

# Epidemic Risk Assessment by a Novel Communication Station Based Method

Zhaoquan Gu<sup>ID</sup>, Le Wang<sup>ID</sup>, Xiaolong Chen, Yunyi Tang, Xingang Wang, Xiaojiang Du<sup>ID</sup>, *Fellow, IEEE*, Mohsen Guizani<sup>ID</sup>, *Fellow, IEEE*, and Zhihong Tian<sup>ID</sup>, *Member, IEEE*

**Abstract**—The COVID-19 pandemic has caused serious consequences in the last few months and trying to control it has been the most important objective. With effective prevention and control methods, the epidemic has been gradually under control in some countries and it is essential to ensure safe work resumption in the future. Although some approaches are proposed to measure people's healthy conditions, such as filling health information forms or evaluating people's travel records, they cannot provide a fine-grained assessment of the epidemic risk. In this paper, we propose a novel epidemic risk assessment method based on the granular data collected by the communication stations. We first compute the epidemic risk of these stations in different intervals by combining the number of infected persons and the way they pass through the station. Then, we calculate the personnel risk in different intervals according to the station trajectory of the queried person. This method could assess people's epidemic risk accurately and efficiently. We also conduct extensive simulations and the results verify the effectiveness of the proposed method.

**Index Terms**—COVID-19, communication station risk, epidemic risk assessment, personnel risk.

## I. INTRODUCTION

THE outbreak of the coronavirus (COVID-19) pandemic has caused very serious consequences in the last few months. According to the statistics, as of July 21 of 2020, the COVID-19 pandemic has spread to 215 counties around the world and a total of 15 million people have been infected. The epidemic has caused nearly 610 000 deaths and caused huge economical losses in almost every country. According to the research report [27], the COVID-19 pandemic has caused 3.8 trillion economic losses and 147 million people lost their

jobs. Due to the strong infectious nature of the virus, many cultural activities and sport events have been suspended or cancelled, which has greatly affected people's daily life. The COVID-19 pandemic has become a very serious problem in every country, and the most important and crucial thing is to fight the epidemic.

Many governments have issued policies to fight the epidemic, such as quarantine, isolation guidelines, social distancing, and suspending many large gathering activities. Meanwhile, many scientific researchers have denoted themselves to the prevention and control of the epidemic. Medical experts are studying the characteristics and the spreading ways of the virus, and are committed to develop vaccines against this virus. For example, the clinical characteristics of COVID-19 patients are studied in [39] and it provided reliable medical evidence for the treatment and prevention of COVID-19. In [37], it verifies that the strict prevention and control methods in China are effective in protecting people from the virus. With the fast development of information technologies [14], [16], [43], [44], many scientific researchers from the computer science field have proposed efficient methods to assist the prevention of the epidemic. For example, a knowledge graph of the COVID-19 is introduced in [7], which discusses the basis knowledge of the epidemic. A mathematical model of predicting the spread of the COVID-19 pandemic is proposed in [20], which helps understand the spreading range and impact of the epidemic. To acquire the public's perception of the COVID-19 epidemic, an intelligent model is introduced in [26] by analyzing twitter data.

Due to the effective prevention and control methods, the epidemic has gradually been under control in some countries, such as in China. Currently, it is an important issue to resume the work and production safely under this situation. *How to assess the personnel's epidemic risk is an essential and important problem in assisting resuming work safely?* There are mainly two approaches taken by the government and enterprises. The first one is to ask people to report his/her daily health information, such as filling health information forms. However, some people might report the information incorrectly intentionally or accidentally, which would cause serious consequences if some infected people are allowed to work with other colleagues. The second approach is adopting applications that show the epidemic risk according to people's travel records (such as flights, trains, etc) and the cities he/she has been in the last half a month. However, many applications

Manuscript received August 1, 2020; revised December 29, 2020; accepted February 7, 2021. Date of publication February 11, 2021; date of current version January 11, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant U1636215, Grant 61902082, the National Key Research and Development Program of China 2018YFB1004003, the Guangdong Key Research and Development Program of China 2019B010136003, and in part by the Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2019). Recommended for acceptance by Dr. Iztok Huma. (*Corresponding authors: Le Wang and Zhihong Tian.*)

Zhaoquan Gu, Le Wang, Xiaolong Chen, Yunyi Tang, Xingang Wang, and Zhihong Tian are with the Cyberspace Institute of Advanced Technology (CIAT), Guangzhou University, Guangzhou 510006, China (e-mail: zqgu@gzhu.edu.cn; wangle@gzhu.edu.cn; 2111806010@e.gzhu.edu.cn; 2111806047@gzhu.edu.cn; 2111906082@e.gzhu.edu.cn; tianzhihong@gzhu.edu.cn).

Xiaojiang Du is with the Department of Computer and Information Sciences, Temple University, Philadelphia, PA 19122 USA (e-mail: dxj@ieee.org).

Mohsen Guizani is with the Computer Science and Engineering Department, Qatar University, Doha 2713, Qatar (e-mail: mguizani@ieee.org).

Digital Object Identifier 10.1109/TNSE.2021.3058762

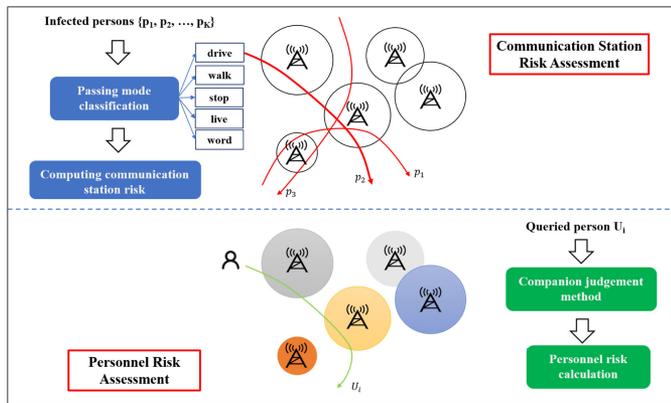


Fig. 1. The epidemic risk assessment method has two stages: the first stage computes the risk of each communication station while the second stage computes the risk of the queried person.

compute the epidemic risk by judging whether the person has been to a city where infected cases exist or not; it is inaccurate such as a healthy person drives through the city. In addition, a city or a district contains a large number of people, computing the epidemic risk in a more refined way is needed in assisting resuming work safely.

In this paper, we propose a novel communication station based method to assess the people's epidemic risk. Since almost everyone is using mobile phones for communication, it is sufficient to judge a person's trajectory according to the mobile phone's trajectory. A mobile phone would exchange information with the nearest communication station timely, and a communication station covers a relative small range, such as a few hundred meters in residential areas and a few kilometers in open areas such as a highway. The station records the time interval that a mobile phone passes through its communication range, and it provides more accurate information for assessing the epidemic risk. Compared with the traditional two above discussed approaches, the communication station based epidemic risk assessment method could utilize more fine-grained data to provide more accurate assessment.

There are three challenges in handling the epidemic risk assessment by the communication station information. First of all, people's trajectory data (collected by the communication stations) has significant temporal and spatial characteristics, such as the time interval of passing through the communication station is the temporal characteristic while the physical locations of the station is the spatial characteristic. It is difficult and inefficient to utilize such data directly. Second, there are many different ways that a person passes through or stays within a station, such as driving a car, walking, staying at home, or staying in the office, it is necessary to consider different situations in assessing a person's epidemic risk. Third, some people may move together with some infected persons, it is needed to find out the person with high epidemic risk according to the station's information.

In this paper, our proposed epidemic risk assessment method addressed all these challenges. As depicted in Fig. 1, there are two stages in assessing a person's epidemic risk. We first evaluate the epidemic risk of each communication station

in different time intervals according to the trajectories of the infected people. Then, we evaluate a person's epidemic risk according to the stations that he/she passes through. By calculating the epidemic risk of each station in each time interval, we could efficiently utilize people's trajectory data, which addresses the first challenge. In the first stage of our method, we propose a *passing mode classification* method where five passing modes (drive, walk, stop, live, and work) reflect how the person passes through a communication station. By assigning different risk weights to these models, such as driving through the station has less risk weight than walking, we can compute the epidemic risk of each station in each time interval and it addresses the second challenge. In the second stage of our method, we introduce a *companion judgement method* which automatically decides whether two people move together. After that, we compute the epidemic risk of a person according to the epidemic risks of the passed communication stations, which addresses the third challenge.

In order to verify the efficiency of the proposed method, we also conduct extensive simulations. We generate a large number of communication stations and the trajectories of some infected persons according to the released information by the Centers for Disease Control (CDC). We generate different types of people's trajectory and our method could compute each person's epidemic risk accurately, which verifies the effectiveness of the proposed method. We summarize the contributions of this work as follows:

- 1) We propose a novel epidemic risk assessment method on the basis of the communication station's information, which provides a fine-grained assessment in assisting resuming work safely;
- 2) We first compute the epidemic risks of each station in different time intervals, and then a person's epidemic risk is calculated by combining the passed communication stations. This two-stage method could help compute a person's epidemic risk accurately and efficiently;
- 3) We conduct extensive simulations to evaluate our method and the results confirm its effectiveness.

