CYBER-PHYSICAL SPATIAL TEMPORAL ANALYTICS FOR DIGITAL TWIN-ENABLED SMART CONTACT TRACING

ABSTRACT

- Purpose: Physical gatherings at social events have been found as one of the main causes of COVID-19 transmission all over the world. Smartphones, exchanging messages have been used to build contact traceability and alerts through their Bluetooth functionality. However, recent confirmed cases found in venues indicated that indirect transmission of the causative virus occurred, resulting from virus contamination of common objects, virus aerosolization in a confined space, or spread from inadequate ventilation environment with no indication of human direct or close contact observed.
- Design/methodology/approach: This paper presents a novel cyber-physical architecture for spatial temporal analytics (iGather for short). Locations with time windows are modeled as digital chromosomes in cyberspace to represent human activity instances in the physical world.
- Findings: Results show that the high spatial temporal correlated but indirect tracing can be realized through the deployment of physical hardware and spatial temporal analytics including mobility and traceability analytics. iGather is tested and verified in different spatial temporal correlated cases. From a management perspective of mobilizing social capacity, the venue not only plays a promotion role in boosting the utilization rates but also a supervision-assisted role for keeping the venue in a safe and healthy situation.
- Originality: This study contributes to the existing literature by cyber-physical spatial temporal analytics to trace COVID-19 indirect contacts through digital chromosome, a representation of digital twin technology. Also, the authors have proposed a venue-oriented management perspective to resolve privacy-preserving and unitization rate concerns.

- Social implications: This research is of particular significance when physical distancing measures are being relaxed with situations gradually become contained. iGather is able to help the general public to ease open questions: Is a venue safe enough? Is there anyone at a gathering at risk? What should one do when someone gets infected without raising privacy issues?

Keywords: Cyber-Physical Systems, Analytics, Digital Twins, Contact Tracing, Spatial Temporal Correlation.

1 INTRODUCTION

COVID-19 outbreak has escalated from epidemic to pandemic. Successful mass vaccination takes time. For individuals of a population to develop natural immunity through cross-infection has been proved morally and scientifically unacceptable as the fatality rate is high (Richards, 2020). Therefore, public health practices and operations management become crucial in fighting against this pandemic. Three levels of operational experiences have demonstrated their effectiveness in limiting the spread within varying scopes and scales, including social distancing, quarantine, and isolation (Wilder-Smith and Freedman, 2020).

Isolation and quarantine are used to protect the public by preventing exposure to infected persons or to persons who may be susceptible to infection. Isolation is used to separate infected individuals from the population. The movement of patients restricted within an environment such as a hospital to help contain the spread of the disease. In contrast, quarantine is carried out to restrict the movement of suspected individuals who may have been exposed to a communicable disease and observe their health conditions. These people may not be aware of their exposure to COVID-19, or they may carry the disease but do not show symptoms. Most recently, quarantine has been extensively adopted in the COVID-19 outbreak (Nussbaumer-Streit et al., 2020). This outbreak has seen entire cities in China effectively placed under mass quarantine, while many thousands of foreign nationals returning home from China have been asked to self-isolate at home or in state-run facilities. There are precedents for such measures (Taghrir et al., 2020). Citywide quarantines were also imposed in areas of China and Canada during the 2003 outbreak of severe acute respiratory syndrome (SARS) (Hawryluck et al., 2004, Zhou et al., 2004), whereas entire villages in many West African countries were quarantined during the 2014 Ebola outbreak (Drazen et al., 2014). The first-wave of COVID-19 shock in China has abated through mass quarantine. However, the contact tracing of incidental and sporadic COVID-19 cases had pose new challenges for quarantine.

Recent research and developments using the Bluetooth communication module embedded in the smartphone to realize contact tracing has gained research and industrial attention rapidly (Hatke et al., 2020). Even Apple and Google decided to work together to release application programming interfaces that enable interoperability between Android and iOS devices (Panzzarino, 2020). Governments and public authorities of different countries published contact tracing applications to enable timely identification of proximity contactors upon confirmed cases (Watts, 2020, Stevens and Haines, 2020). The basic idea behind the Bluetooth enabled contact tracking applications is that the smartphone carriers are broadcasting their own encrypted Bluetooth radio packets and receiving from others in a non-connectable manner under certain privacy-preserving methods. Bluetooth contact records either in a local smartphone or a remote server are used to extract records related to activities of any smartphone user is diagnosed as a confirmed case. Close contacts are therefore immediately identified. However, epidemic prevention can only be effective when a significant proportion (more than 80%) of the population uses the smartphone application. Moreover, recent medical news has posed new challenges to current contact tracing applications, data of several confirmed cases found in venues indicated that indirect transmission of the causative virus occurred, perhaps resulting from virus contamination of common objects, virus aerosolization in a confined space, or spread from asymptomatic infected persons. In these incidents, no indication of human direct or close contact has been observed.

Despite the recent achievements and research that have pointed out the potential solutions, the medical findings and utility analysis still poses several challenges. The first challenge is how to track the potentially associated person in a certain spatial temporal correlation. People may or may not simultaneously appear in the same time window, since the virus exists in a confined space. If other people appear in the same place afterward, there is a risk of infection, and immediate notification is required. The second challenge is a certain venue is usually related to mass infection cases. How to establish spatial temporal model related to the venue is open for investigation. The third challenge is that even contact tracing applications claim that the privacy-preserving method prevents disclosure of personal information. However, malicious hackers still can capture encrypted data packages from smartphone carriers during Bluetooth transmission. Every disclosure of personal information comes with some latent risk that it will be used in the future for purposes not disclosed at collection. No matter how complicated the method of data encryption, there is a need to fundamentally solve the contact tracing issue without Bluetooth broadcasting from smartphone carriers.

This research establishes iGather: a cyber-physical smart contact tracing architecture as a technical pillar, privacy-preserving and unitization rate concerns are innovatively resolved from a venue-oriented management perspective. A spatial temporal analytics is proposed for contact tracing issues, especially for indirect contact. Digital chromosome, a representation of digital twin technology, is developed to synchronize real and virtual spatial temporal instances. Several scenarios are illustrated to demonstrate the feasibility and efficacy of iGather.

