Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Assessing personal exposure to COVID-19 transmission in public indoor spaces based on fine-grained trajectory data: A simulation study

Pengfei Chen a,b, Dongchu Zhang a, Jianxiao Liu c,d, Izzy Yi Jian d

a School of Geospatial Engineering and Science, Sun Yat-Sen University, Guangzhou, 510275, Guangdong, China
b The Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, 519082, Guangdong, China
c Department of Real Estate and Construction, Faculty of Architecture, The University of Hong Kong, 999077, Hong Kong, China
d School of Design, The Hong Kong Polytechnic University, 999077, Hong Kong, China

1. Introduction

The coronavirus disease 2019 (COVID-19) pandemic has posed substantial challenges to worldwide health systems in quick response to epidemics. The assessment of personal exposure to COVID-19 in enclosed spaces is critical to identifying potential infectees and preventing outbreaks. However, traditional contact tracing methods rely heavily on a manual interview, which is costly and time consuming given the large population involved. With advanced indoor localisation techniques, it is possible to collect people’s footprints accurately by locating their smartphones. This study presents a new framework for the assessment of personal exposure to COVID-19 carriers using their fine-grained trajectory data. An integral model was established to quantify the exposure risk, in which the spatial and temporal decay effects are simultaneously considered when modelling the airborne transmission of COVID-19. Regarding the obstacle effect of the indoor layout on airborne transmission, a weight graph based on the space syntax technique was further introduced to constrain the transmission strength between subspaces that are less inter-visible. The proposed framework was demonstrated by a simulation study, in which external comparison and internal analysis were conducted to justify its validity and robustness in different scenarios. Our method is expected to promote the efficient identification of potential infectees and provide an extensible spatial–temporal model to simulate different control measures and examine their effectiveness in a built environment.

ARTICLE INFO

Keywords:
Contact tracing
Space syntax
COVID-19
Visibility analysis
Spatio-temporal model

ABSTRACT

The coronavirus disease 2019 (COVID-19) pandemic has posed substantial challenges to worldwide health systems in quick response to epidemics. The assessment of personal exposure to COVID-19 in enclosed spaces is critical to identifying potential infectees and preventing outbreaks. However, traditional contact tracing methods rely heavily on a manual interview, which is costly and time consuming given the large population involved. With advanced indoor localisation techniques, it is possible to collect people’s footprints accurately by locating their smartphones. This study presents a new framework for the assessment of personal exposure to COVID-19 carriers using their fine-grained trajectory data. An integral model was established to quantify the exposure risk, in which the spatial and temporal decay effects are simultaneously considered when modelling the airborne transmission of COVID-19. Regarding the obstacle effect of the indoor layout on airborne transmission, a weight graph based on the space syntax technique was further introduced to constrain the transmission strength between subspaces that are less inter-visible. The proposed framework was demonstrated by a simulation study, in which external comparison and internal analysis were conducted to justify its validity and robustness in different scenarios. Our method is expected to promote the efficient identification of potential infectees and provide an extensible spatial–temporal model to simulate different control measures and examine their effectiveness in a built environment.
Table 1

| Technology | Indoor accuracy | Advantages | Disadvantages |
|------------|----------------|------------|---------------|
| QR code    | Room level     | Large coverage | Low spatial granularity |
|            |                | Easy deployment | No actual exposure is captured |
|            |                | Low cost       | Large positioning error |
|            |                | Actual contact can be identified | Signal loss in indoor or underground spaces |
|            |                | Movement information is available | Specific App is required |
|            |                | Passive collection | Sensitive to spatial blockage |
|            |                | Actual contact can be identified | No movement information |
|            |                | Preserved privacy | Specific App is required |
| WiFi       | <2 m           | Passive collection | Fails without a wireless environment |
|            |                | Actual contact can be identified | WiFi access needs to be turned on |
|            |                | High positioning accuracy for indoor environment | |
|            |                | Movement information is available | |

2. Related works

2.1. Overview of digital contact tracing approaches

During the fight against COVID-19, a series of smartphone-based digital approaches has been implemented to facilitate contact tracing by recording the user’s close exposure to a person who was later found to be infected. These approaches generally rely on QR codes, Global Positioning System (GPS), Bluetooth or WiFi technology [25]. These technologies are compared in Table 1.