The rest of the paper is organized as follows. The next section highlights the related work about the COVID-19 epidemic. Then, we introduce the preliminaries in Section III. We introduce the first stage of computing the epidemic risk of each station in Section IV and present the second stage of calculating a person's epidemic risk in Section V. We describe the simulation results in Section VI and we conclude the paper in Section VII.

## II. RELATED WORK

During the outbreak of COVID-19 epidemic, there have been a large number of research attempts analyzing the epidemic's characteristics, providing prevention and control applications, and proposing epidemic risk assessment algorithms. In this section, we introduce these related attempts in three categories. First, we introduce research on the characteristics of the epidemic in Section II-A, then we provide an overview of various methods and applications to prevent and

control the epidemic in Section II-B. Finally, we discuss some methods in assessing the epidemic risk in Section II-C.

### A. Characteristics of the COVID-19 Epidemic

In December of 2019, the first released case of the unknown pneumonia appeared in Wuhan [18]. Though high-throughput sequencing, it revealed a novel coronavirus that is named COVID-19 officially [25]. The new virus resembled severe acute respiratory syndrome coronavirus (SARS-CoV) [50].

The COVID-19 coronavirus spreads rapidly from person to person, with a median incubation period of three days and a relatively low fatality rate. By studying some infected patients, some preliminary studies immediately concluded that older males and those with comorbid conditions are more likely to become severe cases. Considering the spreading routes, the virus spreads mainly through direct contact and droplets, and it is likely to be transmitted even in a brief contact with an infected person, such as chatting and/or eating. Until now, there is no directly effective vaccine against the virus and the treatments are mainly symptomatic support therapy, as many antibiotics and antiviral drugs were confirmed ineffective. Even though the COVID-19 virus is highly contagious, according to the latest research, some control methods have proven to be very effective, particularly wearing masks, which could prevent most infections [3], [11], [13], [21], [28], [29], [35], [36], [47].

Some works characterized the temporal and spatial dynamics of the COVID-19 epidemic. The spatial and temporal scan statistics are adopted to quantify the trends of the COVID-19 according to the weekly released infected cases in [46], and it is suggested that the temporal and spatial trends during the spread of the COVID-19 epidemic should be taken into consideration when people are prepared to return to their normal routines.

### B. Applications of Epidemic Prevention and Control

With the fast development of information technologies [2], [8]–[10], [15], [42], many methods have been proposed for the epidemic prevention and control.

The Internet of Things (IoT) have been adopted in medicine and it is referred to as the medical IoT (MIoT) [4], [40]. MIoT aims to establish a decision-oriented big data analysis model that is supported by many latest information technologies such as communication technologies, electronics, biology, and medicine. A point of view is proposed in [33] that human were ill-prepared to deal with the challenges the COVID-19 epidemic has posed. Future research should attempt to address the uses and implications of IoT technologies for mapping the spread of infection.

In China, nCapp (COVID-19 Intelligent Diagnosis and Treatment Assistant Program) [1] is proposed for early COVID-19 diagnose and enhance treatment. According to the available data, questionnaire and examination results, the program can automatically classify a case to different diagnostic results such as confirmed case, suspected case, or normal cases. The application also updates a real-time database of the

COVID-19 epidemic, and updates the diagnostic model timely on the basis of the latest confirmed case to improve the diagnostic accuracy. In addition, nCapp can guide physicians, specialists, and managers in counseling and prevention of the epidemic. The ultimate goal of the application is to raise the level of COVID-19 diagnosis and to improve the treatment of different doctors in different hospitals to a high level (such as national and international levels) through the assistance of the system. 5 G wireless transmission technology is also introduced in the field of medical care for faster video transmission in [6], and it is suggested that more cutting-edge technologies such as IoT, big data, and artificial intelligence could largely improve the medical services [38].

Some works also analyze the spreading of the epidemic. The susceptible exposed infectious recovered (SEIR) model is introduced in [19] to analyze the effectiveness of the quarantine approaches of Wuhan, and the result shows that by reducing the contacts with infected cases or potentially infected cases, many interventions could reduce the number of infected cases and help prevent the spread of the virus efficiently by these quarantine and isolation methods. A disease spreading model based on SEIR is proposed in [5] to simulate the spreading of the disease and it uses several parameters to quantify different ways to control the disease spreading. A probabilistic model that describes the worldwide spread of the infectious diseases is proposed in [17] and it demonstrates that a forecast of the geographical spread of the epidemic is indeed possible.

### C. Assessment of the Epidemic Risk

The outbreak and spread of the COVID-19 epidemic is a complex system constituted by etiology, host, and environment. It is not only related to the spread method between people, but also related to the crowd of social relations and daily behaviors [30]. Particularly, in a large-scale mobile population, close contact with each other would directly accelerate the spread of the virus [32], [45]. To assess the epidemic risk of a specific district of a specific person accurately is important for timely epidemic control, but is also very challenging.

It is examined in [31] that several key factors are important in forecasting the epidemic risk. It is assumed that several factors such as the geographical scale and the temporal duration are identified as important components in predicting the epidemic risk accurately. A similar assessment system has been developed in [41] to forecast the probabilities of anomalously high and low malaria incidence, with a dynamical, seasonal-timescale, multi-model ensemble prediction model of the climate. This forecast system has been successfully applied to the prediction of malaria risk in Botswana, where links between malaria and climate variability are well established, and it has a comparably high level of prediction accuracy. A risk assessment method of the infectious disease is presented in [12], which calculates the risk based on the spatiotemporal trajectory data. This method uses the spatial density and the average speed of the users to determine the risk of infectious disease, and it is assumed that it can be utilized to evaluate the

TABLE I  
NOTATIONS FOR THE EPIDEMIC RISK ASSESSMENT PROBLEMS

Notation	Description
$S = \{S_1, S_2, \dots, S_N\}$	The communication stations
$(x_i, y_i)$	The coordinates of station $S_i$
$\{U_1, U_2, \dots, U_M\}$	$M$ different persons
$Q(U_i)$	The communication station trajectory of user $U_i$
$[t_{j_1}, t_{j_2})$	The time interval of passing $S_j$
$\Delta$	The time length of each slot
$[n\Delta, (n+1)\Delta)$	The time interval of $t_n$
$R(S_i, t_n)$	The communication station risk of $S_i$ during $t_n$
$R(U_i, t_n)$	The personnel risk of $U_i$ during $t_n$
$P = \{P_1, P_2, \dots, P_K\}$	A set of infected and suspected persons

high-risk areas according to different data sources. A new approach of the epidemic risk assessment is proposed in [34], which constructs a novel data-driven framework for assessing the epidemic risk of a geographical area and identifying high-risk areas within a country. The risk value is evaluated as a function of three different components: the hazard of the disease, the exposure of the area and the vulnerability of its inhabitants.

In this paper, we utilize people's trajectory to assess the epidemic risk. Some works utilize WiFi sensing technology or crowdsensed data to track human mobility [22]–[24], [48], [49]. Different from these works, we utilize the trajectory data that are collected by the communication stations. On the basis of that data, we can calculate the epidemic risk of the stations and the epidemic risk of each person using a more integrated approach.

### III. PRELIMINARIES

In this section, we first introduce the system model, including the commonly adopted notations in the paper. Then, we present formal definitions of the epidemic risk assessment problem, including communication station risk assessment and personnel risk assessment.

#### A. System Model

For the convenience of description, we denote the communication stations as set  $S = \{S_1, S_2, S_3, \dots, S_N\}$  where  $S_i$  represents a specific and unique station. Each communication station has its position coordinates and we denote the coordinates of station  $S_i$  as  $(x_i, y_i)$ , which represent the longitude and latitude of the communication station respectively. Denote a set of  $M$  different persons as  $\{U_1, U_2, U_3, \dots, U_M\}$  (we omit the situation that a person carries multiple mobile phones). For each person  $U_i$  who moves among the communication stations, each station  $S_j$  would record the time when the person enters and leaves the communication range and we denote the time interval as  $[t_{j_1}, t_{j_2})$  where  $t_{j_1} < t_{j_2}$ .