The rest of this paper is organized as follows. Section 2 reviews the recent developments, including opportunities and challenges on fighting COVID-19 in practices and academic. iGather, the system architecture for smart contact tracing, is discussed in detail and developed in Section 3. Two types of spatial temporal analytics, especially for indirect contact tracing is discussed in Section 4. Section 5 depicts several scenarios for feasibility and efficiency verification. The limitations of this research and future research directions are concluded in Section 6.

2 LITERATURE REVIEW: OPPORTUNITIES AND CHALLENGES

The contact tracing is considered the most effective step towards containing the COVID-19 outbreak(Kretzschmar et al., 2020). The adoption of emerging technologies has shown significant influence across many countries for leveraging the entire workflow of the carefully designed and customized quarantine measure throughout the planning, monitoring, and controlling operations. A highlight of how technology is currently enabling government bodies to manage the outbreak of COVID-19 is briefly summarized. A mobile application called "Health Code" is developed in China for citizens to broadcast their current health status to several authorities. Designated citizens can obtain color codes ranging from green (low risk) to red (high risk). The application combines geo-tracking with train, plane, and public transport activity records, so those high-risk individuals will not be granted access to shopping malls, communities, offices, travel spots, etc. (CGTN, 2020). In Korea, citizens' activities are monitored by the government with the logs acquired from several sources, such as smartphone location data, CCTV, and credit-card records (Ahn, 2020). The government notifies the public when any medical examination result on the virus is found positive. Recent activities and personal information of an individual whose medical examination is found positive may be revealed by the government. Such information includes the individual's last name, gender, age, district of residence, visited places, and visited restaurants. Residents in Singapore can download a mobile application called "TraceTogether", which uses Bluetooth technology to keep a log of nearby devices. Both phones exchange anonymized IDs. If an individual is found infected, the individual can upload the relevant data to a database maintained by the Ministry of Health, which notifies the owners of all the devices pinged by the infected individual's smartphone (Singapore, 2020). The government of the US is in active talks with Facebook, Google, and a wide array of tech companies and health experts about how they can use location data from smartphones to track whether citizens are keeping one another at safe distances to stem the outbreak. The Australian government launched a mobile application, named "COVIDSafe," using Bluetooth to recognize another user through an encrypted identifier. Over the past months, designated quarantine measures are formulated accordingly by every country, which are, to a certain extent, somewhat different from each other depending on the seriousness of the COVID-19 outbreak.

Establishing digital tracking and monitoring via the development and deployment of emerging IoT and digital technologies largely equips governments with the ability to contain the outbreak of COVID-19. The technology adoption of the abovementioned countries and regions has proven the essence of technology integration as an enabler of the government in effectively developing and exercising contact tracing measures. In the light of requiring extensive information exchange in real-time to facilitate the entire quarantine workflow with the establishment of quarantine village, the integration of multiple technology-enabled containers extends the standalone advantages of single monitoring service, thereby potentially maximizing the effectiveness of total quarantine control. Xia and Lee (2020) outlined Bluetooth enabled contact tracing applications (CTA). They suggested CTA can even replace social distancing given the high adoption rate. McLachlan et al. (2020) proposed a Bayesian model to predict the likelihood of infection supplemented by traditional symptoms and contact tracking. Li and Guo (2020) surveyed CTA in global deployment and challenges. One of the challenges is if the digital ID is unique and static, it runs the risk that certain digital IDs could be hacked and paired with a mobile device, thus compromising individual privacy. Moreover, different mobile devices exhibit a variety of Bluetooth signal intensity, which may confuse signal processing for proximity estimation. Applicability and technology structure of the mobile contact tracing in fighting the pandemic is discussed and the evaluation framework for contact tracing solutions is proposed. The aforementioned measures taken by countries and regions are evaluated in the research(Dar et al., 2020). Then performed a readability analysis to evaluate the comprehensibility of the privacy policies of several CTA(Zhang et al., 2020). Cho et al. (2020) discussed the privacy considerations and trade-offs of CTA. They believed that augmenting with random tokens does not improve privacy and if contacts are malicious, malformed queries will be sent for lowering privacy level. Hernández-Orallo et al. (2020) evaluated the effectiveness of IoT technologies such as GPS, cellular networks, Wi-Fi, and Bluetooth. They determined the impact of contact tracking precision on the spread and control of the infectious disease. The authors believed smartphone-based contact tracing could be extremely useful considering the second wave of infectious diseases. It is noteworthy that the current contact tracing research focus on the direct contact tracing which means people appear at same time and have close contact history. Within ~~in~~ the Bluetooth signals exchanging distance, the contact digital record can be either stored at local smartphone or uploaded to centralized cloud server. The direct contact is depicted in the first scenario in Figure 1 which is the main reason for person-to-person transmission.

Figure 1

Recent medical and aerodynamics research reveals that the COVID-19 is normally non-long-range airborne when there is reasonably sufficient ventilation, but becomes airborne when there is insufficient ventilation which is called opportunistic airborne transmission (Li et al., 2020, Morawska and Milton, 2020). Braithwaite et al. (2020) believed that effect contact tracing depends on the timeliness of case identification. People showed up at the same time but were geographically distant from the others. Infections still occur due to the airflow in the same and inadequate ventilation system. This case is the second scenario in Figure 1, which also has a high spatial temporal correlation. The indirect contact in the third scenario states that people are not showing up simultaneously but in sequent and the geographical location is close. Report of four clinical laboratory technicians in the Department of Clinical Laboratory of Jinyintan Hospital, Wuhan, were infected with COVID-19. Those who had no contact with the confirmed cases may be the evidence for aerosol transmission (Wang and Du, 2020). The virus in microdroplets which was small enough to remain aloft in the air or attached to the surface of the object post threats to late-comers (Morawska and Milton, 2020). Recirculation bubbles (in minimizing inflow of clean air) or stratified air layers (in minimizing dilution) may also enhance the airborne transmission with the unfavorable release of infectious aerosols (Li et al., 2020). The second and third scenarios are indirect contact with high spatial temporal correlation, which also deserves enough attention.

