze the disease spread patterns. Further, the crowdsourcing data from Google Map, OpenStreetMap (OSM), or Android application are also beneficial to get the present status (active cases, number of migrated people, supply of medical resources, vaccination) of the region. The data and health service providers provide the metadata information to the catalog. This catalog is maintained in each of the nearest fog devices, where the data or the service is generated. User can search the catalog any time, and extract the metadata of the information or service he/she requires. Once, the user selects the required service/information, he/she can consume the service. The services can be of different type, like, accessing information regarding vaccination slot at nearest healthcenter, availability of hospital-beds at minimum delay, personalize health checkup or knowing the probability of infected from historical movement data. As it can be seen from the architecture, in the proposed *STOPPAGE* framework, the services are: (i) contact the health care center for medical help; (ii) personalize health-checkup by sending the health-parameters' values collected from the IoT devices/BAN; (iii) get the present disease spread information of any region; (iv) find the risk by analyzing his/her movement history. All of these analytics are computed in the cloud servers, and the user gets the result through the web-dashboard or android app.

3.2 | Health status monitoring and reporting using fog-edge solution

In the previous section, we have discussed the proposed system architecture of *STOPPAGE*. For current health status prediction of people in a region, we have to consider the reports of health care centers/hospitals. Here, Femtolet is used in the health care centers for faster reporting of number of patients who are affected, deceased, and cured. This information is very important to monitor the current status of the pandemic in a geographic region. Femtolet is small cell base station which has storage and processing ability.[13]

In the proposed model, in each ward[‖‖] of the hospital, Femtolets are deployed to store and process the information of the patients currently admitted in that ward. The number of Femtolets to be allocated depends on the coverage area, storage capacity, data processing speed of the Femtolet, and the number of patients currently admitted in that ward. The health status of a patient is collected using BAN and sent to an edge device (for each patient BAN and an edge device are used). The edge device (ED) accumulates the data and sends to the Femtolet, connected to the edge device. The Femtolet processes the incoming data and stores it under the respective patient registry (for each patient a data registry is maintained). If any emergency situation arises, the Femtolet sends signal to the connected alarm which after receiving the signal starts ringing. The health personnel of that ward then takes required action. The Femtolet periodically sends the patient health status to the private cloud servers of the health center, which maintains the information regarding the number of patients cured, deceased, and affected. From the private cloud servers the details of the number of patients affected, cured, and deceased in the hospital with respect to the disease are reported to the public cloud servers periodically. The overall architecture of Femtolet based health care model is shown in Figure 4. As the public cloud has the information regarding the number of patients admitted, cured, and deceased due to COVID-19 in each hospital, the total number of admitted, cured, and deceased patients (considering all the hospitals' information) due to COVID-19 is determined. The people can get this information if he/she wants to know the information of active or suspected cases in his/her region of residence. However, the information of each hospital will not be disclosed without authorization due to privacy management. The cloud already has the information regarding hotspot zone (in Section 3.4 hotspot identification is discussed). By considering both the information, people can be notified about the current scenario.

[††]Survey of India: http://www.surveyofindia.gov.in/
[‡‡]https://www.diva-gis.org/gdata
[§§]Ministry of Health and Family Welfare of India: https://www.mohfw.gov.in/
[¶¶]National Highways Authority of India: https://nhai.gov.in/
[##]Indian Space Research Organization: https://www.isro.gov.in/
[‖‖]Block forming a division of a hospital shared by patients who need a similar kind of care.

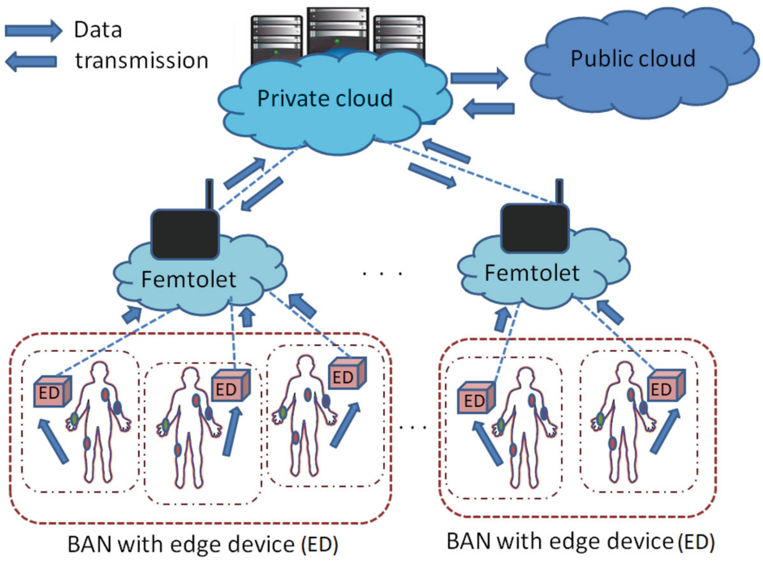


FIGURE 4 Smart health care framework based on Femtolet

3.3 | Spatio-temporal data analytics engine: Pandemic-knowledge graph (PKG) construction

This section describes the features of *PKG* and how it is constructed from the available dataset. This is a multi-layer network which captures the movement behavior individuals, the overall movement semantics of the ROI (Region-of-interest). PKG also captures and stores the statistics related to COVID-19, such as growth pattern of active cases, changes of aggregated mobility patterns, and available medical facilities and so forth. The intuitions behind constructing knowledge graph are: (i) transforming the underlying relations of mobility semantics in a machine-readable format to support information retrieval and query-processing, (ii) the complex spatio-temporal mobility dataset can be represented by graph structure effectively instead of other storage, (iii) due to the dynamic nature of movement patterns, and the correlation with disease spread pattern, SQL-based processing fails to provide flexible storage scheme, and finally (iv) knowledge graph provides a higher level abstraction of information, which may help to extract more complex and previously unknown interrelations merging more than one relations of the graph. For example, merging the relations (visit, stay locations of an infected person) over time, we can detect possible suspected cases (people with overlapping movement history). Therefore, it facilitates building a semantically enriched knowledge-base.

3.3.1 | Pandemic-knowledge graph (PKG)

The PKG is formulated as triplet of $\langle s, r, o\rangle$, where $s \subseteq \Omega$ and $o \subseteq \Omega$ are *entities* and $r \subseteq \varphi$ is the relation between two entities. The sets of entities and relations are denoted by $\Omega\,(\Omega_1 \cup \Omega_2)$ and φ respectively. Notably, in our PKG proposition, each triplet (or fact) has a time-slot, when the fact is valid. Therefore the facts of *PKG* takes the form of $MF : \langle s, r, o, [t_1, t_2], f\rangle$ where the $[t_1, t_2]$ entry denotes the time-interval when the mobility fact (say, two people meet at location p_a for t time-interval) is true. f is the feature value of the relation. This f is introduced to capture the semantic correlation with the COVID data.

The entities consist of: $U \subseteq \Omega_1$ and $P \subseteq \Omega_2$, where U and P are the set of users and set of place-ids of a ROI. Each entity $pl \subseteq P$ has attributes: POI-type, location (latitude and longitude), enclosing area. Few pl has additional features such as *opening/closing time*, and available facilities. Each entity $u \subseteq U$ has unique user-id, age, gender, residence-area, health-profile, and travel-history. The derived attributes of POIs such as *GPS footprint* or *edge segment* are extracted from the GPS log of the users in the ROI. The facts *connectedBy* and *boundingBox* can be directly computed from the road-network and latitude and longitude information of the POI. Based on the historical movement log the relations $r \subseteq \varphi$ are extracted.

Apart from the mobility fact label, each edge in the graph has a feature value associated, which depicts the probability of the occurrence of the edge. Few such facts are represented as follows:

$$visit : MF(u_a, visit, p_i, [t_1, t_2], f_x) = TRUE. \tag{1}$$

User u_a visits place p_i in the time-interval $[t_1, t_2]$ with a probability of being infected f_x. Here, the feature value of the fact is computed by the graph embedding approach.[48]

$$\begin{aligned}
&group : MF(u_a, group, U' \setminus u_a, [t_1, t_2], f_x) = TRUE \\
&\exists(u_1, \dots, u_n)\exists(p_1, \dots, p_m)\forall_j^n \\
&[MF(u_j, visit, p_1, [t_1, t_1^{'}], f_1) \wedge MF(u_j, visit, p_2, [t_2, t_2^{'}], f_2) \\
&\wedge \cdots \wedge MF(u_j, visit, p_m, [t_m, t_m^{'}], f_m)] = TRUE, \quad m \geq 3.
\end{aligned} \tag{2}$$

The sequence of places $(p_1, \dots, p_m)$ are visited by a group of users (U') in the same time-intervals. f_x may have a value within [0,1).