In practical scenarios, different communication stations have different sensing ranges and a person's mobile phone might be connected to multiple stations. In addition, when a person enters indoor or goes underground, the stations might not detect the person correctly. To simplify the problem, we assume a person's mobile phone is only connected to the nearest station at

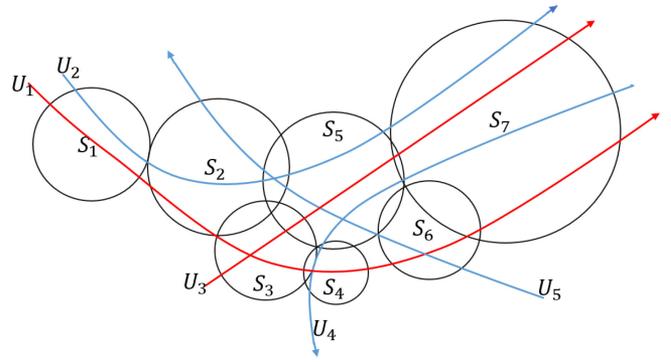


Fig. 2. An illustration example with seven communication stations and five persons.

any time; if a person is not detected by any station, we assume he/she is connected to the previous station. Hence, we can denote the communication station trajectory of person  $U_i$  as a sequence  $Q(U_i) = \{(S_k, [t_{k_1}, t_{k_2}]), (S_j, [t_{j_1}, t_{j_2}]), \dots\}$ , where  $S_k, S_j, \dots$  are the communication stations that the person passed.

In order to count the number of persons passing through a specific communication station in different time intervals, we first divide the time period of everyone passing through the station according to a certain time granularity  $\Delta$ , which implies the time is slotted with length  $\Delta$ . Therefore, denoting the outbreak time of the epidemic as the initial time slot 0, we denote the other time segments as  $t_0 = [0, \Delta)$ ,  $t_1 = [\Delta, 2\Delta)$ ,  $\dots$ ,  $t_i = [i\Delta, (i+1)\Delta)$ ,  $\dots$ . In order to compute the epidemic risk in a more fine-grained way, we set  $\Delta = 600$  s in our simulation and each time segment contains 10 minutes. Actually,  $\Delta$  can be set as other values according to the system requirements. All these notation are listed in Table I.

For example, seven communication stations and five persons are depicted in Fig. 2, where different stations have different communication ranges as the circles. Considering the trajectory of person  $U_1$  who passes through stations  $S_1, S_2, S_3, S_4, S_6, S_7$ , denote the trajectory as  $Q(U_1) = \{(S_1, [0, 3\Delta)), (S_2, [4\Delta, 6\Delta)), (S_3, [6\Delta, 9\Delta)), (S_4, [9\Delta, 11\Delta)), (S_6, [15\Delta, 18\Delta)), (S_7, [18\Delta, 25\Delta))\}$ , which represents the enter and leave time of each communication station. Similarly, suppose person  $U_3$  passes through stations  $S_3, S_5, S_7$  sequentially and we denote the trajectory as  $Q(U_3) = \{(S_3, [8\Delta, 12\Delta)), (S_5, [12\Delta, 16\Delta)), (S_7, [16\Delta, 23\Delta))\}$ . With this method, we can efficiently represent each person's trajectory by the communication stations' information.

#### B. Problem Definition

In this paper, we assess the epidemic risk of both the communication stations and a specific person. As we divided time into different slots, we denote the **communication station risk** as  $R(S_i, [n\Delta, m\Delta))$ , which implies the epidemic risk value of the communication station  $S_i$  during the time period  $[n\Delta, m\Delta)$ ,  $n < m$ . When  $m = n + 1$ ,  $R(S_i, t_n)$  represents the epidemic risk value of  $S_i$  during time segment  $t_n = [n\Delta, (n+1)\Delta)$ . Similarly, we denote the **personnel risk** as

$R(U_i, [n\Delta, m\Delta])$ , which implies the epidemic risk value of person  $U_i$  during the time period  $[n\Delta, m\Delta]$ ,  $n < m$ . Notice that,  $R(U_i, t_n)$  represents the epidemic risk value during time segment  $t_n$ .

Assuming the current time segment is  $t_{cur}$ , we formulate the problem of computing the communication station risk as:

**Problem 1:** Given a set of infected and suspected persons  $P = \{P_1, P_2, \dots, P_K\}$ , denote the trajectories of them as  $\{Q(P_1), Q(P_2), \dots, Q(P_K)\}$ , compute the communication station risk  $R(S_i, t_n)$  for each station  $S_i$  in each time segment  $t_0 \leq t_n \leq t_{cur}$ .

Taking Fig. 2 as the example, supposing  $U_1, U_3$  are patients (red line), it is easy to check that  $R(S_3, [8\Delta, 9\Delta])$  is high since both  $U_1$  and  $U_3$  pass through the station during the time segment. We need to compute the epidemic risk of all these stations and then to calculate the other persons' epidemic risk values.

Computing the communication station risk plays a fundamental role in calculating the personnel risk. We formulate the problem of calculating the personnel risk as:

**Problem 2:** Given a person  $U_i$  and the trajectory  $Q(U_i) = \{(S_k, [t_{k_1}, t_{k_2}]), (S_j, [t_{j_1}, t_{j_2}]), \dots\}$ , calculate the personnel risk  $R(U_i, t_n)$  for the person in each segment  $t_0 \leq t_n \leq t_{cur}$ .

Normally, it is more important to assess the personnel risk of a specific person during a long period, for example many places need to compute the risk within the last 14 days. Hence, we define the personnel risk during time period  $[n\Delta, m\Delta]$  as

$$R(U_i, [n\Delta, m\Delta]) = \sum_{n \leq j < m} R(U_i, t_j).$$

Considering the work resumption and production, many governments and enterprises require a person's risk to be relatively low (even 0) within the past 14 days. Hence, we can calculate the personnel risk and grant the admission if the risk is 0.

#### IV. COMMUNICATION STATION RISK ASSESSMENT

In this section, we introduce the epidemic risk computation method of the communication stations. According to the station trajectories of the patients  $\{P_1, P_2, \dots, P_n\}$ , the goal is to calculate the risk value of each station  $S_i$  in any time interval  $t_n$  as  $R(S_i, t_n)$ . There are many different ways that a person passes through a communication station, such as driving a car or walking. We first propose a **passing mode classification method** in Section IV-A, which introduces five different passing modes. Then, we present a statistic method counting the persons of different passing modes in Section IV-B and we compute the communication station risk in Section IV-C combining these modes.

##### A. Passing Mode Classification

People could pass through a communication station in different ways, which would affect the station risk computation. For example, patients driving in a car have less effect than patients who are walking during the station range, hence it is necessary to classify a person's station trajectory into different

passing modes. In this paper, we consider five different passing modes which are described as follows:

- 1) Driving mode: a person passes through the communication station in a car or bus (denote the mode as *drive* for short);
- 2) Walking mode: a person walks within the communication station's range (denote the mode as *walk* for short);
- 3) Stop mode: a person stays in a communication station for a long time, for example, he is in a restaurant or a market (denote the mode as *stop* for short);
- 4) Living mode: a person lives at home which is within the communication range (denote the mode as *live* for short);
- 5) Working mode: a person works in a building which is within the communication range (denote the mode as *work* for short).

Different modes should be assigned with different risk weights in computing the communication station's risk. For example, patients of walk mode have larger risk weight than patients of drive mode, work mode has larger risk than live mode since it is more likely to contact more people in the working mode. We introduce the details in classifying the passing modes. Considering the station trajectory of a patient  $Q(P_i) = \{(S_k, [t_{k_1}, t_{k_2}]), (S_j, [t_{j_1}, t_{j_2}]), \dots\}$ , the passing mode classification method aims at classifying the trajectory segment in each communication station  $(S_k, [t_{k_1}, t_{k_2}])$  to the five passing modes. Since the presented five modes have significant patterns, we design some features and rules for classifying the passing modes.

1) *Driving Mode:* Commonly speaking, a person would pass a communication station in a short time by driving a car or taking a bus, hence an important feature of the driving mode is that the time period passing through the station is quite short. We formulate the first classifying rule as

$$t_{k_2} - t_{k_1} \leq T_{S_k, drive}, \quad (1)$$

where  $T_{S_k, drive}$  is a parameter that reflects the time threshold of driving through station  $S_k$ . Suppose the station range is  $r_k$ , for example  $r_k = 500$  meters in the resident areas or  $r_k = 2000$  meters in the open areas, we can set  $T_{S_k, drive}$  as 0.1 h and 0.4 h respectively, assuming the driving speed is no less than 10 km/h. If the trajectory segment  $(S_k, [t_{k_1}, t_{k_2}])$  satisfies Eqn. (1), we call it a potential driving segment.

Notice that, a person may be very close to the edge of the station range, when the person walks outside the range, the station trajectory may be classified to the driving mode. Hence, we introduce a new feature which requires at least  $\delta_d$  continuous potential driving segments. For example, the segment  $(S_k, [t_{k_1}, t_{k_2}])$  satisfies Eqn. (1), we need to check whether the following  $\delta - 1$  segments also satisfy the equation. The feature could help better classify the person's station trajectory.

In this paper, we only adopt these two features and they could achieve good result in classifying the driving mode. Obviously, more features could achieve better results and it is an important work to add more features during the classification.