Digital twin and cyber-physical systems as new emerging technologies, which connect the physical and virtual worlds, have attracted much attention worldwide, especially in the manufacturing industry (Tao et al., 2019, Guo et al., 2020, Zhong et al., 2017). Ding et al. (2019) adopted the cyber-physical system and digital twin technologies to build the interconnection and interoperability of a physical shop floor and corresponding cybershop floor to realize real-time monitoring, simulation, and optimization. Luo et al. (2019) proposed a comprehensive framework for integrating big data analytics into cyber-physical system, the redundant raw data can be converted into smart data. Tu et al. (2018) implemented an IoT-based cyber-physical prototype in a production logistics scenario. IoT technology is integrated into cyber-physical system to enable communication between layers.

The location and time information act as core components to conduct the spatial-temporal analytics since it links space, time, and the subject. In the indirect contact tracing area, two challenges are encountered. Firstly, how to track the high spatial temporal correlated but indirect contacts is open for investigation. Current Bluetooth based practices exchange Bluetooth messages among users to identify the detailed contact record through received signal strength indicator (RSSI) within the range of broadcasting distance. Signal exchanging among smartphones requires a very high spatial temporal correlation. For example, two individuals are in close contact within the same time window so that the Bluetooth signals can be both received. However, airborne transmission (airflow) and aerosol may pose challenges to the current solution that the high spatial temporal correlation but with divergence. For example, two people are in distant contact, or two people come to the same location but not in the same time window.

Secondly, current practices raise two issues: privacy and utilization rate. Privacy issues always bring individuals great vigilance, especially the disclosure of location information. People are generally happy to share locations of their goods and belongings, but they have serious concerns about revealing their personal location information to other parties. The aim of the contact tracing is not to track where the user went to but to screen out the indirect contacts at a specific spatial-temporal window. Research shows that contact tracing can be effective when a significant proportion (more than 80%) of the population uses the smartphone-based application. A novel "easy to deploy, plug and play" technology architecture to preserve privacy essentially and facilitate the utilization rate from a management perspective are therefore necessary.

3 iGATHER: SYSTEM ARCHITECTURE FOR SMART CONTACT TRACING

To fulfill the research gap, iGather: a cyber-physical smart contact tracing architecture is proposed. Internet of Things devices are mounted and carried in physical space to collect, broadcast, and generate information. Decisions and analytics are conducted in cyberspace based on the collected data for contact tracing purposes. The concerns of privacy and use proportion is discussed and designed behind the architecture. The conceptual framework of iGather is depicted in Figure 2.

Figure 2

There are three stakeholders in the iGather conceptual framework, namely individuals, venues, and health authority. During the initialization phase, the venues and individuals conduct registration action to the health authority, the venues then receive the venue beacon nodes to deploy onsite. The confirmed and mandatory quarantine cases receive wristband beacon nodes. Once the individuals have physical gatherings at venues, their smartphones receive the signal data emitted from the venue beacon nodes and store it locally as digital chromosome without uploading. Health authority updates the digital chromosome list from newly confirmed cases and pushes to individuals mobile application. The mobile application of individuals conducts spatial-temporal analytics locally to identify the relevancy with confirmed cases. Once been considered as close contact, the mobile application status will be displayed as an abnormal state and urge the user to seek medical service. The close contact analytics result will not automatically upload to any server for the privacy-preserving purpose. The user may take the following actions:

(1) Go for medical service to coronavirus nucleic acid detection which is highly appreciated.

(2) Still go outside and have physical gatherings with others at different venues. The venues have to conduct status checking before enter permission.

(3) Stay at home to conduct home quarantine for certain days.

The health authorities can give COVID-19 safety ratings to venues according to the historical data. The following figure then illustrates the system architecture of iGather for smart indirect contact tracing from a technology perspective.

Figure 3

3.1 Physical space deployment

There are three groups of iGather users: (1) Social individuals, most of whom are just healthy people while a very small minority maybe those under quarantine and even isolation; (2) Corporate venue where people gather for some purposes; and (3) Public health professionals and officials who define measures, execute operations and monitor the situation. Health authorities distribute two kinds of Bluetooth-compliant beacon nodes to two types of carriers, respectively. The first is the wristband beacon node, which distributes to confirmed or mandatory quarantined cases (C/MQ case). Users are required to wear the wristband beacon until they recover or the time reaches the end of mandatory quarantine. The second kind beacon nodes are distributed to public venues such as restaurants, public transport, and parks, etc. The coronavirus spreads rapidly, especially when people gathering in public spaces rather than walking on the street to their destination. The health code system has been widely used especially when visiting public places such as shops and airports. Users are required to provide authorized health code to grant access to facilities. In this case, venue managers actually play a promotion role in boosting utilization rates. It is difficult for the government or official institutions to request populace install or deploy technology application, but the various venues can contribute to encourage and verify the application installation from a management perspective. Venue beacon nodes are stick or attached to the fixed positions in venues such as the entrance of the mall, train compartment, or light poles. Both kinds of beacon nodes are broadcasting specific and customized UUID at certain time intervals and power strength. Healthy, recovered, or asymptomatic users are required to carry smartphones only for receiving Bluetooth signals from beacon nodes. Many studies infer that the message exchange between nearby devices may quickly notify and record the proximity and contact history even privacy-preserving methods have been adopted to encrypt the message. However, the messages sending from the user's smartphone always pose information security risks and make users feel that privacy has been violated. In this architecture, users' smartphones do not broadcast or send any information to nearby devices but only receives the signals from nearby beacon nodes.