2.1.1. QR codes

QR codes are easily deployed and provide a cheap way to collect public mobility information for contact tracing. In China, people are instructed to scan a QR code when visiting public spaces. An alert will be pushed to smartphone users when they are found to be once co-located with an infected person [11]. Although this approach lacks focus on whether a close contact has actually occurred, it is a quick and effective response to the spread of COVID-19 given the vast population of China [26].

2.1.2. GPS

GPS can provide detailed footprints of users and thus support a finer-grained contact tracing than QR codes. Various smartphone GPS-based apps, such as ‘Care19’ from the US, ‘Corona Data’ from Germany or ‘Corona 100 m’ from South Korea, have been established; they allow the quick identification of possible locations of risky contact [10,27,28]. Wang et al. [29] demonstrated a typical example that presented a GPS-based system to reconstruct the spatiotemporal trajectory of a smartphone user in the past 14 days and evaluated the infection risk based on the user’s historic visits to places that contained diagnosed patients. However, GPS data suffer from a profound signal loss caused by dense buildings or indoor environments [30]. In addition, given the inevitable positioning errors in GPS data, setting proper proximity thresholds for distinguishing risky contact is challenging [29].

2.1.3. Bluetooth

Another batch of studies attempted to use Bluetooth for contact tracing. Unlike GPS, which records the coordinates of users, the received signal strength (RSS) of a Bluetooth beacon is utilised to measure the distance between nearby devices. Although Bluetooth technology provides less information on user movement and actual geographic location, which is not conducive to studying the disease spread, user privacy can be largely preserved [31]. In addition, Bluetooth is not restricted by indoor or outdoor conditions. Given these advantages, Bluetooth technology has been widely implemented in contact-tracing apps, such as the ‘NHS COVID-19’ app from the UK and ‘COVIDSafe’ from Australia [32,33]. Although Bluetooth technology can generally achieve a 2 m accuracy in distance estimation [34], the measurement can be less reliable once the smartphone is deep in a pocket or handbag [25]. Furthermore, the technology cannot determine whether people are separated by walls or floors, which will induce a false level of spurious identification in contact tracing [25,35].

2.1.4. WiFi

WiFi can provide positioning data with higher accuracy in indoor spaces compared with GPS. This merit makes WiFi an effective technology for contact-tracing in indoor environments, such as hospitals, airports and supermarkets [35]. Once users enable the WiFi access on their devices, their locations can be timely computed based on the
strength of signals between their devices and WiFi access points (WAPs). Similar to GPS-based contact tracing approach, social distance and duration thresholds are commonly used in WiFi-based approaches to identify close exposure [36]. Section 2.3 will introduce more details about the positioning performance of WiFi technology.

2.2. Spatial model of contact risk

The current methods for contact tracing are mostly qualitative and output only Boolean values to distinguish risky contacts based on predefined proximity thresholds [10,27,28]. However, given the distance-decay effect, contacts that satisfy given thresholds may present varying risks. Yin et al. [37] assumed the contact strength is inverse-distance weighted and introduced a distance-decay model to improve the calculation of encounter probability based on the spatio-temporal trajectories of agents. Nevertheless, only spatial distance was considered in this study, whilst the effect from temporal distance was overlooked.