2) *Walking Mode*: Similarly, we assume the time period passing through the communication station is relatively larger than the driving mode. Hence, we formulate the basic classifying rule as:

$$T_{S_k,drive} < t_{k_2} - t_{k_1} \leq T_{S_k,walk}, \quad (2)$$

where  $T_{S_k,walk}$  is a parameter that reflects the time threshold of walking through station  $S_k$ . We can assume the slowest walk speed as  $2 \text{ km/h}$  and the threshold can be correspondingly calculated. For example,  $T_{S_k,walk} = 0.5 \text{ h}$  when  $r_k = 500 \text{ m}$  and  $T_{S_k,walk} = 2 \text{ h}$  when  $r_k = 2000 \text{ m}$ . These parameters can be designed by different situations. If the trajectory segment  $(S_k, [t_{k_1}, t_{k_2}])$  satisfies Eqn. (1), we call it a potential walking segment.

Since there might also exist some counterexamples, we also add another feature which requires the following  $\delta_w$  segments should also be classified as potential ones.

3) *Stop Mode*: Suppose a person is in a restaurant or a market, he/she would stay within the communication range for a long time. Hence, we formulate the basic classifying rule as:

$$t_{k_2} - t_{k_1} > T_{S_k,walk}. \quad (3)$$

More features can be extracted and we only use the simple one in classifying the stop mode in this paper. In order to better assist work resumption, we introduce two more modes: living mode and working mode. These two modes share the same classifying rule as the stop mode, and we introduce how to distinguish them as follows.

4) *Living Mode*: Living mode has some special features other than the stop mode. First, the station range is normally small since the many people are communicating through the station. Second, the time interval belongs to a person's rest time, such as in the evening. Third, the living mode segments of a person's trajectory follow a quite similar schedule, since a normal person go back home regularly.

Hence, we can first adopt the feature: the station range  $r_k \leq R_L$  where  $R_L$  is a range parameter such as  $500 \text{ m}$ . Then, we describe the second feature as  $[t_{k_1}, t_{k_2}] \subseteq T_{Rest}$  which implies the time interval belongs to the person's rest time, such as  $T_{Rest} = [21 : 00, 24 : 00] \cup [0 : 00, 8 : 00]$ . The third feature is extracted by checking the patterns in several days. With these additional features, we can classify a trajectory segment from the stop mode to the living mode in a more refined way.

5) *Working Mode*: Similar as the living mode, we can extract more features other than the stop mode. First, the station range is normally small and it can be formulated as  $r_k \leq R_W$  where  $R_W$  is a range parameter such as  $500 \text{ m}$ . Second, the time interval belongs to a person's working time and we formulate it as  $[t_{k_1}, t_{k_2}] \subseteq T_{Work}$ , where  $T_{Work}$  represents a person's working time such as  $T_{Work} = [9 : 00 - 17 : 00]$ . Third, the working mode segments of a person's trajectory also follow a quite similar schedule, and we can extract the feature by checking the segments in several days or weeks.

There might exist many methods to classify the passing mode according to the station trajectory. In this section, we

only present a simple method which works on some predefined detection rules. However, this classification method might make incorrect results. For example, on some crowded roads, walking might be faster than driving; the method would not classify the passing mode correctly. Deep neural networks based methods could improve the accuracy and this would be an important and interesting work in the future.

After classifying a person's station trajectory into different passing modes as drive, walk, stop, live, or work, we can represent the trajectory as  $Q(P_i) = \{(S_k, [t_{k_1}, t_{k_2}], mode), (S_j, [t_{j_1}, t_{j_2}], mode), \dots\}$ . Since the proposed features are not definitely correct, we need to correct some abnormal results of these segments. For example, the segments of the living mode are usually continuous, and we need to modify some incorrect modes; the segments of the driving mode would be continuous for a sufficient time, and it is incorrect when some scattered living or working segments are located within a long driving segments. Hence, with these modification principles, we can correct some abnormal modes and the method could achieve better classification results.

## B. Statistics on Different Passing Modes

In order to compute the epidemic risk of each communication station, we need to find out how many patients have passed through the station and how do they pass through it. Hence, we introduce the method to computing the number of different passing patients.

Since the station's trajectory has temporal and spatial characteristics, we need to reduce the problem from different dimensions. After dividing the time into slots of equal length  $\Delta$ , we can find out the statistical features in each time segment. Denote  $R(S_k, t_n)$  as the station's risk in time segment  $t_n = [n\Delta, (n+1)\Delta)$ , we need to compute the number of patients who are in different passing modes for each  $t_n$ . Denote  $N_{drive}(S_k, t_n)$  as the number of patients who are in the driving mode when passing  $S_k$  during  $t_n$ . Similarly, denote  $N_{walk}(S_k, t_n)$ ,  $N_{stop}(S_k, t_n)$ ,  $N_{live}(S_k, t_n)$ ,  $N_{work}(S_k, t_n)$  as the number of patients who are in the walking, stop, living, working modes correspondingly.

To account the number, we first convert a person's station trajectory to a detailed time-segment representation. For each trajectory segment  $(S_k, [t_{k_1}, t_{k_2}], mode)$  where the passing mode is one of the five introduced modes, find out  $0 \leq i \leq j$  such that  $i\Delta \leq t_{k_1} < (i+1)\Delta$  and  $j\Delta \leq t_{k_2} < (j+1)\Delta$ , we convert the segment to time-segment representation as:

$$\{(S_k, t_i, mode), (S_k, t_{i+1}, mode), \dots, (S_k, t_j, mode)\}.$$

Then, the number of patients in the related time segments (from  $t_i$  to  $t_{j-1}$ ) can be updated. For example, if the mode is drive,  $N_{drive}(S_k, t_i), N_{drive}(S_k, t_{i+1}), \dots, N_{drive}(S_k, t_j)$  could be updated by adding this instance.

Supposing time length  $\Delta = 10$  and the trajectory segment is  $(S_k, [22, 96], drive)$  as depicted in Fig. 3, we convert the segment to  $(S_k, t_2, drive), (S_k, t_3, drive), \dots, (S_k, t_9, drive)$  as the upper part of the figure. Then,  $N_{drive}(S_k, t_i) = N_{drive}(S_k, t_i) +$

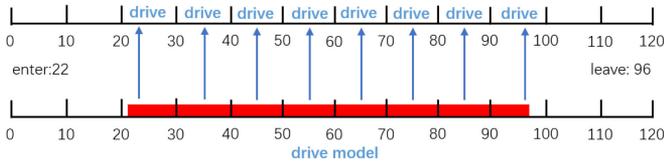


Fig. 3. An illustration example of converting a trajectory segment to a detailed time-segment representation.

$1, \forall i \in [2, 9]$ . This step could convert non-aligned time segments to aligned ones for computation.

After converting all patients' trajectories, we can compute the number of patients passing through these communication stations in each time segment.

### C. The Computation of the Communication Station Risk

As stated above, we assign different risk weights to different passing modes. Denote  $\mu_{drive}, \mu_{walk}, \mu_{stop}, \mu_{live}, \mu_{work}$  as the risk weights of the driving mode, walking mode, stop mode, living mode and working mode respectively. According to the epidemic features of these modes, the living mode has the highest risk since people living together or in the same building would be infected easily as people rarely take precautions at home. After that, the working mode also has large risk since many people are working together and only a few of them may take precautions. The risk of the stop mode is larger than the walking and driving mode, since people staying in a restaurant, market, or a hospital has larger probability being infected. The walking mode has small risk since most people would take precautions such as wearing masks, while the driving mode has the smallest risk value since patients in a car affect little to the station. Hence, these risk weights satisfy the following inequality:

$$\mu_{live} > \mu_{work} > \mu_{stop} > \mu_{walk} > \mu_{drive}.$$

Then, we can compute the epidemic risk of the communication station by combining the risks of each passing modes. We formulate the computation process of  $R(S_k, t_n)$  as the following equation:

$$R(S_k, t_n) = \sum_{mode} N(S_k, t_n, mode) \cdot \mu_{mode}, \quad (4)$$

where  $mode \in \{drive, walk, stop, live, work\}$ .

Putting it together, we describe the method of computing the communication station risk as Alg. 1. The first two lines initialize the related parameters. The passing mode classification is applied as Lines 4-7 and abnormal modes are corrected on Line 8. Then the statistics on different passing modes are computed as Lines 10-15, and we finally generate the epidemic risk of each communication station as Lines 16-18. After the assessment, the risks are normalized to  $[0,1]$  to represent the communication station's risk value.

## V. PERSONNEL RISK ASSESSMENT

In this section, we introduce the method of calculating personnel risk on the basis of the communication station epidemic.