3.2 Cyberspace information encapsulation

To realize contact tracing purposes, information encapsulation should be established in cyberspace based on the physical space deployment. Wristband beacon nodes attached to C/MQ cases request registration to the iGather cloud server, an encrypted user ID is then generated to the wristband beacon node. Bluetooth signal carried with the encrypted user ID is then broadcasted to nearby smartphones. For instance, user D is a mandatory quarantine case required to stay home for 14 consecutive days. The health authority gives a wristband beacon node and requests to be carried in the following quarantine days. If user D goes to public places, the signals emitted from wristband beacon nodes will be received by other smartphones.

Thereby, an immediate alert and notice will be shown on other users' smartphones for further attention.

Venue managers register the received the venue beacon nodes to iGather cloud server and the corresponding ID is generated and returned. The specific information such as venue name, address, and contact person is associated with ID for quick response if an outbreak occurs in the venue. Venue managers are required to check the user iGather status and only the correct status will be allowed to enter the venue, which serves the purpose of promoting users to download and use. The venue beacon nodes deployed at fixed positions broadcast the message to the user's smartphone. A "digital chromosome" is then generated, the digital chromosome contains the time and received broadcast packets including ID and received signal strength indicator (RSSI) from the venue beacon nodes. Users store the digital chromosomes locally in the smartphone rather than uploading them to the cloud server. If one of the users gets infected, the health authority requires to invoke the digital chromosome from the user side and transfer it to iGather cloud server. iGather cloud server pushes daily updates of suspected digital chromosome lists to application users. The spatial-temporal mobility and traceability analytics are then distributed executed in the user smartphone to filter the close contacts. Close contacts with special attention will receive alerts and instructions from public health authorities. The iGather status of close contacts is changed to suspected cases, which are refused to enter public venues by venue manager until they are tested or follow relative instructions by the public health authorities.

3.3 Privacy and use proportion concerns

During the pandemic of COVID-19, various privacy-preserving methods are discussed and evaluated since the start of contact tracing applications. WHO affirms the digital tools offer opportunities to strengthen contact tracking while raises privacy concerns about the disclosure of personal data(Organization, 2020). Complex encryption and disguise can protect privacy to some degree but malicious hackers may capture the information during transmission from user smartphone which is worrisome. In iGather architecture, the smartphone of the user is not broadcasting any Bluetooth enabled signals to nearby devices and not uploading any information to the cloud. All the data is locally stored and clear after 14 days. iGather users only receive the signals broadcast sending from public venue beacon nodes and wristband beacon nodes. In order to maintain an updated situation, users download suspected digital chromosome list from the iGather cloud server to execute local analytics.

People may not accept or believe the technology method to realize contact tracing for common good. They are reluctant to use these contact tracing solutions. The large-scale use of the Health Code system in China has brought new thinking. Public train stations, shopping malls, and even parks have set up checkpoints to view whether the incomer has "Green Code (Health)". In the proposed iGather architecture, venues play an important role in promoting the use of contact tracing applications. Only normal or healthy iGather users are permitted to enter the venue. From a management perspective, it is hard for the governments and health authorities to urge and regulate every single citizen to download or use the contact tracing, which sounds like a mandatory requirement for people to accept, but it is flexible for the venue to follow and regulate.

4 SPATIAL TEMPORAL ANALYTICS FOR DIGITAL TWIN ENABLED INDIRECT CONTACT TRACING

Recent research efforts have indicated that COVID-19 viruses are transmitted during exhalation, talking and coughing in droplets small enough to remain aloft in the air and inadequate ventilation system caused infections, which pose risks to people who do not even have direct contact history. Current contact tracing applications, including the collaboration between Google and Apple only consider the direct contact scenarios where users gather in the coincident time window and nearby location. How to trace the spatial-temporal correlated but indirect contacts is discussed in this section. The Department of Defense of USA first proposed digital twin to conduct a health check of aerospace vehicles(Glaessgen and Stargel, 2012). It adopts physical modeling, sensor updating, and running history to realize the mapping from the full life cycle of the physical asset in the physical world to virtual representation in the information space. In order to realize the indirect contact tracing in the information space, a realistic depiction of what happens in the physical world is necessary. In the light of digital twin conception, the human movement trajectory and stay time can be portrayed and mapped in the information space through the proper signal information model. The signal data contributes to simulate the attributes of geometry, rules, and constraints of human activity history in the physical world. With the help of digital twin, the indirect contact tracing is supposed to be orchestrated in the physical world and information space.

Figure 4

Figure 4 explains the two high spatial temporal correlated situations in the venue scenario. It depicts a venue where it has multiple levels and relatively wide environment such as shopping malls or train stations. The airflows of the ventilation system for air circulation are labeled on the figure. Venue beacon nodes are deployed at venues via a “stick and play manner” without the support of electricity and network, which only broadcasts Bluetooth compliant signal packages to nearby smartphones. The basic idea behind the current solutions for direct contact tracing is the Bluetooth message exchanging between two users’ smartphones. However, there are two challenging situations where the current solution cannot trace the contacts. Two illustrative cases are listed to clarify two high spatial-temporal correlated but indirect contact situations. The first case is user A and user B showed up at the same time window but geographically distant, which exceeds the smartphone Bluetooth transmission distance. If user B gets infected, the ventilation of the venue may bring the aerosol with the virus to every corner of the venue at the same time window. Thus, user A is supposed to be tracked and take immediate actions (NAT and self-isolation). The second case is that user B and user C visit the same place, which is geographically close but not in the same window. User C comes to the same place that user B had been. If user B is tested COVID-19 positive, the droplets caused by exhalation, coughing, talking remain aloft in the air, and even touching makes the virus stay on the surface of the object. User C will face a great risk of infection, considering the attachment and stay of the virus from user B. The smartphone devices contribute to collect the data that broadcasted from fixed venue beacon nodes, a digital twin representation, a digital chromosome is then established with a time window and received signals strength indicators in local storage. Thus, physical staying and passing by historical data is recorded locally for further mobility and traceability analytics, which targeted at the high spatial temporal correlated indirect contact tracing.