$$\begin{aligned}
&flow : MF(P_1, flow, P_2, [t_1, t_2], f_x) = TRUE \\
&\exists(u_1, \dots, u_n)\forall_j^n[MF(u_j, visit, p_a, [t_1, t_1^{'}], f_x) \\
&\wedge MF(u_j, visit, p_b, [t_2, t_2^{'}], f_x) = TRUE, \quad n \geq \nu.
\end{aligned} \tag{3}$$

The above mentioned movement flow (MF) is detected from p_a to p_b in a specific time-slot, when a set of users $(u_1, \dots, u_n)$ visit the POI p_a followed by p_b in the same time-interval. The count (n) of GPS footprints satisfying the fact must be greater than a threshold value ν. In other words, if $\geq \nu$ people visit the sequence of POIs in same time-interval, then it is considered that there is a mobility flow from p_a to p_b.

$$hotspot : MF(boundingbox(\{p\}), infected, \{u\}, [t_1, -], co) = TRUE. \tag{4}$$

The bounding box of the set of regions (p) is a hotspot with *co* number of infected people.

$$connectivity : MF(p_a, connect, p_b, [t_1, t_2], f_p) = TRUE. \tag{5}$$

The place p_b can be visited from place p_a in the time-interval $[t_1, t_2]$ where f_p denotes the number of available routes. It may be noted that due to the lockdown measure, this fact changes frequently for all pairs of places. The connectivity index (CI) of a place p is computed based on the *connectivity* and the in-flow and out-flow of traffic from a particular place p. The high connectivity index between two place represents higher mobility flow, and thus can be a potent feature for the hotspot detection.

Next, we define *cascading pattern* (P'_{ca}) and *co-occurrence pattern* (P'_{co}) extracted from PKG. These patterns can be analyzed from the travel history of the persons diagnosed with COVID-19 positive. Cascading patterns represent events whose instances are located together and occur serially. For instance, analyzing the PKG data of a hotspot may reveal that a *group* of users *visit* a POI on a day and spends a particular time-duration within a distance threshold. Also, air-travel of a group of users from a region with high number of active cases and new hotspot emergence in the destination region within a time-span is another example of cascading pattern. Also, P'_{ca} covers patterns such as: higher number of medical facilities in a city-region, more cases of COVID-19 patient transfer and emergence of hotspot. On the other side, *co-occurrence patterns* reflects patterns occurring frequently within a spatial range and a temporal span. This patterns comprise properties from different contexts. For example, given three different contexts, namely, population density, aggregate movement pattern and COVID-19 active/new cases (say, within a spatial buffer of 2 km and time-span of 7 days): $P'_{co} : \{population - density > \delta; aggregate\ movement\ pattern > \gamma; Hotspot\_prob = HIGH\}$. Also, P'_{co} covers patterns such as: festival/gathering events in a region and emergence of new hotspot; or more amount of location-based service requests (ride sharing services/food delivery services) and hotspot detection in the same region within a temporal bound. Algorithm 1 shows the steps of extracting cascading and co-occurrence patterns from PKG, which in turn helps in extracting the knowledge of disease spread, and finding the possible hotspot zones in future. The algorithm needs a threshold value called *pattern participation index (PI)* to prune the irrelevant patterns.

Algorithm 1. Extracting cascading and co-occurrence patterns from PKG

Input: Pandemic knowledge graph (*PKG*), neighbor ST relation ($NR[]$), set of event/fact types (*MF*), pattern participation index thresholds (PI_1, PI_2)
Output: Set of P'_{co} and P'_{ca} patterns

$P'_{ca}[\,] \leftarrow NULL$ *and* $P'_{co}[\,] \leftarrow NULL$ ▷ `Create NULL arrays for storing patterns`
for $i = 1$ to *size* **do** ▷ `size: maximum length of the patterns`
  filter($P'_{ca}[i], PI_1$) ▷ `Remove candidates by analyzing threshold value`
  Genpat(i) $\leftarrow P'_{ca}$ *of length* $(i-1)$ ▷ `Extract new patterns of i length from patterns with length (i-1)`
  $temp \leftarrow ST-Join(P'_{ca}[i], NR)$ ▷ `Perform spatio-temporal join with instances from direct neighbor relation`
  if $checkThreshold(temp) \geq PI_1$ **then**
    $P'_{ca}[i] \leftarrow Append(temp)$ ▷ `Append the new pattern if it satisfy the threshold condition`
  end if
end for
for $t = 1$ $to(n-1)$ **do**
  $te \leftarrow GenCand(PKG(MF, NR))$ ▷ `Generate candidates from PKG using possible facts/events and neighbor relation`
  $P'_{co}[t] \leftarrow sweepLine(te, P_2)$ ▷ `Sweep-line based method to extract events satisfying ST-property and threshold`
end for

$$PI = \min[Pr(P'_{ca}|MF) \text{ or } Pr(P'_{co}|MF)]$$
$$= \min\left[\frac{\text{facts or events participating in } P'_{ca} \text{ or } P'_{co}}{\text{Total no. of facts or events in PKG}}\right]. \tag{6}$$

Here, *MF* represents the mobility facts and events present in PKG. It may be noted that the conditional probability of a pattern; given the possible similar type of events/facts is measured by *PI*. It is also an useful measure for predicting the near future occurrence of a pattern in the spatio-temporal proximity of an observed instance of a participating event-type. We set two threshold values namely PI_1 and PI_2 for pruning the irrelevant patterns from PKG. The neighbor spatio-temporal or ST relation represents the time and spatial constraint on finding the co-occurrence patterns. In brief, the algorithm first extracts the plausible candidates from PKG, followed by pruning technique to eliminate irrelevant patterns. Construction of the PKG consists of following steps: (i) In the first step, the POIs and users along with the attributes are extracted and stored. The information are discovered by geo-tagging step and segmenting the trajectory of individuals at different time-scales. (ii) Next, the links between the entities (or facts) at different time-instances are discovered by knowledge graph embedding approach. The relations such as visit, group, flow and so forth are defined. The movement log is analyzed to measure the plausibility of any such facts. However, the facts are checked over different temporal instances, since the facts or relations of the knowledge graph change with time-instances. The backbone SDI helps in sharing heterogeneous data, maintaining the authenticity and integrity of the data to build PKG, and facilitating services seamlessly to the user.

3.4 | Spatio-temporal data analytics engine: Deep learning architecture to find out hotspots

"You cannot fight a fire blindfolded. And we cannot stop this pandemic if we don't know who is infected."[***] Analyzing spatio-temporal pattern of disease outbreak is an integral part of pandemic control. It helps in identifying spatio-temporal

[***]WHO Director-General's opening remarks at the media briefing on COVID-19, dated March 16, 2020.

hotspots (disease emerging areas), and subsequently assists the planning of emergency measures for monitoring, surveillance, and prevention of the disease spread. Most of the governments have enforced partial or complete lockdown to prevent the disease spread. However, the complete lockdown measure has negative impact on the socio-economic condition of a country, and more feasible solution is required, such as identifying high-risk zones (large number of affected person) and impose lockdown at those regions. In such a scenario, it is absolutely necessary to find the correlation of the disease spread and other additional information from spatio-temporal context. For instance, whether the disease spread is related to population density, demography data and so forth. Moreover, whether specific regions like, railway junction, commercial area is more susceptible to infection—needs to be analyzed.

3.4.1 | Predict the hotspots areas

The primary objective of the analytics module is to identify the high-risk areas of disease spread in a region. We aim to find out the correlation of contextual information such as mobility with the probable spread of the disease.

Of late, *deep learning* has gained significant research interest to utilize the correlations between related tasks and improve the classification accuracy by jointly learning more than one task. Inspired by this paradigm, *STOPPAGE* models the deep learning module to predict the hotspot areas. It may be noted, that in our problem set-up, the mobility patterns have significant impact, and learning the representation of these contextual variables is one of the important aspects of our deep learning architecture. Figure 5 illustrates the recurrent architecture of the proposed deep learning module. In our framework, we have considered the following features:

1. Demographic data (D): Demographic data is the statistical data collected about the characteristics of the population. We have considered *age*, *gender*, *employment status*, and *literacy rate* of people of a region.
2. Population density (PD): It provides the number of people in each region along with the condition of their household (slum area/houseless etc.). The datasets D and PD have been collected from the census information board of India[†††], where district and zone-wise (specific locality) population and demographic data are available.
3. Point of Interest (POI): The POI information, such as the locations of the airports, railway junctions, hospitals, and other commercial areas are analyzed to extract insights into the possible disease spread pattern and in identifying the hotspot/containment zones. The data has been extracted from OSM and Google place API[‡‡‡].
4. Mobility report (MR): The individual and aggregated mobility information are analyzed for predicting hotspots. The aggregated mobility report is collected from *Google Community mobility Report*[§§§]. The individual mobility history is collected from users' smart-device and Google map timeline of individuals[¶¶¶] in the study-region.

All of these information is managed using an efficient storage method. The region is segmented into different grids, and the fog devices store the information of all the grids within its coverage using a hashing scheme.[49] Thus, the information is segregated into different hierarchical segments, and data extraction becomes efficient.