### Algorithm 1: Communication Station Risk Assessment.

- 1: List the station trajectories  $\{Q(P_1), Q(P_2), \dots, Q(P_K)\}$  of the patients;
- 2: Initialize  $N(S_k, t_n, mode) = 0$  for each station  $S_k$ , each time  $t_n \geq 0$ , and each passing mode;
- 3: **for** each trajectory  $Q(P_i)$  **do**
- 4:   **for** each segment  $(S_k, [t_{k_1}, t_{k_2}])$  in  $Q(P_i)$  **do**
- 5:     Classify the segment to a specific passing mode  $mode \in \{drive, walk, stop, live, work\}$ ;
- 6:     Update the segment as  $(S_k, [t_{k_1}, t_{k_2}], mode)$ ;
- 7:   **end for**
- 8:   Modify the passing mode of each segment by correcting abnormal modes;
- 9: **end for**
- 10: **for** each trajectory  $Q(P_i)$  **do**
- 11:   **for** each segment  $(S_k, [t_{k_1}, t_{k_2}], mode)$  of  $Q(P_i)$  **do**
- 12:     Find out  $0 \leq i \leq j$  such that  $i\Delta \leq t_{k_1} < (i+1)\Delta$  and  $j\Delta \leq t_{k_2} < (j+1)\Delta$ ;
- 13:      $N_{mode}(S_k, t_g) = N_{mode}(S_k, t_g) + 1, \forall i \leq g \leq j$ ;
- 14:   **end for**
- 15: **end for**
- 16: **for** each station  $S_k$  and each time segment  $t_n$  **do**
- 17:    $R(S_k, t_n) = \sum_{mode} N(S_k, t_n, mode) \cdot \mu_{mode}$ .
- 18: **end for**

For any person  $U_i$  who is to be enquired, for example a person is going to enter a building for work, denote the trajectory as  $Q(U_i) = \{(S_k, [t_{k_1}, t_{k_2}]), (S_j, [t_{j_1}, t_{j_2}]), \dots, (S_i, [t_{i_1}, t_{i_2}])\}$  where  $t_{i_2}$  is the current time  $t_{cur}$  and  $t_{k_1}$  is the start time of the querying (such as 14 days before the current time), the goal is to calculate his personnel risk  $R(U_i, t_n)$  in each time segment  $t_n$ . Commonly speaking, the officer would check the current personnel risk to decide whether the person can return to work.

Before invoking the communication stations' risk, we first introduce a companion judgement model which decides whether the queried person has been with any patient together in Section V-A. Then, we propose the personnel risk calculation method according to communication stations the person passed through in Section V-B.

### A. Companion Judgement Method

Traditional companion judgement can be fulfilled by checking whether the person has travelled with a infected person simultaneously. For example, the flight information contains each person's detailed data (such as the date of the flight, the seat number of the person, etc). Such travel information could be utilized to judge whether they are in the same plane or in the same train. However, when two persons drive or walk together, such travel information is useless in the judgement. So we introduce a companion judgement method based on the communication stations' information.

Considering the station trajectory of the queried person  $Q(U_i) = \{(S_k, [t_{k_1}, t_{k_2}]), (S_j, [t_{j_1}, t_{j_2}]), \dots, (S_i, [t_{i_1}, t_{cur}])\}$ , we first adopt the passing mode classification method in Section IV-A and each trajectory segment  $(S_k, [t_{k_1}, t_{k_2}])$  can be classified to a specific passing mode.

---

**Algorithm 2:** Companion Judgement Method.

---

- 1: Denote person  $U_i$ 's station trajectory as  $Q(U_i) = \{(S_k, [t_{k_1}, t_{k_2}]), (S_j, [t_{j_1}, t_{j_2}]), \dots (S_i, [t_{i_1}, t_{cur}])\}$ ;
  - 2: Invoke the passing mode classification method to each segment as  $(S_k, [t_{k_1}, t_{k_2}], mode)$ ;
  - 3: **for** each infected person's trajectory  $Q(P_j)$  **do**
  - 4: Find segments  $CS = \{(S_k, [t_1, t_2], mode), \dots\}$  such that they are passing through  $S_k$  during time period  $[t_1, t_2]$  with the same passing mode;
  - 5: **for** each segment  $(S_k, [t_1, t_2], mode) \in CS$  **do**
  - 6: **if** they share the same passing mode in the same station within next  $t_2 + T_{mode}$  time **then**
  - 7: Return the companion infected person  $P_j$ ;
  - 8: **end if**
  - 9: **end for**
  - 10: **end for**
  - 11: Return the judgement that no companion exists.
- 

Since the communication station risk computation method in Section IV has already classified each infected person's trajectory with different passing modes, and these information are stored in memory. Then, we need to judge whether person  $U_i$  has been together with any infected person  $P_j$ . We describe the method as Alg. 2.

The first two lines describe the passing mode classification results of the queried person. For the station trajectory of each infected person  $P_j$ , we find the set of common segments  $CS = \{(S_k, [t_1, t_2], mode), \dots\}$  such that both people are passing through the station  $S_k$  during the time period  $[t_1, t_2]$  simultaneously with the same passing mode (as Line 4). For each common segment  $(S_k, [t_1, t_2], mode) \in CS$ , two people satisfy the companion judgement requirement if they share the passing mode in the next  $t_2 + T_{mode}$  time, where  $T_{mode}$  is defined according to different passing modes. For example, the living mode has larger  $T_{live}$  as persons are likely to stay at home for a long time (such as more than 10 hours), while the driving mode or the walking mode have smaller  $T_{drive}, T_{walk}$ . This judgement is described as Lines 6-8. Finally, if the queried person's trajectory does not match the companion requirement with any infected person, the algorithm outputs no companion is detected.

Notice that, if the queried person once moved together with any infected person, we set the risk as  $R(U_i, t_n) = 1$ ; otherwise we need to calculate the personnel risk based on the communication stations' risk.

**B. Personal Risk Assessment Based on the Communication Station Risk**

If the queried person does not satisfy the companion judgement with any infected person, we need to calculate the person's risk value. For the station trajectory  $Q(U_i) = \{(S_k, [t_{k_1}, t_{k_2}]), (S_j, [t_{j_1}, t_{j_2}]), \dots (S_i, [t_{i_1}, t_{cur}])\}$  and the passing mode of each segment  $(S_k, [t_{k_1}, t_{k_2}], mode)$ , we also convert the trajectory segment to a time-segment representation  $\{(S_k, t_i, mode), (S_k, t_{i+1}, mode), \dots, (S_k, t_j, mode)\}$  where  $i\Delta \leq t_{k_1} < (i+1)\Delta$  and  $j\Delta \leq t_{k_2} < (j+1)\Delta$ .

---

**Algorithm 3:** Personal Risk Assessment.

---

- 1: For a queried person  $U_i$  and the trajectory  $Q(U_i)$ ;
  - 2: Invoke the Companion Judgement Method (Alg. 2) to the queried person, if it detects companion, return the personnel risk as high as 1;
  - 3: **for** each trajectory segment  $(S_k, [t_{k_1}, t_{k_2}], mode)$  **do**
  - 4: Find out  $0 \leq i \leq j$  such that  $i\Delta \leq t_{k_1} < (i+1)\Delta$  and  $j\Delta \leq t_{k_2} < (j+1)\Delta$ ;
  - 5: Compute  $R(U_i, t_n) = R(U_i, t_n) + \gamma_{mode} \cdot R(S_k, t_n)$  where  $i \leq n \leq j$ ;
  - 6: **end for**
  - 7: Set  $T_{period}$  as the risk querying time;
  - 8: Compute  $R(U_i, [t_n - T_{period}, t_n])$  as Eqn. (6).
- 

We design infection weights for each passing modes as  $\gamma_{drive}, \gamma_{walk}, \gamma_{stop}, \gamma_{live}, \gamma_{work}$ , where  $\gamma_{mode}$  ( $mode \in \{drive, walk, stop, live, work\}$ ) implies the probability that a person might be infected by the passing mode. Hence, we can calculate the personnel risk when the person passes through station  $S_k$  in time  $t_n$  as

$$R(U_i, t_n) = \gamma_{mode} \cdot R(S_k, t_n). \quad (5)$$

Normally, it is useless to merely compute the personnel risk in a time segment  $t_n$ . For example, many places require that the queried person should be at low risk for the past 14 days, thus we need to compute the personnel risk in the past  $T_{period}$  time. We formulate the personnel risk calculation method as

$$R(U_i, [t_n - T_{period}, t_n]) = \sum_{t_j \in [t_n - T_{period}, t_n]} r(U_i, t_j). \quad (6)$$

Combining these steps, the whole process the personnel risk assessment is described as Alg. 3.

The first two lines of Alg. 3 invokes the companion judgement method and the personnel risk is set as high as 1 if companion with any infected person is detected. Otherwise, the personnel risk in each time segment  $t_n$  is computed as Lines 3-6. Finally, we set the querying period and compute the personnel risk on Line 8. When the risk  $R(U_i, t_{cur})$  is very small, we can regard the queried person as healthy.

VI. SIMULATIONS

In this section, we generate some station trajectory data for both infected persons and queried persons since the real data might cause privacy leakage. We first introduce the settings in our simulations, including the generation of the communication stations and the method of generating the station trajectory of each person. Then, we show the simulation results of the proposed epidemic risk assessment method. Finally, we introduce the implemented system for querying a person's epidemic risk, which could help assist safe work resumption.