4.1 Mobility Analytics

The mobility analytics is concerned with (1) people gathered in the same time window and (2) they are geographically distant from others (long-range) which exceeds Bluetooth transmission distance from the user smartphone but within a coverage area of the same ventilation system. The inadequate ventilation or unpredictable airflow pose threats to people who may be in different rooms or levels in the same venue but also get infected. The airborne transmission of COVID-19 is still under research given existing ventilation infection examples. The indirect contact is therefore supposed to be tracked for a quick response.

Figure 5

The mobility analytics procedure includes two phases: collection and analytics described in Figure 5 (a) and (b), respectively. The mobility analytics conducts at each smartphone adopting edge-computing capability to offload the centralized server pressure and preserve privacy maximally. During the collection phase, venue beacon nodes are deployed at fixed positions in the venue. If the venue beacon nodes are under the same ventilation system, they are clustered as one same set. The indirect time window *T* is defined by the Department of Health, the National Health Commission, or other government authorities, which refer to the possible time window for infection occurs. Each venue beacon node then starts broadcasting Bluetooth signal packages. iGather user receives and acquires the encrypted name of the venue beacon node *UUID*, RSSI *R*, and the timestamp *t* from the signal packages. A digital twin is hereby generated. A physical instance in the digital twin is the event that a human has been in a specific situation where spatially and temporally existing, and the virtual instance in a digital twin is the received and locally stored signal packets as digital chromosomes. The digital chromosome can reflect that the anonymous but unique human has been to anonymous but unique places at a specific time. If the user is diagnosed to be positive for COVID-19, the digital chromosome is therefore transferred to iGather cloud server through permission from health authorities. iGather users regularly update the suspected digital chromosome lists *L* from iGather cloud server. Each update triggers the local analytics of potential indirect risks. The edge computing capacity of the smartphone enables the select suspected venue beacon nodes *SVBN* and timestamp *st* from suspected digital chromosomes and extract venue beacon nodes *VBN* and timestamp *t* from local stored digital chromosomes. If *SVBN* and *VBN* belong to the same venue beacon set, which means they are in the same venue and under the same ventilation system, and the timestamp of the two records both in the indirect time window *T*. In this case, iGather confirms that the user is the indirect contact of confirmed cases. Thus the alert is triggered.

4.2 Traceability Analytics

The traceability analytics is concerned with (1) people gathered not in the same time window and (2) they are geographically close (short-range). The aerosol with the virus remains aloft in the air and the surface of items is contagious. People even do not show up at the same time but still have a chance to get infected if they stay in the nearby location chronologically by exhaling and breathing aerosol and touching contaminated objects. Traceability analytics is based upon the digital chromosomes spatial similarity given indirect contact interval time window. Figure 6 demonstrates the traceability analytics procedure of how to investigate indirect contacts with high spatial temporal correlation.

Figure 6

Traceability analytics also includes the collection phase which we already discussed in the previous subsection, the analytics phase is executed at the user smartphone to leverage the edge computing capacity. We first acquire the indirect contact interval time window *IT,* which refers to the interval appearance time of two subjects given by the health authorities, which usually take 1 or 2 weeks. The spatial similarity threshold value is predefined for similarity classification. If the venue beacon nodes from local stored digital chromosomes are the same as the one from the suspected fingerprint list by judging its encrypted UUID while the interval time of two timestamps of local record and the suspected list is less than the indirect contact interval time window, a k-NN approach for digital similarity value calculation is conducted.

We first transfer the RSSI to distance using a signal propagation model. Before converting the RSSI to distance, The RSSI should be filtered since the multipath effect and interference. Kalman Filter is then conducted to remove RSSI noise to refine results. Kalman Filter is one of the most widely used methods for tracking and estimation due to its simplicity, optimality, tractability, and robustness. It is a state estimator that estimates some unobserved variable based on noisy measurements. Given the filtered RSSI, the log-normal shadowing path loss model is established to generate the physical distance between the emitter to the receiver. It is shown as:

$$R = R_0 - 10n\log_{10}\left({}^{d_m}/{d_0}\right) + X_\sigma \quad (1)$$

where the received signal strength R (expressed in *dBm*) is regarded as a logarithmic function of the distance (d_m in meters) between the receiver and the emitter from an unknown surrounding target. R_0 is the mean power received at the reference distance d_0, usually in 1 meter, the n variable is the propagation constant which is environment-dependent. X_σ is a zero-mean Gaussian random variable. These parameters depend on the environment and the operating frequency. The estimated distance d_m can be calculated. Due to the privacy issue, the relative location is recorded instead of an absolute location. The coordinates of the venue beacon nodes are not concerned however the relationship of the iGather user with venue beacon nodes is logged. In order to make the tracking mechanism clearly stated. A biology metaphor is adopted for the problem. A large amount of genes forms up DNA, the chromosome is finally established with DNA and protein. One piece of the gene can be denoted as $G_u^i = \{PD_{(i,u)}\}$ which represents the one physical distance between venue beacon node i to the iGather user u. The DNA of iGather user u is actually the combination of the gene from iGather user u to each venue beacon node which can be established as $DNA_u = \{PD_{(1,u)}, PD_{(2,u)} ... PD_{(i,u)}\}, i \in N^*$. The protein is actually the timestamp t, therefore the chromosome of the iGather user u at the time window t can be defined as $CRS_u^t = \{DNA_u^t\}$. This chromosome does not reveal the detailed location information of the iGather user, it only states the relationship (distance) from iGather user to each receivable venue beacon node with the correct timestamp. A virtual location is generated which depicts the relative location of iGather user and venue. The spatial temporal chromosome is stored locally as contact history in the iGather user's smartphone without uploading to the cloud. If one iGather user is diagnosed as a COVID-19 patient, public health professionals have the privilege to access the user data including its UID, locally stored digital chromosomes. The information is synchronized to update the suspected digital chromosome list. Other iGather users automatically download the caution list in a real-time manner. The iGather starts to filter the possible indirect contact history. The indirect contact tracing in traceability analytics means the iGather user may not appear simultaneously with the confirmed case, but still have certain commonness (e.g. confirmed case sits on the table to eat dinner at 18:00 and the other user sits on the same table at 21:00). Similar chromosome refers to the similar spatial venue visiting history given a specific contact time window. A k-nearest neighbour algorithm is adopted to measuring similarity between two chromosomes as follows:

$$d_E(CRS_u^t, CRS_c^{St}) = \frac{1}{n}\sqrt{\sum_{i\,=\,1,\ldots,n}(PD_{(i,c)} - PD_{(i,u)})^2}$$

The Euclidian distance between the two chromosomes is the quantitative indicator of similarity. If the calculated Euclidian distance is less than the spatial similarity threshold value, therefore the two chromosomes have great similarities. The indirect contact can be confirmed given a high spatial correlation. Moreover, if the time window is restricted to the same certain period, the direct contact history can also be tracked through traceability analytics.