Next, it is important to find out the spatio-temporal correlation (SC) factor of the COVID-19 spread in neighboring regions, which has been computed based on Moran's I[50] as follows:

$$SC = \frac{o}{\sum_a \sum_b w_{ab}} \times \frac{\sum_a \sum_b w_{ab}(v_a - \bar{v})(v_b - \bar{v})}{\sum_a (v_b - \bar{v})^2}. \quad (7)$$

Here, v_a represents the number of events/incidence (here, number of new cases reported in region a) observed in a region a, the mean of the newly confirmed cases in the whole region is $\bar{v}$, the total number of observations is o. w_{ab} represents the spatial adjacency between a and b. Here, we have used different types of adjacency relation instead of only distance measure. The direct spatial distance measures as follows: *(1) two regions share a common border; (2) direct route is present*

†††https://www.censusindia.gov.in/(S(esr3lm45pksguc451d45sp55))/2011census/population_enumeration.aspx
‡‡‡https://developers.google.com/places/web-service/search
§§§https://www.google.com/covid19/mobility/
¶¶¶https://www.google.com/maps/timeline?pb

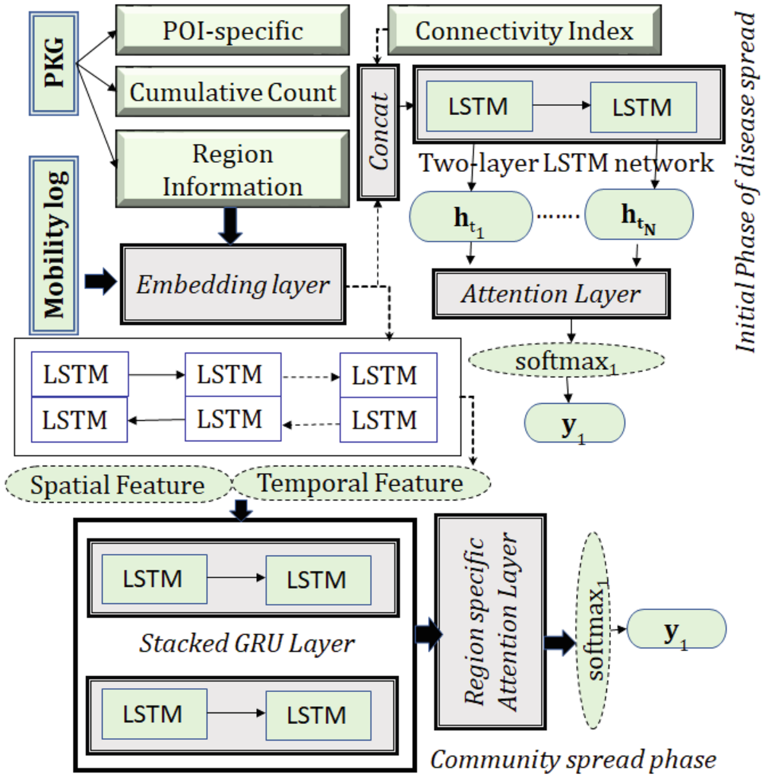


FIGURE 5 Deep learning architecture of **STOPPAGE**

to reach b directly from a. Next, we sort the regions based on their *(3) population density, (4) literacy rate, (5) available medical facilities, (6) aggregate movement flow*. Based on each of the variable, we compute *SC* where a region is adjacent with its previous ranked region and the next ranked region. We compute these values at different time-spans (specifically in each week of the study period).

$$SC = \begin{cases} 0 & \text{No spatial autocorrelation} \\ > 0 & \text{Positive spatial autocorrelation} \\ < 0 & \text{Negative spatial autocorrelation.} \end{cases} \tag{8}$$

As shown in the Equation (8), when the computed value of SC is 0, it represents no spatial autocorrelation. Similarly, the larger absolute SC value demonstrates stronger spatial autocorrelation. This information (SC on six different adjacency metric) is used in the deep learning module to efficiently predict next possible hotspot zones.

Specifically, a trajectory segment/movement history of an individual records the consecutive stay-points in the path of the user, which are continuous spatial locations. Each grid is identified with unique id. Thus, we discretize the continuous spatial location information. We use the skip-gram model to learn the representation of spatial locations $(l_1, l_2, \ldots, l_T)$ in the training dataset:

$$\frac{1}{T}\sum_{t=1}^{T} \log[p(l_{t-c}, \ldots, l_{t-1}, l_{t+1}, \ldots, l_{t+c}|l_t)], \tag{9}$$

which can be written as:

$$= \frac{1}{T}\sum_{t=1}\sum_{con} \log p(l_{t+j}|l_t)$$
$$\text{where context: } -c \le j \le c, j \ne 0, \tag{10}$$

where l_{t+j} is the neighboring location of the present location l_t. Based on the spatial proximity rule that two neighboring locations have similar representation, we use softmax function as defined:

$$p(l_{t+j}|l_t) = \frac{\exp\left(a^N_{l_{t+j}} a_{l_t}\right)}{\sum_{l_i \in l} \exp\left(a^N_{l_i} a_{l_t}\right)}. \tag{11}$$

Thus, the embedding layer converts the spatial locations into:

$$a_{l_1}, a_{l_2} \ \dots \ , a_{l_N}, \tag{12}$$

where N is the number of the regions. Next, we embed the temporal information. The representation of the temporal information should comprises of day of the week, timestamp, and time-duration spent in a location. Here, the skip-gram model is not efficient, and we use paragraph-vector model and get the vector representation of temporal information. After embedding the spatial and temporal features, we deploy a bidirectional LSTM layer to capture the shared information of both the tasks. The basic building block of the LSTM layer is as follows:

$$\begin{aligned} & i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i),\ f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ & o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o); \\ & c_t = f_t \cdot c_{t-1} + i_t \cdot \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \end{aligned} \tag{13}$$

where the input, output, and forget gates are represented by i, o, and f, respectively. The hidden representation is denoted by h_t. Since, *STOPPAGE* uses a bidirectional LSTM, the output of the layer is modified as:

$$h_t = \overrightarrow{h_t} + \overleftarrow{h_t}. \tag{14}$$

Here, the output from the forward and backward propagation layer are represented by $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ respectively.

The next layer utilizes a *gated recurrent unit (GRU)* which is similar to *LSTM*. This layer is used for extracting the impact of the other contexts such as demography, age and aggregated movement and so forth. GRU has reset and update gate which are formally defined as:

$$\begin{aligned} & z'_t = \sigma(W'_z \cdot [h'_{t-1}, x'_t]) \ \ r'_t = \sigma(W'_r \cdot [h'_{t-1}, x'_t]) \\ & h'_t = (1 - z'_t) * h'_{t-1} + z'_t * \tanh(W'_h \cdot [r'_t * h'_{t-1}, x'_t], \end{aligned} \tag{15}$$

where W'_z, W'_r, and W'_h are weight matrices. The *update gate* (z'_t) helps to extract the required information from the past time-step and pass to the future. On the other hand, *reset gate* (r'_t) is used to determine how much past information needs to be eliminated. Therefore, in the proposed model, the GRU is capable of filtering and storing information by utilizing the reset and forget gates. Here, we use specific information such as SC, P'_{ca}, P'_{co} and so forth. This is crucial to capture the different phase of the disease spread, since GRU layer eliminates the vanishing gradient problem, but pass the relevant information to the next steps of the network. Note that in the proposed approach, deep LSTM architecture is used, allowing the network to learn at different time scales over the input. Furthermore, they can make better use of parameters by distributing over the space through multiple layers.

Another important module of the architecture is *attention mechanism*. It is used to capture the relationship between all of these contextual factors and hotspots in different spatio-temporal resolutions. Here, we have used the dot product attention function f_{att} and the representation is defined as:

$$r^{att} = \sum_{i=1}^{N} \frac{\exp(f_{att}(h'_{t_i}, m_{t_i}))}{\sum_{i=1}^{N} \exp(f_{att}(h'_{t_i}, m_{t_i}))} h'_{t_i}. \tag{16}$$

Here, m is the vector representation of spatio-temporal features from the previous layer. The GRU input layer is replaced by the weighted representation (r_{att}). On the other side, air travel connectivity index is utilized for the initial phase of disease spread. Here, *attention layer* provides more *attention* to the long-distance travels and connectivity between regions. For both the phases of disease spread, the final layer is softmax layer where all the outputs from different layers are fed.

$$y' = softmax(W'h' + b'). \tag{17}$$

The network is trained to minimize the cross-entropy loss as:

$$L(y, y') = -\sum_{i=1}^{TS}\sum_{j=1}^{Cl} y_i^{'j} \log(y_i^j). \tag{18}$$

The number of training samples and number of classes are represented by *TS* and *Cl*. We define 4 classes here: (i) > 20 cases within 500 m; (ii) > 50 cases within 1 km; (iii) < 5 cases within 1 km and (iv) < 10 cases within 2 km. We define (i) and (ii) as hotspot zones. y' is the ground truth and y is the prediction probability.