TABLE II  
STATION CHARACTERISTICS OF DIFFERENT SUB-AREAS

ID (sub-areas)	Area type	Radius (km)	# of stations
1	Remote suburb	[1, 3]	300
2	Industrial area	[0.5, 1]	500
3	Forest park area	[0.8, 2]	100
4	Science and technology park	[0.3, 0.5]	500
5	Forest park area	[0.8, 2]	100
6	Urban area	[0.3, 0.5]	300
7	Remote suburb	[1, 3]	200
8	Urban area	[0.3, 0.5]	500
9	Park area	[0.5, 1]	50
10	Urban area	[0.3, 0.5]	500
11	Urban area	[0.3, 0.5]	500
12	Urban area	[0.3, 0.5]	500
13	Remote suburb	[1, 3]	300

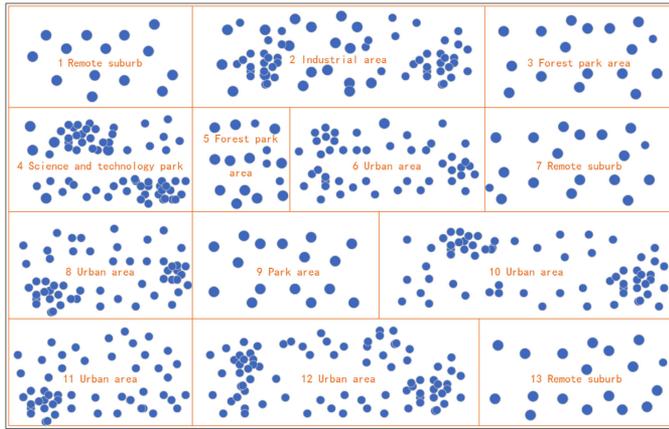


Fig. 4. The area is divided into 13 sub-area of different types.

### A. Simulation Settings

We introduce the settings for evaluating the assessment method. First, we introduce the generation of the communication stations that are deployed in a certain area. Then, we generate the station trajectories of users (including infected persons and the person to be queried).

1) *Generation of the Communication Stations*: Assume all communication stations are deployed in a certain geographic range and we divide the area into 13 sub-areas (districts) of different types for more realistic simulation. As shown in Table II, the second column describes the main type of the area such as remote suburb, industrial area, urban area, etc. The third column describes the radius range of the deployed stations in each sub-area, and the fourth column is the number of the stations. Generally speaking, the station's range in the remote suburb is larger than that in the urban area or industrial area. The generated sub-areas are illustrated in Fig. 4.

We divide the type of communication stations into 8 different categories according to the covered places: residence (Re), supermarket (Sp), restaurant (Rt), workplace (Wp), hotel (Ho), shopping mall (M), park (P), and hospital (Hp). We introduce the characteristics of these types in Table III, where the second column shows the enter time range, the third column lists the normally passing modes, and the fourth column

TABLE III  
CHARACTERISTICS OF DIFFERENT TYPES OF COMMUNICATION STATIONS

Station Type	Enter time range	Passing mode	Time period (h)
residence	00:00:00-23:59:59	{walk, drive, live, stay}	9 – 12
supermarket	9:00:00-22:00:00	{walk, drive, stay}	0.5 – 1
restaurant	11:00:00-21:00:00	{walk, drive, stay}	0.5 – 2
workplace	8:30:00-14:00:00	{walk, drive, work, stay}	5 – 8
hotel	00:00:00-23:59:59	{walk, drive, stay}	9 – 12
mall	9:00:00-23:00:00	{walk, drive, stay}	2 – 4
park	6:00:00-22:00:00	{walk, drive, stay}	0.5 – 1
hospital	00:00:00-23:59:59	{walk, drive, stay}	0.5 – 2

TABLE IV  
THE DISTRIBUTION OF DIFFERENT STATIONS IN THE SUB-AREAS

ID	Station Types	Distribution
1	{Re, Sp, Rt}	{0.6,0.2,0.2}
2	{Re, Sp, Rt, Wp, Ho}	{0.1,0.1,0.2,0.5,0.1}
3	{Sp, P}	{0.2,0.8}
4	{Sp, Re, Wp, Ho}	{0.1,0.2,0.6,0.1}
5	{Sp, P}	{0.2,0.8}
6	{Hp, M, Res, Sp, Re, Ho}	{0.1,0.3,0.2,0.2,0.1,0.1}
7	{Re, Sp, Rt}	{0.6,0.2,0.2}
8	{Hp, M, Res, Sp, Re, Ho}	{0.1,0.3,0.2,0.2,0.1,0.1}
9	{Sp, P}	{0.2,0.8}
10	{Hp, M, Res, Sp, Re, Ho}	{0.1,0.3,0.2,0.2,0.1,0.1}
11	{Hp, M, Res, Sp, Re, Ho}	{0.1,0.3,0.2,0.2,0.1,0.1}
12	{Hp, M, Res, Sp, Re, Ho}	{0.1,0.3,0.2,0.2,0.1,0.1}
13	{Re, Sp, Rt}	{0.6,0.2,0.2}

shows the time period a person may stay within the station. Different sub-areas may contain different types of stations, and we introduce the distribution of these stations in each sub-area as Table IV.

2) *Generation the Station Trajectory of the Users*: We show the method to generate the station trajectory of the users. We first simulate 10 000 users with distinct name and identifiers, and we suppose 2000 persons are infected. Hence, the other 8000 users can be utilized for the querying test.

In order to simulate the trajectories of the patients, we collect some released news from the Centers for Disease Control (CDC) and simulate the trajectories of some infected persons according to the released locations the person passed by, such as in some hospital or stay in some market. The generated trajectories can reflect some characteristics of the infected persons.

To evaluate the risk assessment method, we generate four different types of station trajectories for the queried persons.

- Non-overlapping trajectory implies the queried person has no overlapping trajectory segments with any infected persons;
- Overlapping trajectory implies the queried person has some overlapping trajectory segments with some infected person, and different passing modes are simulated;
- Companion trajectory implies the queried person has once moved together with an infected person;
- Home/work trajectory simulates a periodical pattern that the queried person goes to work or stays at home.

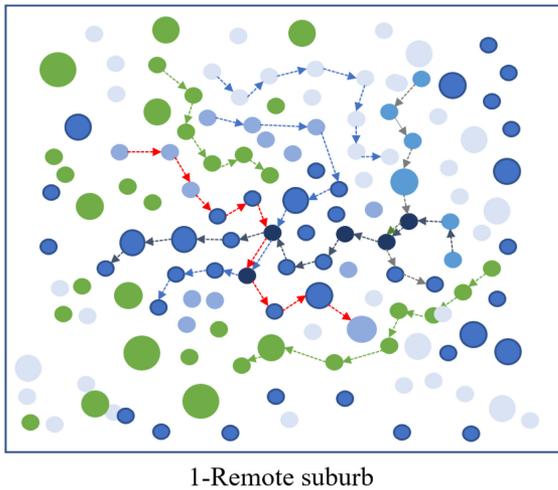


Fig. 5. An example of the communication station risk in the sub-area 1.

Notice that, the enter time range of the home/work trajectory satisfies the characteristics in Table III.

We generate these types of trajectories with distribution  $\{0.5, 0.2, 0.1, 0.2\}$  which means half of the queried persons have non-overlapping with the infected person, while 20% people has overlapping trajectory segments but some of them might be at low risk by the driving or walking passing mode.

Notice that, we can obtain some real information of the communication stations through some public data, such as <https://opencellid.org/>. However, the trajectory of each person is sensitive and such information is unobtained in real experiment. In order to protect person's privacy, we generate the communication stations and people's trajectory to show the performance of the proposed method.

### B. Communication Station Risk Assessment

We implement the assessment method in Section IV and we can compute the epidemic risk values of each station in each time segment. To show some intuitive results, we fix any time segment and the risk values of these communication station is depicted with different colors. We depict an example in Fig. 5 when we choose a specific time segment, and the figure shows the epidemic risk of the communication stations in the 1-th sub-area. Green means the station has no risk while the degree of the blue color means the risk level. The communication station with larger risk value is depicted with darker blue in the figure. Notice that, the blue arrows imply the station trajectories of the infected persons, while the green arrows imply the station trajectory of the queried person. From the figure, with more infected persons passing through, the communication station has larger risk value.

### C. Personnel Risk Assessment

We first show the distribution of the personnel risks of different types of trajectories. As shown in Table V, if a queried person has no overlapping trajectory segments with any infected person, the person is assessed with no risk (4000 such persons have low risk within  $[0, 0.1)$ ). For the 800 companion

TABLE V  
THE PERSONNEL RISK OF THE QUERIED PERSONS

Trajectory Type	Number of Persons	$[0, 0.1)$	$[0.1, 0.5)$	$[0.5, 1]$
Non-overlapping	4000	4000	0	0
Overlapping	1600	832	545	223
Companion	800	0	0	800
Home	800	715	3	82
Work	800	550	190	60

trajectories, the proposed algorithm could detect the companion person and the risk is set as high as 1; hence all these trajectories have high risk value within  $[0.5, 1]$ . Considering the overlapping trajectories, many persons are in the driving or the walking mode, and they are not affected too much. In addition, only a small portion of the queried persons have high risk value for this type of trajectory. Similarly, most of the home/work trajectories have low risk value, and only a few of them have high risk value.