5 DEMONSTRATION AND TESTING

5.1 Experimental studies

A group of lab mates acts as individuals. Only smartphones and venue beacon nodes are used for data collection and transmission through Bluetooth functions, respectively. Two scenarios in one venue are demonstrated and tested.

One restaurant is selected as the venue where it has multi-levels and ventilation systems. Five venue beacon nodes (VBN_1, VBN_2, VBN_3, VBN_4, and VBN_5) are dispersed deployed at fixed locations in the venue. iBeacon protocol, released by Apple. Inc., is adopted to regulate the broadcasting standards. The UUID of iBeacon is 16 bytes string used to differentiate the venue identification. The major index is a 2-byte string meant to identify the ventilation cluster. Venue beacon nodes in the same venue and under the same ventilation system have the same UUID and major value. The minor value is also a 2-byte string meant to identify individual venue beacon nodes. These fine venue beacon nodes broadcast Bluetooth signal packets to nearby devices constantly.

Figure 7

Figure 7 demonstrates the digital chromosome collection phase during spatial temporal analytics in iGather mobile application. The digital chromosome is generated when the smartphone receives the Bluetooth signal packets. Lab mates are asked to act as individuals to visit the restaurant to have lunch in order to simulate the situation of two high spatial temporal correlated indirect contact tracing. iGather mobile application is installed to collect their spatial temporal data as input for analytics. Table I illustrates the digital chromosome of collected RSSI of 4 individuals in the venue. Beacon nodes in the same venue have the same UUID during the registration. VBN_1, VBN_2, and VBN_3 have the same major value which means these three beacon nodes are deployed under the same ventilation system and VBN_4 and VBN_5 under another ventilation system. The minor value specifies the unique identification of the venue beacon node. Two time windows *Ta* (14:06-14:17) and *Tb* (14:17-14:21) are defined. R_{Ta} refers to the received signal strength indicator at time window *Ta*. All the digital chromosomes are locally stored at smartphone. If I_1 is confirmed to be COVID-19 positive, the digital chromosome is then acquired by the public health authorities and push to all iGather users for mobility analytics and traceability analytics.

Table I

Mobility analytics: The mobility analytics is concerned with people gathered in the same time window and they are geographically distant from others (long-range), which exceeds Bluetooth transmission distance, from user smartphone but within the coverage area of the same ventilation system. Other individuals I_2, I_3, and I_4 receive suspected digital chromosomes list from iGather cloud server. VBN_1, VBN_2, and VBN_3 have the same UUID and major value which means these three venue beacon nodes are deployed in the same venue and under the same ventilation system. During the time window *Ta* where the infected case I_1 appears, I_2 also receives signals from the venue beacon node in the same ventilation system, which can be inferred that I_2 is an indirect contactor since the venue beacon node has the same UUID and major value at the same time window.

Traceability analytics: The traceability analytics is concerned with people gathered not in the same time window and they are geographically close (short-range). The digital chromosomes are organized for traceability analytics in Figure 8.

Figure 8

iGather user I_1 is unfortunately diagnosed as COVID-19 confirmed case. iGather User I_3 and I_4 download the updated suspect digital chromosomes list real-timely. Despite the received *VBN* signals reveals that I_3 or I_4 does not appear simultaneously with I_1 considering the time window is not consent. However, the adjacent time window explains they may have been to similar (nearby) places in time sequence. The genes are the distance between user and venue beacon nodes using the signal propagation model. The indirect contact history of mobility analytics is conducted. The k nearest neighbour calculates the average Euclidian distance among the chromosomes is listed as follows:

Table II

If the predefined similarity threshold value is 10. From Table II, the similarity value of iGather user I_3 with I_1 is less than the threshold value which geographically close to I_1 with a high level of risk. (e.g. After I_1 leaving, I_3 may use the same table or close table). The similarity value of iGather user I_4 with I_1 is over the threshold value which has relatively low level of risk. (e.g. I_4 may also enter the same restaurant, but sit far from the table that I_1 used to sit). Traceability analytics can also realize the direct contact tracing by confining the timestamp of contactors in the same time window.

To verify the effectiveness of iGather solution, we conduct quantitative analysis in The University of Hong Kong-Zhejiang Institute of Research and Innovation. The campus includes various offices, pantries, elevators, and other venues for contact tracing testing. 15 venue beacon nodes are deployed at a specific location in the physical space in Figure 9 and configured in the cyber world in Figure 10(a). Office 301 to 304 in level 3 share the same ventilation system for mobility analytics. 20 staff are invited as volunteers for contact tracing purposes. The iGather smartphone application records the digital chromosome locally when visiting 15 venues. This test records the volunteer's movement history as a digital chromosome for one week time. The individuals are randomly picked as contagious ones, the iGather manager then requests them to upload digital chromosomes in Figure 10(b) and other users conduct the spatial-temporal analysis to alert the close contacts.

Figure 9

Figure 10

The contact tracing results in Table III show that the mobility analytics can reach 100% close contact tracing. However, the traceability analytics show satisfying results only when the similarity threshold value sets to 20. The reason is that the multi-path effect and non-line of sight confuses the signal readings which has attenuation effect. If the threshold value is too small, the tractability analytics may decide the people gathered not geographically close in different time windows.