It may be noted that, in the *MR* feature analysis, we also combine the information if any neighborhood region was a containment zone/hotspot zone (high infection rate) in last 14 days. This information is stored in the respective fog devices of the region. Thus, the architecture encodes and learns different mobility semantics and other parameters at varied contexts and finds out whether the region is the next hotspot zone.

4 | APPLICATIONS OF STOPPAGE

In this section, we show how *STOPPAGE* can be utilized to combat pandemic situation. We have discussed two use-cases, and how data analytics engine and SDI of *STOPPAGE* can be utilized in such applications. An Android application along with the API-endpoints have been developed to retrieve and share information. We have also shown delay and power consumption to illustrate the feasibility of *STOPPAGE* in terms of latency and energy consumption.

4.1 | Use-case: Find the suspected crowd

There are varied contributing factors in spreading the disease. The probability of infecting other people are much higher when the asymptomatic person is in the workplace or shopping mall compared to when he is in a less crowded park or open playground[###]. Also, the time-duration of contact with other person is also an important factor. Therefore, *STOPPAGE* can also use the individual mobility history to find the suspected people from the PKG analysis. If any of the users is tested COVID-19-positive, then the information is sent to the nearby fog device and based on the PKG analysis, all other users are notified. Alongside, using the reverse pruning technique, we can also find out the asymptotic person by analyzing the mobility data. Analyzing and assisting people in real-time is the major challenge here. The fog devices keep track of a particular locality and assist users within that region. Since, the aggregated movement dataset is huge, any compute-intensive task (training phase of deep learning or mobility knowledge graph embedding) is carried out in the distant cloud servers. However, to reduce the overall execution time, fog nodes store the trained deep learning model (deep learning inference) to make predictions based on the local dataset. The local dataset is the data accumulated from the coverage area of the particular fog node. Therefore, there is no need to communicate with cloud server for each prediction. Further, the small-scale processing such as aggregating the number of active cases, availability of medical resources and so forth are computed in the fog nodes itself. Thus, the overall execution time is reduced.

4.2 | Use-case: Health status monitoring and assistance at home

In this crucial time, it may not be possible to clinically test everyone. However, the preliminary testing can be done at home. The basic health parameter values such as body temperature, blood pressure (systolic and diastolic), pulse rate,

[###]Avoid 3C's: closed, crowded, close contact: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public

SPO2 (oxygen saturation) level can be collected using BAN and sent to the smart phone of the user. In case of COVID-19, the common symptoms are fever, abnormality in breathing, cough and so forth. Here, fever and breathing abnormality can be predicted from the collected body temperature, pulse rate and SPO2 level. The collected health parameter values are sent to the smart phone. The smart phones already have GPS (Global Positioning System) tracking module, the environmental temperature, and humidity detecting module. It can check that whether the collected data values fall in the normal range with respect to the user's health profile and contextual information (location, motion and acceleration, environmental temperature, and humidity). If the collected value falls in the normal range, the smart phone will show that the health status is normal in the preliminary checkup. Even then, if it is observed from the historical record of the user, that the user came in close contact with an infected or suspected COVID-19 person, then STOPPAGE alerts the user, and recommends the user to self-quarantine. Let the normal range of a health parameter hp is hp_{up_u} to hp_{low_u}, that is, the upper value is hp_{up_u} and lower value is hp_{low_u} for a user u with respect his/her health profile and contextual information, where $hp \in H$ and H is the set of health parameters. Let hp_{cl_u} is the collected value of the health parameter hp for the user u. If $hp_{up_u} \geq hp_{cl_u} \geq hp_{low_u}$, then the collected health parameter value resides in normal range. This checking is done for all health parameters considered in the experiment. If hp_{cl_u} falls in normal range for $hp\ \forall H$, then the status is predicted as normal[‖‖‖]. Otherwise, the predicted health status is abnormal and the smart phone will show an alert message, and send the collected data with user's health profile and contextual information to the cloud. The health care center of the region where the user belongs will access the data and contact the user for further checkup. In this way a preliminary health monitoring service can be provided to the user.

4.2.1 | MyHealth: An Android app for health monitoring

We have designed an Android application (app) called, MyHealth which can be used for personal health status checking. The collected health data, profile and contextual information are processed, and if the user's health status seems to be abnormal (e.g., body temperature is high, pulse rate is abnormal, SPO2 level is low), then the user is advised to self-quarantine, and medical person will contact him/her soon for further medical tests (refer Figure 6). It indicates faster health care provisioning than the traditional system, where the user has to contact a medical staff by himself/herself for further medical tests all the time.

4.2.2 | Delay in health status prediction

The health profile is already present inside the smart phone. As the contextual information collection takes place simultaneously while the smart phone receives health parameter data from BAN, the total delay in collecting and accumulating health data with profile and contextual information is given as:

$$Delay_{ca} = Delay_{mob} + max(Delay_h, Delay_c). \tag{19}$$

If there are m uplink communications and n downlink communications take place, then the total communication delay is given as,

$$Delay_{com} = \sum_{i=1}^{m}(1 + f_{ui}) \cdot (D_i/R_{ui}) + \sum_{j=1}^{n}(1 + f_{dj}) \cdot (D_j/R_{dj}). \tag{20}$$

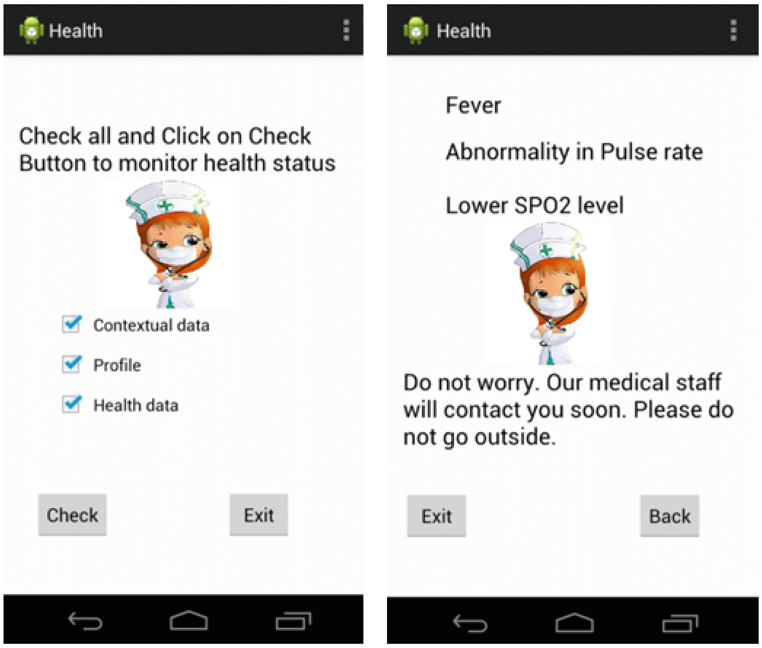
The delay in processing the data is given as,

$$Delay_{pro} = (D_{mob}/S_{mob}) + (D_f/S_f) + (D_c/S_c). \tag{21}$$

The total delay in predicting health status is given as,

$$Delay_{tot} = Delay_{ca} + Delay_{com} + Delay_{pro}. \tag{22}$$

[‖‖‖]The normal range of health parameter values are determined by medical practitioner.



(A) (B)

FIGURE 6 MyHealth: Android app for health monitoring. (A) Home page of MyHealth app, (B) Recommendation page of MyHealth app

4.2.3 | Power consumption of smartphone during health status prediction/monitoring

In the module, smartphones are used as an edge device, which collects, accumulates and sends health and context related information. In this context, it is important to measure and optimize the power consumption of the mobile device.

The power consumption of the smart phone during data collection and accumulation period is given as,

$$Po_{ca} = (Po_a \cdot Delay_{mob}) + (Po_r \cdot \max(Delay_h, Delay_c)). \tag{23}$$

If for the smart phone the number of uplink communication is k and downlink communication is q, the power consumption of the smart phone during communication period is given as,

$$\begin{aligned} Po_{com} = &\left(Po_t \cdot \left(\sum_{i=1}^{k} (1 + f_{ui}) \cdot (D_i / R_{ui}) \right) \right) \\ &+ \left(Po_r \cdot \left(\sum_{j=1}^{q} (1 + f_{dj}) \cdot (D_j / R_{dj}) \right) \right) + \left(Po_i \cdot \left(\sum_{i=1}^{(m-k)} (1 + f_{ui}) \cdot (D_i / R_{ui}) \right) \right) \\ &+ \left(Po_i \cdot \left(\sum_{j=1}^{(n-q)} (1 + f_{dj}) \cdot (D_j / R_{dj}) \right) \right). \end{aligned} \tag{24}$$

The power consumption of the smart phone during data processing period is given as,

$$Po_{pro} = Po_a \cdot (D_{mob} / S_{mob}) + Po_i \cdot (D_f / S_f) + Po_i \cdot (D_c / S_c). \tag{25}$$

In the conventional strategies followed in most of the developing countries, the user sees the parameter values, answers to the questions regarding health monitoring over voice call/message. Then based on his/her answers, the health status is predicted. If it seems to be abnormal, then the user is asked to visit a health center for further checkup. It is quite obvious this whole procedure takes much time. If the mobile health monitoring takes place, the latency is quite less than the conventional health monitoring system. Moreover, as the data collection and processing entirely takes place in the device, the accuracy of the prediction and reliability are also higher. On the other hand, the user can do this residing at home, without visiting to health center for preliminary check up until an emergency occurs. This in turn helps to avoid unnecessary gathering and helps in social distancing. Thus, we strongly recommend to use mobile health monitoring in such situations to avoid unnecessary delay and gathering.