To better illustrate different situations of the queried person, we introduce and evaluate the following special cases.

#### 1) The User Has Been to a Certain Sub-Area, but he/she Has Not Contacted Any Infected Person in the Same Time:

When our algorithm calculates the user's personnel risk, it would find out the stations he/she has passed through. Since no infected persons pass through the station simultaneously, the communication station risk is low when the user passes through the station. Hence, the personnel risk is low and the simulation results also corroborate our analyses.

#### 2) The User Drives Through Some Communication Stations With High Risk, but he/she Has Not Contacted Any Infected Person:

This is a classic overlapping trajectory where the user passes through some communication stations with high risk values, but the personnel risk is low since the driving mode is not affected by the station. By setting a very small  $\gamma_{drive}$  value (Section V-B), the simulation results verify that the user has low epidemic risk.

#### 3) The User Drives in a Sub-Area With an Infected Person:

For example, the user has a good relationship with an infected person and they have driven together in some area. Though the driving mode has low infection weight, the companion judgement method could detect the companion trajectory and assess the user's epidemic risk value as high as 1. The simulation results of the companion trajectories corroborate the analyses.

### D. The Epidemic Risk Assessment System

On the basis of the proposed method, we implemented the epidemic risk assessment system as Fig. 6. The left side of the system shows the distribution of the communication stations in a specific area. The middle part shows the characteristics of the communication stations, including the number of persons who pass through the station and the distribution of the station risk. The right side is utilized to query and search the personnel risk. By inputting the personnel information, the system would output the calculated risk value. This system can help assist safe work resumption by verifying the personnel risk.



Fig. 6. The epidemic risk assessment system.

## VII. CONCLUSION

In this paper, we propose a novel epidemic risk assessment method on the basis of the granular data collected by the communication stations. This method is composed of two stages: computing the communication station risk and calculating the personnel risk. The first stage aims at evaluating the stations' risk values in different time intervals, while the second stage calculates a fine-grained personnel risk value according to the station trajectory of the specific person. In order to compute the communication station risk more accurately, we define five different passing modes which reflect different ways of passing through the station and these modes are assigned with different risk weights. The risk values of the stations are computed according to the number of infected persons and the passing mode of each person. To calculate the personnel risk, we propose a companion judgement method which could automatically decide whether two people have moved together. If no companion trajectory is detected, we then calculate a person's epidemic risk according to the risks of stations he/she has passed through. This two-stage assessment method could better calculate people's epidemic risk, which plays an important role in assisting safe work resumption.

In the future, we would like to extend and apply the proposed method in practice by collecting real data from the telecommunications operators. Furthermore, there exist many important and interesting problems to handle, such as improving the passing mode classification accuracy by more advanced methods.

## REFERENCES

- [1] L. Bai, D. Yang, X. Wang, L. Tong, X. Zhu, and N. Zhong, "Chinese experts' consensus on the Internet of Things-aided diagnosis and treatment of coronavirus disease 2019 (COVID-19)," *Clin. ehealth*, vol. 3, pp. 7–15, 2020.
- [2] J. Brown and X. Du, "Detection of selective forwarding attacks in heterogeneous sensor networks," in *Proc. IEEE Int. Conf. Commun.*, 2008, pp. 1583–1587.
- [3] J. F.-W. Chan *et al.*, "A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster," *The Lancet*, vol. 395, no. 10223, pp. 514–523, Feb. 2020.
- [4] C. Bai and J. Zhao, *Medical Internet of Things*. Beijing: Science Press, 2016.
- [5] K. L. Cooke and P. V. D. Driessche, "Analysis of an SEIRS epidemic model with two delays," *J. Math. Biol.*, vol. 35, pp. 240–260, 1996.
- [6] L. Dong, "5G and intelligence medicine how the next generation of wireless technology will reconstruct healthcare?," *Precis. Clin. Med.*, vol. 2, no. 4, pp. 205–208, Dec. 2019.
- [7] D. Domingo-Fernandez *et al.*, "COVID-19 knowledge graph: A computable, multi-modal, cause-and-effect knowledge model of COVID-19 pathophysiology," *Biorxiv*, 2020, doi: 10.1101/2020.04.14.040667.
- [8] X. Du, "QoS routing based on multi-class nodes for mobile ad hoc networks," *Ad Hoc Netw.*, vol. 2, no. 3, pp. 241–254, 2004.
- [9] X. Du and F. Lin, "Improving sensor network performance by deploying mobile sensors," in *Proc. 24th IEEE Int. Perform., Comput. Commun. Conf.*, 2005, pp. 67–71.
- [10] X. Du and D. Wu, "Adaptive cell-relay routing protocol for mobile ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 278–285, Jan. 2006.
- [11] Q. Gao, Y. Hu, Z. Dai, J. Wu, F. Xiao, and J. Wang, "The epidemiological characteristics of 2019 novel coronavirus diseases (COVID-19) in Jingmen, Hubei, China," *Medicine*, vol. 99, no. 23, 2020, Art. no. e20605.
- [12] L. Gong, X. N. Liu, and X. Zou, "Spread of infectious disease risk assessment based on the spatial-temporal trajectory data analysis," *Acta Geodaetica Et Cartographica Sinica*, 2015, doi: 10.11947/j.AGCS.2015.F062.
- [13] W. Guan *et al.*, "Clinical characteristics of 2019 novel coronavirus infection in China," *Medrxiv*, 2020, doi: 10.1101/2020.02.06.20020974.
- [14] Z. Gu, W. Hu, C. Zhang, H. Lu, L. Yin, and L. Wang, "Gradient shielding: Towards understanding vulnerability of deep neural networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 921–932, Apr.–Jun. 2021.
- [15] Z. Gu, Y. Wang, T. Shen, and F. C. M. Lau, "On heterogeneous sensing capability for distributed rendezvous in cognitive radio networks," *IEEE Trans. Mobile Comput.*, vol. 20, no. 11, pp. 3211–3226, Nov. 2021.
- [16] Z. Gu, T. Shen, Y. Wang, and F. C. M. Lau, "Efficient rendezvous for heterogeneous interference in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 91–105, Jan. 2020.
- [17] L. Hufnagel, D. Brockmann, and T. Geisel, "Forecast and control of epidemics in a globalized world," *Proc. Nat. Acad. Sci. USA*, 2004. [Online]. Available: <https://doi.org/10.1073/pnas.0308344101>
- [18] C. Huang *et al.*, "Clinical features of patients with 2019 novel coronavirus in Wuhan, China," *The Lancet*, vol. 395, no. 10223, pp. 497–506, Feb. 2020.
- [19] C. Hou *et al.*, "The effectiveness of quarantine of Wuhan city against the corona virus disease 2019 (COVID-19): A well-mixed SEIR model analysis," *J. Med. Virol.*, vol. 92, no. 7, pp. 841–848, Jul. 2020.
- [20] L. Jun-feng, "Assessment and prediction of COVID-19 based on SEIR model with undiscovered people," *J. Univ. Electron. Sci. Technol. China* vol. 49, no. 3, pp. 375–382, 2020, doi: 10.12178/1001-0548.2020083.
- [21] M. Khalili, M. Karamouzian, N. Nasiri, S. Javadi, A. Mirzazadeh, and H. Sharifi, "Epidemiological characteristics of COVID-19: A systematic review and meta-analysis," *Epidemiol. Infect.*, vol. 148, 2020, Art. no. E130.
- [22] B. P. L. Lau *et al.*, "A survey of data fusion in smart city applications," *Inf. Fusion*, vol. 52, pp. 357–374, 2019.
- [23] K. Li *et al.*, "An experimental study for tracking crowd in smart cities," *IEEE System J.*, vol. 13, no. 3, pp. 2966–2977, Sep. 2019.
- [24] W. Li, R. J. Piechocki, K. Woodbridge, C. Tang, and K. Chetty, "Passive WiFi radar for human sensing using a stand-alone access point," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 1986–1998, Mar. 2021.
- [25] R. Lu *et al.*, "Genomic characterization and epidemiology of 2019 novel coronavirus: Implications of virus origins and receptor binding," *The Lancet*, vol. 395, no. 10224, pp. 565–574, Feb. 2020.
- [26] C. E. Lopez, M. Vasu, and C. Gallemore, "Understanding the perception of COVID-19 policies by mining a multilanguage twitter dataset," 2020, *arXiv:2003.10359*.
- [27] M. Lenzen *et al.*, "Global socio-economic losses and environmental gains from the coronavirus pandemic," *PLoS One*, vol. 15, no. 7, 2020, Art. no. e0235654 <https://doi.org/10.1371/journal.pone.0235654>.
- [28] B. M. Liu, Q. Q. Yang, L. Y. Zhao, W. Xie, and Y. Si, "Epidemiological characteristics of COVID-19 patients in convalescence period," *Epidemiol. Infect.*, vol. 148, no. e108, pp. 1–9, 2020.
- [29] Q. Li *et al.*, "Early transmission dynamics in wuhan, china, of novel coronavirus-infected pneumonia," *New Engl. J. Med.*, vol. 382, pp. 1199–1207, 2020.
- [30] C. Moore and M. E. J. Newman, "Epidemics and percolation in small-world networks," *Phys. Rev. E*, vol. 61, 2000, Art. no. 5678.