Table III

5.2 Comparative analysis

In order to evaluate the contact tracing application, a comparative analysis is conducted based on the contact tracing evaluation framework from Dar et al. (2020). It is a five-step method for evaluating a particular contact tracing solution including the nature of the model, technology employed, privacy, tracing model, and scalability.

The nature of the model concerns the architecture of the solution. A centralized system is where all the users store the data in one server or authority. The tracing calculations are conducted in the central server to compare and match all relevant users' data. While the decentralized system is one where no entity control over the whole users' data. All the information is stored in a distributive manner. Tracing calculations are computed in edge devices to offload the computing pressure of a single server. The technology employed mainly consider the underlying technology for contact tracing purpose. Passive tracking manner such as health code using QR code is based on the active report from the user. Active tracking such as GPS and proximity-based approach can record the user movement action without inputting information manually. The privacy is always the first step in devising the contact tracing solutions. We adopt the notions from Cho et al. (2020) to label them as L1: privacy from snoopers, L2: Privacy from contacts, and L3: Privacy from authorities. The tracing model is the key index to evaluate whether the solution is able to identify close contact based on different spatial-temporal correlations. Scalability concentrates on whether the solution can remain effective when encountered mass users and data. Several CTAs are selected to conduct the comparative analysis. The following figure is the comparative analysis of evaluated solutions.

Table IV

As can be seen from Table III, the health code system relies on the active report from the user, and the information is collected and matched on the centralized server. The centralized architecture requires all the information uploaded to central server which may have potential privacy concerns toward authorities. Moreover, the centralized architecture may affect the scalability since more user and data generated from the edge side increase the burden of the server. Decentralized architecture takes full advantage of edge computing capability in users' smartphones to alleviate pressure on a single server. Currently, the iGather is the only solution that can realize the indirect tracing.

5.3 Managerial insights

The results show that the indirect contact tracing can be realized through spatial temporal analytics. Smartphones do not send but only receive signals to alleviate privacy concerns. Edge computing is fully leveraged to analyse locally stored data without uploading it to the central server. Furthermore, from a management perspective of mobilizing social capacity, the venue not only plays a promotion role in boosting the utilization rates but also a supervision-assisted role for keeping the venue in a safe and healthy situation. The health authorities can protect the user's privacy while urging close contacts to seek medical services through venue management.

6 CONCLUDING DISCUSSIONS

Physical gathering in venues such as offices, sports clubs, restaurants, pubs, and public transports will resume normal working and living styles to keep economic and social development. In addition to wearing masks, technological means are also needed to assist in tracking close contacts. Bluetooth, together with other emerging technologies, contributes to realizing the direct contact tracing through exchanging messages among user smartphones given the high adoption rate. However, recent research efforts have indicated that COVID-19 viruses are transmitted during exhalation, talking and coughing in droplets small enough to remain aloft in the air and inadequate ventilation system caused infections, which pose risks to people who do not even have direct contact history. To answer these questions, this study proposed iGather: IoT-enabled cyber-physical architecture as a technical pillar to track indirect contacts oriented in venues. Two spatial temporal analytics: mobility and traceability analytics are presented to realize the tracing of two high spatial-temporal correlated scenarios. Locations with timestamps are modeled as digital chromosome to represent the real world instance through digital twin conceptions. Demonstrations and testing are conducted in a restaurant to verify the feasibility and efficacy of the overall solution. Results show that the indirect tracing can be realized through the deployment of physical hardware and spatial temporal analytics.

This research makes some contributions to the joint fighting of the COVID-19 pandemic as follows. First, the airborne transmission of viruses may pose threats to people who gathered in an inadequate ventilation venue, the mobility analytics contributes to tracking indirect contacts at relatively long distance but within the same ventilated environment and time window. Second, aerosol, the microdroplets remain aloft in air generated through exhalation, talking, and coughing, still a risk for latecomers who do not appear simultaneously with a confirmed case. The traceability analytics realize the tracking indirect contacts with different time windows but similar locations. Third, concerning privacy-preserving of Bluetooth enabled contact tracing, the information or message sending from the user smartphone to confirm nearby close contact cause user discomfort even encryption or anonymization adopted. We propose venue beacon nodes in cyber-physical architecture to emit signal packets instead of using smartphones to protect privacy from snoopers and other users. The edge computing allows local calculations for close contacts which alleviate central server pressure but also guarantee privacy from health authorities.

REFERENCES

AHN, M. J. (2020), "Combating COVID-19: Lessons from South Korea".

BRAITHWAITE, I.CALLENDER, T.BULLOCK, M. and ALDRIDGE, R. W. (2020), "Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19". *The Lancet Digital Health,* Vol. 2 No. 11. e607-e621.

CGTN. 2020. *How China moved to contain COVID-19: lessons and reflections* [Online]. Available: https://news.cgtn.com/news/2020-06-07/How-China-moved-to-contain-COVID-19-lessons-and-reflections--R7KfgfWt2M/index.html [Accessed 2020-11-03.

CHO, H.IPPOLITO, D. and YU, Y. W. (2020), "Contact tracing mobile apps for COVID-19: Privacy considerations and related trade-offs". *arXiv preprint arXiv:2003.11511*.

DAR, A. B.LONE, A. H.ZAHOOR, S.KHAN, A. A. and NAAZ, R. (2020), "Applicability of mobile contact tracing in fighting pandemic (covid-19): Issues, challenges and solutions". *Computer Science Review*. 100307.

DING, K.CHAN, F. T. S.ZHANG, X.ZHOU, G. and ZHANG, F. (2019), "Defining a Digital Twin-based Cyber-Physical Production System for autonomous manufacturing in smart shop floors". *International Journal of Production Research,* Vol. 57 No. 20. 6315-6334.

DRAZEN, J. M.KANAPATHIPILLAI, R.CAMPION, E. W.RUBIN, E. J.HAMMER, S. M.MORRISSEY, S. and BADEN, L. R. 2014. Ebola and quarantine. Mass Medical Soc.

GLAESSGEN, E. and STARGEL, D. The digital twin paradigm for future NASA and US Air Force vehicles. 53rd AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and materials conference 20th AIAA/ASME/AHS adaptive structures conference 14th AIAA, 2012. 1818.