5 | PERFORMANCE EVALUATION

In this section, we present the implementation details and experimental evaluations of the proposed framework.

5.1 | Implementation details of STOPPAGE and dataset

We have used the compute engine and app engine of Google Cloud Platform (GCP) to carry out the spatio-temporal data analysis. In the test-bed, we have used Raspberry Pi 3 as the fog device. We have also designed an Android application, using Android Studio 4.1 with *Firebase database* support, which collects the location, acceleration, proximity, temperature and light sensor data from the smartphone's in-built sensors using the *Android sensor framework*. The application is also capable of communicating with the wearable devices such as smart-watch (Fitbit). In the Raspberry Pi 3, we have installed the *Eddystone Bluetooth Beacon*, for sending data periodically. The Android application can be used to monitor the current pandemic status of the country, and for assisting the users. It retrieves the data from the nearest fog device regarding the predicted hotspot areas and sends the notification to the users within the predicted hotspots. The services of SDI is provisioned by a web-dashboard. In our SDI, the metadata information is retrieved in *.xml* or *.json* format. The architecture is implemented using Django, which is a high-level Python web framework. Gunicorn (Python Web Server Gateway Interface HTTP server) is used to integrate with backend database and web-application is provisioned.

The dataset contains 6 month mobility report of India which contains more than 280,000 data points. The attributes of the data are location, time, number of active cases, number of deceased cases, transportation data: outflow and inflow (air-travel, public, and private vehicles) from different locations (inside and outside India), weather variables at different time-interval, population density (block wise[****]), demographics (age, household types, average income), and distribution of medical resources (public and private hospitals). Apart from that, we have considered historical mobility patterns of 145 users (timestamped sequence of latitude, longitude) for 6 months. We have selected "West Bengal" as our study region (22.9868N, 87.8550E) (area: $\approx$ 88,752 km^2). The total size of the data is 5.12 GB.

For training the deep learning module, we have used popular Adam algorithm to update the network weights iteratively and optimize the parameters. The run-time (training time) of the model is 86 min for all five regions for 100 rounds (epochs). We also observe the performance (accuracy@5 for hotspot prediction) variation of the model using different cell size and batch size, and get optimal results in cell size 64 and batch size 10. For evaluating the spatio-temporal analysis, we have implemented the methods in GCP VM and QGIS framework. The overall mobility change in India is shown in Figure 7. It also shows the phase-wise lockdown and unlockdown measures taken by the government.

5.2 | Experimental evaluation on hotspot prediction

Based on the overall situation of the disease spread, we have evaluated the prediction in two phases: (i) initial phase of the disease spread (upto unlock 1.0) and (ii) next phase (till mid Dec). The accuracy of predicting the hotspots is shown in Figure 8. For this prediction, we have considered the study region as West Bengal, a state of India. We have executed the proposed method on five selected regions of the state where we have the access of individual mobility history of the residents. The precision and recall value computation captures the efficacy of *STOPPAGE* to predict the boundary of the

[****]Based on the administrative structure of India, a block or a community development block is a district subdivision, consisting of a cluster of villages or cities.

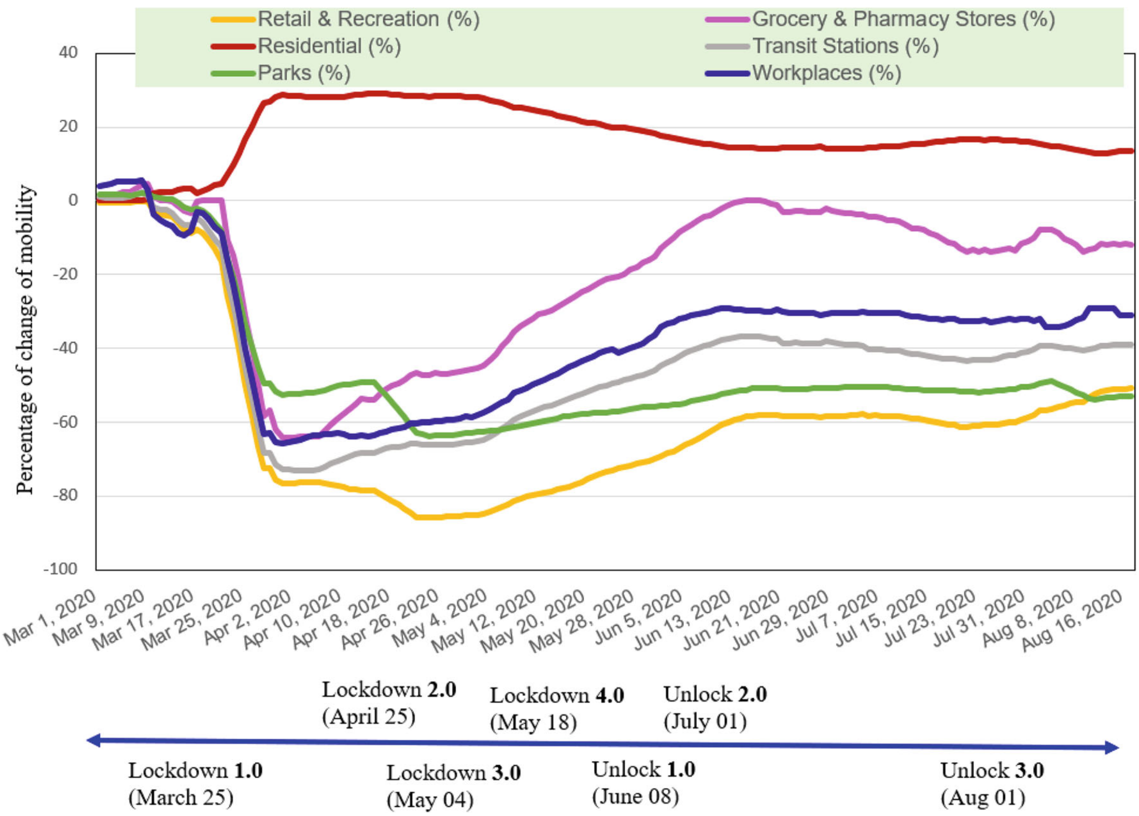


FIGURE 7 Mobility pattern change across India

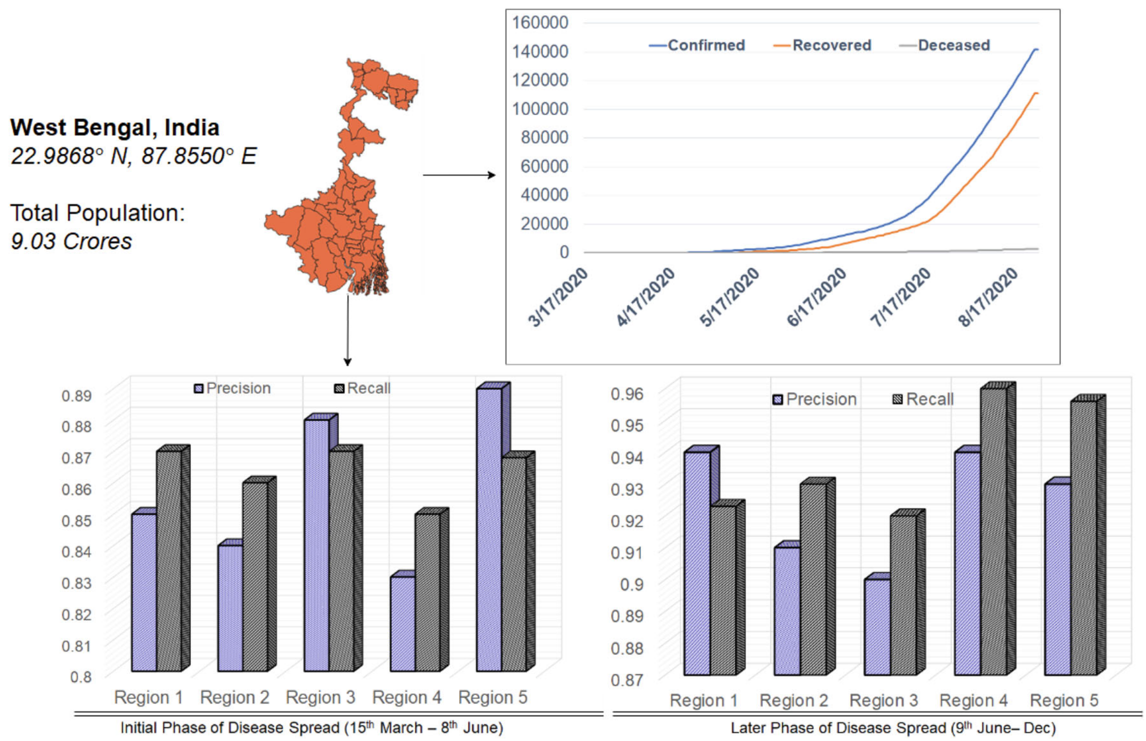


FIGURE 8 Experimental result *(five regions have been selected and the boundary of all hotspots within the regions have been predicted by* ***STOPPAGE****)*