[31] M. F. Myers, D. J. Rogers, J. Cox, A. Flahault, and S. I. Hay, "Forecasting disease risk for increased epidemic preparedness in public health," *Adv. Parasitol.*, vol. 47, pp. 309–330, 2000.

[32] B. Meng, J. F. Wang, J. Y. Liu, and J. L. Wu, "Understanding the spatial diffusion process of SARS in Beijing," *Public Health*, vol. 119, no. 12, pp. 1080–1087, Dec. 2005.

[33] N. Peeri *et al.*, "The SARS, MERS and novel coronavirus(COVID-19) epidemics, the newest and biggest global health threats: What lessons have we learned?," *Int. J. Epidemiol.*, vol. 49, no. 3, pp. 717–726, Jun. 2020.

[34] A. Pluchino *et al.*, "A novel methodology for epidemic risk assessment: The case of COVID-19 outbreak in Italy," 2020, *arXiv: 2004.02739*.

[35] L.T. *et al.*, "Importation and human-to-human transmission of a novel coronavirus in Vietnam," *New Engl. J. Med.*, vol. 382, pp. 872–874, 2020.

[36] C. Rothe *et al.*, "Transmission of 2019-nCoV infection from an asymptomatic contact in Germany," *N. Engl. J. Med.*, vol. 382, pp. 970–971, 2020.

[37] K. Roosa *et al.*, "Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020," *Infect. Dis. Modell.*, vol. 5, pp. 256–263, 2020.

[38] Y. Song, J. Jiang, X. Wang, D. Yang, and C. Bai, "Prospect and application of Internet of Things technology for prevention of SARIs," *Clin. Ehealth*, vol. 3, pp. 1–4, 2020.

[39] P. Sun *et al.*, "Clinical characteristics of hospitalized patients with SARS-CoV-2 infection: A single arm meta-analysis," *J. Med. Virol* vol. 92, no. 6, pp. 612–617, 2020, doi: 10.1002/jmv.25735.

[40] Y. Song, J. Jiang, X. Wang, D. Yang, and C. Bai, "Prospect and application of Internet of Things technology for prevention of SARIs," *Clin. ehealth*, vol. 3, pp. 1–4, 2020. [Online]. Available: <https://doi.org/10.1016/j.ceh.2020.02.001>

[41] M. C. Thomson *et al.*, "Malaria early warnings based on seasonal climate forecasts from multi-model ensembles," *Nature*, vol. 439, pp. 576–579, 2006.

[42] Z. Tian, C. Luo, J. Qiu, X. Du, and M. Guizani, "A Distributed deep learning system for web attack detection on edge devices," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 1963–1971, Mar. 2020.

[43] Z. Tian *et al.*, "Real time lateral movement detection based on evidence reasoning network for edge computing environment," *IEEE Trans. Ind. Inform.*, vol. 15, no. 7, pp. 4285–4294, Jul. 2019.

[44] Z. Tian, X. Gao, S. Su and J. Qiu, "Vcash: A novel reputation framework for identifying denial of traffic service in internet of connected vehicles," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 3901–3909, May 2020.

[45] D. J. Watts, R. Muhamad, D. C. Medina, and P. S. Dodds, "Multiscale, resurgent epidemics in a hierarchical metapopulation model," *Proc. Nat. Acad. USA*, 2005, doi: 10.1073/pnas.0501226102.

[46] Y. Wang, Y. Liu, J. Struthers, and M. Lian, "Spatiotemporal characteristics of COVID-19 epidemic in the United States," *Clin. Infect. Dis.*, vol. 72, no. 4, pp. 643–651, Feb. 2021.

[47] J.T. K. WuLeung, and GM. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modeling study," *The Lancet*, vol. 395, no. 10225, pp. 689–697, Feb./Mar. 2020.

[48] Y. Zhou, B. P. L. Lau, C. Yuen, B. Tuncer, and E. Wilhelm, "Understand urban human mobility through crowdsensed data," *IEEE Commun. Mag.*, vol. 56, no. 11, pp. 52–59, Nov. 2018.

[49] Y. Zhou, B. P. L. Lau, Z. Koh, C. Yuen, and B. K. K. Ng, "Understanding crowd behaviors in a social event by passive WiFi sensing and data mining," *IEEE Internet Things J.*, vol. 7, no. 5, pp. 4442–4454, May 2020.

[50] N. Zhu *et al.*, "A novel coronavirus from patients with pneumonia in China, 2019," *New Engl J. Med.*, doi: 10.1056/NEJMoa2001017.



**Le Wang** received the Ph.D. degree in computer science from the National University of Defense Technology, Changsha, China. He is currently an Associate Professor with the Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou, China. His current research interests include network and big data security. He was a Member of the China Computer Federation.



**Xiaolong Chen** received the bachelor's degree in computer science from Sichuan Agricultural University, Ya'an, China, in 2018. He is currently working toward the master's degree in computer science and technology from Guangzhou University, Guangzhou, China. His research interests include knowledge graph and data mining.



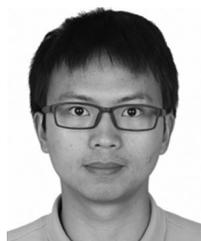
**Yunyi Tang** received the bachelor's degree in information security from Guizhou University, Guiyang, China, in 2018. He is currently working toward the master's degree in computer science and technology from Guangzhou University, Guangzhou, China. His research interests include knowledge graph and natural language processing.



**Xingang Wang** received the bachelor's degree in computer science and technology from the Taiyuan University of Technology, Taiyuan, China, in 2016. He is currently working toward the master's degree in computer science and technology from Guangzhou University, Guangzhou, China. His research interests include deep learning, knowledge graph, and malware detection.



**Xiaojiang Du** (Fellow, IEEE) received the B.S. and M.S. degrees in electrical engineering (Automation Department) from Tsinghua University, Beijing, China in 1996 and 1998, respectively. He received the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, MD, USA, in 2002 and 2003, respectively. He is a tenured Full Professor and the Director of the Security And Networking (SAN) Lab in the Department of Computer and Information Sciences at Temple University, Philadelphia, PA, USA. His research interests are security, wireless networks, and systems. He has authored more than 500 journal and conference papers in these areas, as well as a book published by Springer. He won the best paper award at IEEE ICC 2020, IEEE GLOBECOM 2014 and the best poster runner-up award at the ACM MobiHoc 2014. He serves on the editorial boards of three IEEE journals. He is a Life Member of ACM.



**Zhaoquan Gu** received the bachelor's and Ph.D. degrees in computer science from Tsinghua University, Beijing, China, in 2011 and 2015, respectively. He is currently a Professor with the Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou, China. His research interests include wireless networks, distributed computing, big data analysis, and artificial intelligence security.



**Mohsen Guizani** (Fellow, IEEE) received the B.S. (with distinction) and M.S. degrees in electrical engineering, the M.S. and Ph.D. degrees in computer engineering from Syracuse University, Syracuse, NY, USA, in 1984, 1986, 1987, and 1990, respectively. He is currently a Professor with the Computer Science and Engineering Department, Qatar University, Qatar. His research interests include wireless communications and mobile computing, computer networks, mobile cloud computing, security, and smart grid. Throughout his career, he received three

teaching awards and four research awards. He also received the 2017 IEEE Communications Society WTC Recognition Award as well as the 2018 AdHoc Technical Committee Recognition Award for his contribution to outstanding research in wireless communications and Ad-Hoc Sensor networks. He was the Chair of the IEEE Communications Society Wireless Technical Committee and the Chair of the TAOS Technical Committee. He served as the IEEE Computer Society Distinguished Speaker and is currently the IEEE ComSoc Distinguished Lecturer. He is a Senior Member of ACM.



**Zhihong Tian** (Member, IEEE) is currently a Professor, and Dean, with the Cyberspace Institute of Advanced Technology, Guangzhou University, Guangdong Province, China. Guangdong Province Universities and Colleges Pearl River Scholar (Distinguished Professor). He is also a part-time Professor at Carlton University, Ottawa, ON, Canada. Previously, he served in different academic and administrative positions at the Harbin Institute of Technology. He has authored more than 200 journal and conference papers in these areas. His research interests include computer networks and

cyberspace security. His research has been supported in part by the National Natural Science Foundation of China, National Key research and Development Plan of China, National High-tech R&D Program of China (863 Program), and National Basic Research Program of China (973 Program). He also served as a member, Chair, and General Chair of a number of international conferences. He is a Senior Member of the China Computer Federation.