GUO, D.LI, M.ZHONG, R. and HUANG, G. (2020), "Graduation Intelligent Manufacturing System (GiMS): an Industry 4.0 paradigm for production and operations management". *Industrial Management & Data Systems*.

HATKE, G. F.MONTANARI, M.APPADWEDULA, S.WENTZ, M.MEKLENBURG, J.IVERS, L.WATSON, J. and FIORE, P. (2020), "Using Bluetooth Low Energy (BLE) signal strength estimation to facilitate contact tracing for COVID-19". *arXiv preprint arXiv:2006.15711*.

HAWRYLUCK, L.GOLD, W. L.ROBINSON, S.POGORSKI, S.GALEA, S. and STYRA, R. (2004), "SARS control and psychological effects of quarantine, Toronto, Canada". *Emerging infectious diseases,* Vol. 10 No. 7. 1206.

HERNáNDEZ-ORALLO, E.MANZONI, P.CALAFATE, C. T. and CANO, J.-C. (2020), "Evaluating how smartphone contact tracing technology can reduce the spread of infectious diseases: the case of COVID-19". *IEEE Access*.

KRETZSCHMAR, M. E.ROZHNOVA, G.BOOTSMA, M. C.VAN BOVEN, M.VAN DE WIJGERT, J. H. and BONTEN, M. J. (2020), "Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study". *The Lancet Public Health,* Vol. 5 No. 8. e452-e459.

LI, J. and GUO, X. (2020), "Global Deployment Mappings and Challenges of Contact-tracing Apps for COVID-19". *Available at SSRN 3609516*.

LI, Y.QIAN, H.HANG, J.CHEN, X.HONG, L.LIANG, P.LI, J.XIAO, S.WEI, J. and LIU, L. (2020), "Evidence for probable aerosol transmission of SARS-CoV-2 in a poorly ventilated restaurant". *medRxiv*.

LUO, S.LIU, H. and QI, E. (2019), "Big data analytics – enabled cyber-physical system: model and applications". *Industrial Management & Data Systems,* Vol. 119 No. 5. 1072-1088.

MCLACHLAN, S.LUCAS, P.DUBE, K.HITMAN, G. A.OSMAN, M.KYRIMI, E.NEIL, M. and FENTON, N. E. (2020), "Bluetooth Smartphone Apps: Are they the most private and effective solution for COVID-19 contact tracing?". *arXiv preprint arXiv:2005.06621*.

MORAWSKA, L. and MILTON, D. K. (2020), "It is time to address airborne transmission of covid-19". *Clinical Infectious Diseases*.

NUSSBAUMER-STREIT, B.MAYR, V.DOBRESCU, A. I.CHAPMAN, A.PERSAD, E.KLERINGS, I.WAGNER, G.SIEBERT, U.CHRISTOF, C. and ZACHARIAH, C. (2020), "Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review". *Cochrane Database of Systematic Reviews,* No. 4.

ORGANIZATION, W. H. 2020. Digital tools for COVID-19 contact tracing: annex: contact tracing in the context of COVID-19, 2 June 2020. World Health Organization.

PANZZARINO, M. (2020), "Apple and Google are launching a joint COVID-19 tracing tool for iOS and Android". *Tech Crunch,* Vol. 10.

RICHARDS, A. D. (2020), "Ethical guidelines for deliberately infecting volunteers with COVID-19". *Journal of Medical Ethics*.

SINGAPORE, M. O. H. 2020. *TraceTogether* [Online]. Available: https://www.tracetogether.gov.sg/ [Accessed 2020-11-03.

STEVENS, H. and HAINES, M. B. (2020), "TraceTogether: Pandemic Response, Democracy, and Technology". *East Asian Science, Technology and Society*.

TAGHRIR, M. H.AKBARIALIABAD, H. and MARZALEH, M. A. (2020), "Efficacy of mass quarantine as leverage of health system governance during COVID-19 outbreak: a mini policy review". *Archives of Iranian medicine,* Vol. 23 No. 4. 265-267.

TAO, F.SUI, F.LIU, A.QI, Q.ZHANG, M.SONG, B.GUO, Z.LU, S. C. Y. and NEE, A. Y. C. (2019), "Digital twin-driven product design framework". *International Journal of Production Research,* Vol. 57 No. 12. 3935-3953.

TU, M.MING, K. L. and YANG, M.-F. (2018), "IoT-based production logistics and supply chain system – Part 2: IoT-based cyber-physical system: a framework and evaluation". *Industrial Management & Data Systems,* Vol. 118 No. 1. 96-125.

WANG, J. and DU, G. (2020), "COVID-19 may transmit through aerosol". *Irish Journal of Medical Science (1971 -)*.

WATTS, D. (2020), "COVIDSafe, Australia's Digital Contact Tracing App: The Legal Issues". *Australia's Digital Contact Tracing App: The Legal Issues (May 2, 2020)*.

WILDER-SMITH, A. and FREEDMAN, D. O. (2020), "Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak". *Journal of travel medicine,* Vol. 27 No. 2. taaa020.

XIA, Y. and LEE, G. (2020), "How to return to normalcy: Fast and comprehensive contact tracing of COVID-19 through proximity sensing using mobile devices". *arXiv preprint arXiv:2004.12576*.

ZHANG, M.CHOW, A. and SMITH, H. (2020), "COVID-19 Contact-Tracing Apps: Analysis of the Readability of Privacy Policies". *J Med Internet Res,* Vol. 22 No. 12. e21572.

ZHONG, R. Y.XU, C.CHEN, C. and HUANG, G. Q. (2017), "Big Data Analytics for Physical Internet-based intelligent manufacturing shop floors". *International Journal of Production Research,* Vol. 55 No. 9. 2610-2621.

ZHOU, Y.MA, Z. and BRAUER, F. (2004), "A discrete epidemic model for SARS transmission and control in China". *Mathematical and Computer Modelling,* Vol. 40 No. 13. 1491-1506.