TABLE 2 Comparison of accuracy measure and ablation study of the deep learning module for identifying COVID-19 hotspots

Model	Precision	Recall	F1-score	Accuracy
KNN	54.03%	54.8%	54.415%	56.05%
DT	50.08%	51.87%	50.972%	54.9%
SVM	60.04%	58.56%	59.3%	62.48%
CNN	75.12%	72.30%	73.71%	77.15%
ST-RNN	77.05%	72.68%	74.865%	78.72%
Heiler et al.[51]	82.18%	81.89%	82.035%	84.80%
Ghahramani et al.[52]	78.45%	85.92%	82.185%	82.08%
Parvin et al.[53]	83.18%	81.56%	82.37%	82.05%
Ablation study of the deep learning module of STOPPAGE				
w/o PKG integration	82.8%	83.74%	83.27%	86.18%
w/o attention layer	84.4%	85.78%	85.09%	88.09%
w/o bi-LSTM	85.20%	87.51%	86.355%	93.01%
w/o two-phase analysis	82.04%	86.18%	84.11%	90.45%
STOPPAGE (full)	**88.15**%	**90.25**%	**89.20%**	**93.12**%

hotspot region. The ground-truth data has been collected from the state government dashboard[††††] in daily basis. It is observed that the precision and recall in both phases of the disease spread is quite high. *STOPPAGE* can identify the hotspot areas a priori with more than 85% precision and recall values. Table 2 presents the accuracy measure of *STOPPAGE* along with the baseline methods namely, *KNN*, *Decision Tree (DT)*, *SVM*, *CNN*,[55] and *ST-RNN*.[56] The parameter for KNN is selected as 5. We have chosen *radial basis function (RF)* as the activation function in NN. A linear kernel is selected for SVM. The results for different runs are captured and average precision, recall, F1-score, and accuracy measure is reported. Apart from these, we have evaluated STOPPAGE with existing works.[51-53] Parvin et al.[53] analyses spatial autocorrelation using COVID-19 dataset of India, while socio-economic distribution is considered as a determining factor in Reference 52. The objective of the work[53] and proposed geostatistical technique are closely related to STOPPAGE. The authors measure spatial autocorrelation of different factors (population density, foreign tourist arrival, reported cases etc.) using Moran's I statistics. Then, spatial autocorrelation value is used for hotspot analysis and detection of vulnerable zones. However, the authors do not consider individual movement patterns which is one of the important factors in hotspot detection. Further, direct use of spatial autocorrelation value significantly reduces the efficacy of the system to predict hotspot. STOPPAGE uses spatial autocorrelation value to analyze disease spread pattern in neighboring region. Then, we utilize deep learning network to predict hotspots considering several factors, such as individual movement patterns, population density, outflow and inflows and so forth. Heiler et al.[51] studies the country-wide mobility pattern using mobile phone data to understand the disease spread patterns. The authors demonstrate the effect of activity-space, POI visits, and graph based community analysis using mobile phone data in Austria. The work mainly quantifies mobility data using OD-matrix (origin destination) and clustering. However, it is observed that simple counting based method and graph based analysis of movement fail to identify hotspots effectively. Socio-economic determinants are analyzed in Dublin using AI-based spatial clustering technique.[52] In our experimentation, we have considered their proposed association analysis technique between geodemographic clustering and number of confirmed cases to predict the hotspots. All of these existing works produce accuracy in the range of 82.05% to 84.80% in the same experimental set-up and dataset. It is observed that none of the existing works can outperform STOPPAGE in terms of precision (88.15%), recall (90.25%), F1-score (89.20%), and accuracy (93.12%) measure.

We have performed the ablation study of the deep learning architecture of *STOPPAGE* to demonstrate the significance of each of the modules. Table 2 shows the comparison values of the baselines and the ablation result. It is observed that *STOPPAGE* outperforms all other baselines in a significant margin. The ablation study proves that the architecture is

[††††] https://wb.gov.in/containment-zones-in-west-bengal.aspx

TABLE 3 Comparison of execution time of **STOPPAGE** for retrieving PKG information

	Execution time (in min)		
No. of entities	**Huang et al.[48]**	**Trivedi et al.[54]**	**STOPPAGE**
1×10^3	2.2	2.5	2.6
5×10^3	8.9	7.4	3.23
10×10^3	26.01	20.4	5.2
20×10^3	35.18	28.7	9.12
40×10^3	40.54	34.2	16.5
50×10^3	49.2	40.4	21.32

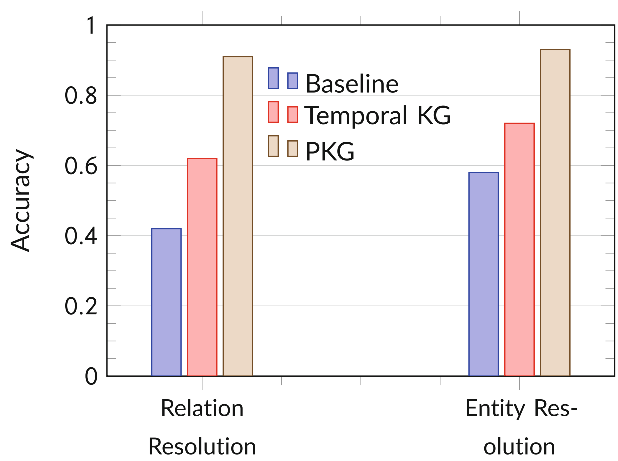


FIGURE 9 Accuracy of PKG (STOPPAGE)

suitable and better compared to all other possible set-ups. This shows the novelty and the significance of the deep learning module of *STOPPAGE* in COVID-19 context.

5.3 | Experimental evaluation of PKG and scalability study

One of the major contributions of this article is constructing the PKG and summarizing the information in the knowledge graph structure. We have implemented PKG with real-life dataset and compared the execution time to extract relations or facts with different number of entities ranging from 1000 to 50,000. The performance of PKG is compared with baseline knowledge graph[48] and temporal knowledge graph.[54] The execution time to extract the facts or relations is shown in Table 3. It has been observed that our proposed framework, *STOPPAGE* has lower execution time than the other existing approaches, which indicates *STOPPAGE* provides better time-efficiency. Figure 9 shows the accuracy measures for two tasks namely, *relation resolution* and *entity resolution*. The *relation resolution* finds out facts or relations from the knowledge graph, and *entity resolution* refines the entity information effectively. In the temporal knowledge graph,[54] the authors present a novel deep learning based framework to combine the evolving entities and their dynamically changing relationships over time, and represented that as *knowledge evolution*. However, in the context of COVID-19, this approach does not work well. It is observed that the baseline and temporal KG have achieved accuracy within 0.42–0.72 range, while *STOPPAGE* has an accuracy of 0.91–0.93 range. Therefore, it is observed empirically that *STOPPAGE* has outperformed other existing approaches in a larger margin in terms of accuracy and execution time.

We have evaluated STOPPAGE with varying user-queries and presented average execution time to demonstrate the scalability of the end-to-end system. Here, we have considered five services, namely, *hotspot prediction*, retrieving information regarding *availability of medical facilities*, *vaccination*, *active cases at a particular location during a time-interval* and *home-health assistance (preliminary health check-up, finding abnormal condition and reporting to nearest health-center)*.

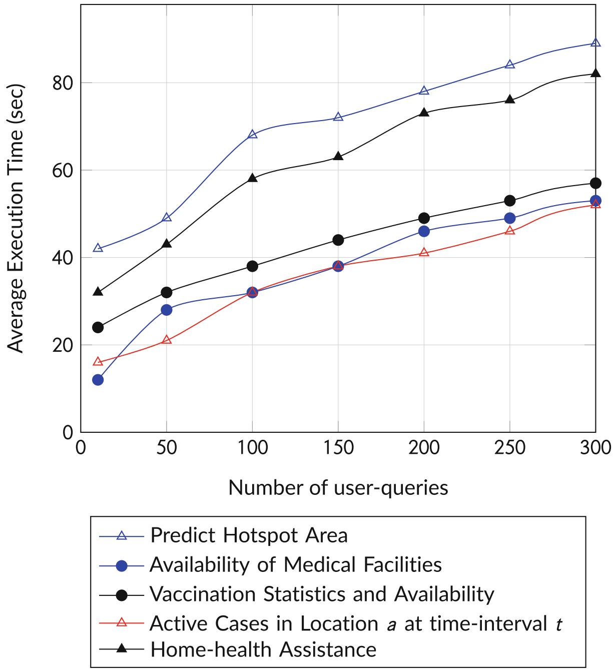


FIGURE 10 Comparison of execution time for different services

Figure 10 depicts the average execution time for above-mentioned services with increasing number of user-queries (10–300). It can be observed that even with large number of concurrent user-queries, STOPPAGE can provide service within reasonable time. The maximum execution time is required for hotspot detection in the range of 42–98 s. Obvious reason is that hotspot detection service needs to integrate several datapoints to predict the hotspot zones. In home-health assistance, we assume that user-requests are being generated from the coverage of a particular fog node. This assumption is made to provide a fair comparison, and to demonstrate the scalability of the fog node. The maximum time taken to provision home-health assistance service is 84 s for 300 concurrent user-requests. The execution time taken by other services are in the range of 12–62 s.

5.4 | Delay and power consumption of the user device in health status monitoring

The delay in predicting health status in case of proposed system and in case of cloud only system are presented in Figure 11. This is observed that the use of the proposed health care system reduces the delay ~(21–60)% than the cloud only health

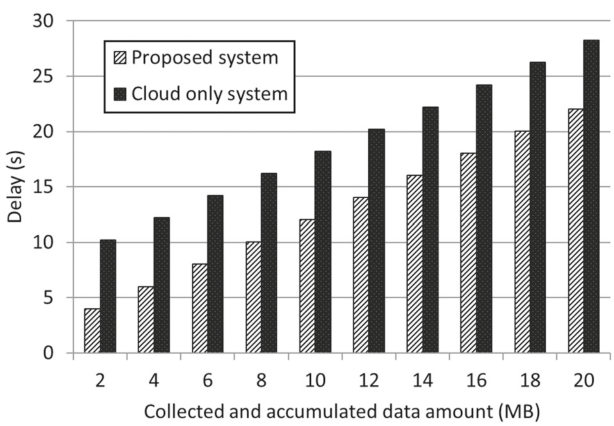


FIGURE 11 Delay in health status prediction

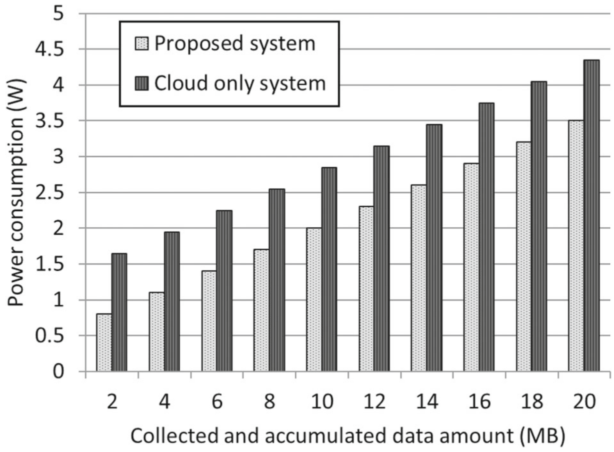


FIGURE 12 Power consumption of the smart phone during health status prediction period

care system. In cloud only system the communication delay is higher, where as in proposed system the intermediate nodes participate in data processing, which in turn reduces the communication delay. Subsequently, the total delay is reduced in the proposed system. The power consumption of the smart phone in the proposed system and cloud only system are presented in Figure 12. As in the proposed system the intermediate nodes participate in data processing the delay is reduced, and consequently the power consumption of the smart phone is also reduced. This is observed that the proposed system reduces the power consumption of the smart phone by ~(19–50)% than the cloud only system. Thus, we observe that the proposed framework predicts health status at lower latency and power consumption of the user device.

6 | CONCLUSIONS AND FUTURE WORK

In this article, we have presented an efficient pandemic management and monitoring framework, namely, *STOPPAGE* leveraging the Internet of Spatio-Health Things—a Cloud/Fog/Edge/IoT based efficient SDI to overcome the various present challenges due to the outbreak of COVID-19. Specifically, we have presented a novel spatio-temporal data analytics method using deep learning to augment varied contextual information to identify probable hotspots for taking preventive measures. A novel pandemic knowledge graph (PKG) is proposed to find out the correlations of disease outbreak and semantic spatio-temporal information. *STOPPAGE* presents varied types of movement semantics (cascading and co-occurrence patterns), and correlation with the COVID-19 spread which are potent features to predict next hotspot zones. The proposed cloud/fog/edge integrated architecture outperforms existing approaches in terms of delay, power consumption, and accuracy. It is shown that SDI is one of the major components to combat the pandemic situation.

As part of the future work, we will explore the possibility of incorporating textual data (say, Tweets) to enhance the accuracy of *STOPPAGE*. *STOPPAGE* facilitates a spatio-temporal data-driven analytics engine running over SDI backbone which is capable to extract meaningful insights from available data utilizing technologies like IoT, mobility analytics in spatial context. However, in future, we will extend *STOPPAGE* by incorporating domain knowledge from medical and epidemiology experts to make the system more robust and reliable to combat any type of pandemic situation efficiently.

AUTHOR CONTRIBUTIONS

Shreya Ghosh: Conceptualization, Methodology, Investigation, Data curation, Software, Writing - original draft, Writing - review & editing. **Anwesha Mukherjee**: Formal Analysis, Investigation, Writing - original draft, Writing - review & editing. **Soumya K Ghosh**: Conceptualization, Resources, Writing - review & editing, Supervision, Validation. **Rajkumar Buyya**: Conceptualization, Investigation, Writing - review & editing, Supervision, Validation.

FUNDING INFORMATION

The work is partially supported by the research grant from Department of Science and Technology, Government of India, under National Geospatial Chair Professorship scheme.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Shreya Ghosh https://orcid.org/0000-0002-6970-8889
Anwesha Mukherjee https://orcid.org/0000-0001-9110-8591
Soumya K. Ghosh https://orcid.org/0000-0001-8359-581X
Rajkumar Buyya https://orcid.org/0000-0001-9754-6496

REFERENCES

1. Neiderud C-J. How urbanization affects the epidemiology of emerging infectious diseases. *Infect Ecol Epidemiol*. 2015;5(1):27060.
2. Ghosh S, Mukherjee A, Ghosh SK, Buyya R. Mobi-iost: mobility-aware cloud-fog-edge-IoT collaborative framework for time-critical applications. *IEEE Trans Netw Sci Eng*. 2020;7(4):2271-2285.
3. Das J, Ghosh S, Ghosh SK, Buyya R. Lyric: deadline and budget aware spatio-temporal query processing in cloud. *IEEE Trans Serv Comput*. 2021:1-14.
4. Mukherjee A, Ghosh S, Behere A, Ghosh SK, Buyya R. Internet of health things (IOHT) for personalized health care using integrated edge-fog-cloud network. *J Ambient Intell Humaniz Comput*. 2021;12(1):943-959.
5. Ghosh S, Das J, Ghosh SK, Buyya R. Clawer: context-aware cloud-fog based workflow management framework for health emergency services. Proceedings of the 2020 20th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing (CCGRID); 2020:810-817; IEEE.
6. Mahmud R, Ramamohanarao K, Buyya R. Application management in fog computing environments: a taxonomy, review and future directions. *ACM Comput Surv*. 2020;53(4):1-43.
7. Poonia A, Ghosh S, Ghosh A, Nath SB, Ghosh SK, Buyya R. Confront: cloud-fog-dew based monitoring framework for COVID-19 management. *IoT*. 2021;16:100459.
8. Ghosh S, Das J, Ghosh SK. Locator: a cloud-fog-enabled framework for facilitating efficient location based services. Proceedings of the 2020 International Conference on COMmunication Systems & Networks (COMSNETS); 2020:87-92; IEEE.
9. Ghosh S, Ghosh SK. Modeling individual's movement patterns to infer next location from sparse trajectory traces. Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC); 2018:693-698; IEEE.
10. Vinten-Johansen P, Brody H, Paneth N, et al. Cholera, chloroform, and the science of medicine: a life of John snow. Oxford University Press; 2003.
11. Eldrandaly KA, Abdel-Basset M, Shawky LA. Internet of spatial things: a new reference model with insight analysis. *IEEE Access*. 2019;7:19653-19669.
12. Verschuur J, Koks EE, Hall JW. Global economic impacts of COVID-19 lockdown measures stand out in high-frequency shipping data. *PLoS One*. 2021;16(4):e0248818.
13. Mukherjee A, De D. Femtolet: a novel fifth generation network device for green mobile cloud computing. *Simul Model Pract Theory*. 2016;62:68-87.
14. Wang Q, Li M, Wang X, et al. COVID-19 literature knowledge graph construction and drug repurposing report generation. Proceedings of the NAACL-HLT 2021, Association for Computational Linguistics; 2021:66-77.
15. Esteva A, Kale A, Paulus R, et al. COVID-19 information retrieval with deep-learning based semantic search, question answering, and abstractive summarization. *NPJ Digit Med*. 2021;4(1):1-9.
16. Domingo-Fernández D, Baksi S, Schultz B, et al. COVID-19 knowledge graph: a computable, multi-modal, cause-and-effect knowledge model of COVID-19 pathophysiology. *Bioinformatics*. 2021;37(9):1332-1334.
17. Xu J, Kim S, Song M, et al. Building a pubmed knowledge graph. *Sci Data*. 2020;7(1):1-15.
18. Reese JT, Unni D, Callahan TJ, et al. Kg-COVID-19: a framework to produce customized knowledge graphs for COVID-19 response. *Patterns*. 2021;2(1):100155.
19. Kraemer MU, Yang C-H, Gutierrez B, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*. 2020;368(6490):493-497.
20. Pepe E, Bajardi P, Gauvin L, et al. COVID-19 outbreak response, a dataset to assess mobility changes in Italy following national lockdown. *Sci Data*. 2020;7(1):1-7.
21. He H, Li R, Wang R, Bao J, Zheng Y, Li T. Efficient suspected infected crowds detection based on spatio-temporal trajectories, arXiv preprint arXiv:2004.06653.
22. Badr HS, Du H, Marshall M, Dong E, Squire MM, Gardner LM. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect Dis*. 2020;20:1247-1254.

23. Rahman MA, Hossain MS, Alrajeh NA, Guizani N. B5g and explainable deep learning assisted healthcare vertical at the edge: COVID-19 perspective. *IEEE Netw*. 2020;34(4):98-105.
24. Kapoor A, Ben X, Liu L, et al. Examining COVID-19 forecasting using spatio-temporal graph neural networks. arXiv preprint arXiv:2007.03113.
25. Luz E, Silva P, Silva R, et al. Towards an effective and efficient deep learning model for COVID-19 patterns detection in x-ray images. *Res Biomed Eng*. 2022;38:149-162.
26. Tuli S, Tuli S, Tuli R, Gill SS. Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing. *IoT*. 2020;11:100222.
27. Whaiduzzaman M, Hossain MR, Shovon AR, et al. A privacy-preserving mobile and fog computing framework to trace and prevent COVID-19 community transmission. *IEEE J Biomed Health Inform*. 2020;24(12):3564-3575.
28. Kumar A, Sharma K, Singh H, Naugriya SG, Gill SS, Buyya R. A drone-based networked system and methods for combating coronavirus disease (COVID-19) pandemic. *Future Generation Computer Systems (FGCS)*. Vol 115. Elsevier Press; 2021:1-19.
29. da Costa CA, Pasluosta CF, Eskofier B, da Silva DB, da Rosa Righi R. Internet of health things: toward intelligent vital signs monitoring in hospital wards. *Artif Intell Med*. 2018;89:61-69.
30. Santos MA, Munoz R, Olivares R, Rebouças Filho PP, Del Ser J, de Albuquerque VHC. Online heart monitoring systems on the internet of health things environments: a survey, a reference model and an outlook. *Inf Fusion*. 2020;53:222-239.
31. Aman AHM, Hassan WH, Sameen S, Attarbashi ZS, Alizadeh M, Latiff LA. Iomt amid COVID-19 pandemic: application, architecture, technology, and security. *J Netw Comput Appl*. 2021;174:102886.
32. Sood SK, Kaur A, Sood V. Energy efficient iot-fog based architectural paradigm for prevention of dengue fever infection. *J Parallel Distrib Comput*. 2021;150:46-59.
33. Das J, Ghosh S, Mukherjee A, Ghosh SK, Buyya R. Rescue: enabling green healthcare services using integrated IoT-edge-fog-cloud computing environments. *Softw Pract Exp*. 2022;52:1615-1642.
34. Vedam N, Ghose D. The impact of mobility and interventions on the spread of diseases. *IEEE Trans Comput Soc Syst*. 2022:1-21.
35. Ienca M, Vayena E. On the responsible use of digital data to tackle the COVID-19 pandemic. *Nat Med*. 2020;26(4):463-464.
36. Wang P, Lin C, Obaidat MS, Yu Z, Wei Z, Zhang Q. Contact tracing incentive for COVID-19 and other pandemic diseases from a crowdsourcing perspective. *IEEE Internet Things J*. 2021;8(21):15863-15874.
37. Song Q, Zheng Y-J, Yang J, Huang Y-J, Sheng W-G, Chen S-Y. Predicting demands of COVID-19 prevention and control materials via co-evolutionary transfer learning. *IEEE Trans Cybern*. 2022:1-14.
38. Yang Z, Liang B, Ji W. An intelligent End–Edge–Cloud architecture for visual IoT-assisted healthcare systems. *IEEE Internet Things J*. 2021;8(23):16779-16786.
39. Smith M, Ponce-de Leon M, Valencia A. Evaluating the policy of closing bars and restaurants in cataluña and its effects on mobility and covid19 incidence. *Sci Rep*. 2022;12(1):1-19.
40. Singh A, Kaur A, Dhillon A, Ahuja S, Vohra H. Software system to predict the infection in COVID-19 patients using deep learning and web of things. *Softw Pract Exp*. 2022;52(4):841-867.
41. Abid A, Cheikhrouhou S, Kallel S, Jmaiel M. Novidchain: blockchain-based privacy-preserving platform for COVID-19 test/vaccine certificates. *Softw Pract Exp*. 2022;52(4):868-886.
42. Ghosh S, Mukherjee A. Strove: spatial data infrastructure enabled cloud–fog–edge computing framework for combating COVID-19 pandemic. *Innov Syst Softw Eng*. 2022;1-17.
43. Ghosh S, Mukherjee A. Cloud–fog–edge computing framework for combating COVID-19 pandemic. Proceedings of International Conference on Advanced Computing Applications; 2022:247-255; Springer, New York.
44. Mohan S, Abugabah A, Kumar Singh S, Bashir AK, Sanzogni L. An approach to forecast impact of COVID-19 using supervised machine learning model. *Softw Pract Exp*. 2022;52(4):824-840.
45. Tuli S, Mahmud R, Tuli S, Buyya R. Fogbus: a blockchain-based lightweight framework for edge and fog computing. *J Syst Softw*. 2019;154:22-36.
46. Kallel A, Rekik M, Khemakhem M. Iot-fog-cloud based architecture for smart systems: prototypes of autism and COVID-19 monitoring systems. *Softw Pract Exp*. 2021;51(1):91-116.
47. Goodchild MF. Citizens as voluntary sensors: spatial data infrastructure in the world of web 20. *Int J Spatial Data Infrastruct Res*. 2007;2(2):24-32.
48. Huang X, Zhang J, Li D, Li P. Knowledge graph embedding based question answering. Proceedings of the 12th ACM International Conference on Web Search and Data Mining; 2019:105-113; ACM.
49. Ghosh S, Ghosh SK, Buyya R. Mario: a spatio-temporal data mining framework on google cloud to explore mobility dynamics from taxi trajectories. *J Netw Comput Appl*. 2020;164:102692.
50. Moran PA. Notes on continuous stochastic phenomena. *Biometrika*. 1950;37(1/2):17-23.
51. Heiler G, Reisch T, Hurt J, et al. Country-wide mobility changes observed using mobile phone data during COVID-19 pandemic. Proceedings of the 2020 IEEE International Conference on Big Data (Big Data); 2020:3123-3132; IEEE.
52. Ghahramani M, Pilla F. Leveraging artificial intelligence to analyze the COVID-19 distribution pattern based on socio-economic determinants. *Sustain Cities Soc*. 2021;69:102848.
53. Parvin F, Ali SA, Hashmi S, Ahmad A. Spatial prediction and mapping of the COVID-19 hotspot in India using geostatistical technique. *Spat Inf Res*. 2021;29(4):479-494.

54. Trivedi R, Dai H, Wang Y, Song L. Know-evolve: deep temporal reasoning for dynamic knowledge graphs. Proceedings of the 34th International Conference on Machine Learning; Vol. 70, 2017:3462-3471.
55. Karatzoglou A, Schnell N, Beigl M. A convolutional neural network approach for modeling semantic trajectories and predicting future locations. Proceedings of the International Conference on Artificial Neural Networks; 2018:61-72; Springer, New York.
56. Liu Q, Wu S, Wang L, Tan T. Predicting the next location: a recurrent model with spatial and temporal contexts. Proceedings of the 30th AAAI Conference on Artificial Intelligence; 2016.

How to cite this article: Ghosh S, Mukherjee A, Ghosh SK, Buyya R. STOPPAGE: Spatio-temporal data driven cloud-fog-edge computing framework for pandemic monitoring and management. *Softw Pract Exper*. 2022;52(12):2700-2726. doi: 10.1002/spe.3144