With awareness of the elevated transmission risk during prolonged exposure, Ying and O’Clery [38] measured the risk of viral transmission by the duration when a customer was exposed to an infected person in the same zone of a supermarket. This kind of method, which substantially estimates the contact strength by the frequency of co-location, is common in current studies on COVID-19, especially in large-scale surveillances [39–41]. However, such co-location-based methods have been argued as being coarsely grained without consideration on the precise distance between individuals [38,42]. Moreover, given the airborne transmission property of SARS-CoV-2, even without direct contact with infected people, virus-laden aerosols remain risky to those who walk along the patient’s path [17]. The above issues must be further considered when building a sophisticated risk model.

2.3. Indoor localisation technology

Advanced indoor localisation approaches offer new chances to develop better contact tracing methods. WiFi and active radio frequency identification (RFID) are essential technologies used in indoor real-time positioning [43]. Technically, WiFi and RFID share similar positioning principles, in which the location of a target device is estimated based on the geometric relationship between multiple RFID readers or WAPs with known positions. However, given the limited coverage of the RFID reader, RFID is often used only when room-level accuracy is required [44], such as the tracking of goods, whilst WiFi is more accurate and easily deployed to track people in a large and complex scenario [22,43].

WiFi-based localisation technology generally utilises the RSS to estimate the distance between a smartphone and a WAP: the higher the RSS value, the smaller the distance [24]. By combining the estimated distances to at least three reference points, the real-time location of a smartphone in a predefined coordinate system can be possibly obtain based on trilateration or the least square method [24]. The positioning accuracy of WiFi technology ranges within several meters in an obstructionless space and can be significantly improved to the centimetre level using ultra-wideband techniques [22,45].

To deal with highly obstructed indoor environments, Chen et al. [23] presented a frequency-hopping approach and achieved an accuracy of 5 cm with a standard bandwidth. In the recent study of Zhang et al. [46], an adaptive access point selection method, which can achieve a positioning accuracy within 2 cm, was proposed to deal with the positioning problem of smartphones with a low scanning frequency. Although the accuracy of WiFi-based localisation is affected by the environments and hardware, advanced studies have proven that WiFi technology can be readily used to collect people’s fine-grained trajectory.

3. Methodology

3.1. Spatiotemporal decay in the airborne transmission of SARS-CoV-2

According to the study of Chu et al. [16], the absolute risk of SARS-CoV-2 infection is generally estimated at 12.8% with a physical distance (PD) of less than 1 m, and for every additional 1 m, the risk decreases approximately by half (change factor: 2.02). Hence, regardless of additional protection (e.g. face mask and goggles), the infection risk concerning spatial distance can be modelled as follows:

$$C_{\text{risk}}(D, P) = 12.8\% \times 2.02^{(-D/P)}$$

where $P$ and $D$ are the spatial coordinate vector of the exposed and infected persons, respectively. Notably, although the distance decay effect in Equation (1) is widely acknowledged in the scientific community [47,48], the coefficients 12.8% and 2.02 can be arguable because they are the result of regression from multiple clinical studies without consideration of the duration of exposure, size of droplets, and other factors, such as wearing of masks and face shields [16,49,50]. However, we kept these coefficients following the original paper because they have negligible impact on the assessment of the relative risk of transmission and can be easily amended in different scenarios.

In addition, previous studies revealed that the infectiousness of virus-laden aerosols decays along with time, whilst the decay rate can be affected by numerous environmental factors. According to the study of Dubisch et al. [51], temperature, humidity and simulated sunlight have a significant influence on the persistence of SARS-CoV-2 in aerosols, and the decay rate can be calculated as follows:

$$k_{\text{decay}}(T, RH, S) = 0.16030 + 0.04018 \left(\frac{T - 20.615}{10.585}\right) + 0.02176 \left(\frac{RH - 45.235}{28.665}\right) + 0.14369 \left(\frac{S - 0.95}{0.95}\right) + 0.02636 \left(\frac{T - 20.615}{10.585}\right) \left(\frac{S - 0.95}{0.95}\right).$$

$$\frac{d}{dt} S(t) = k_{\text{decay}}(T, RH, S) \cdot S(t) - C_{\text{risk}}(D, P) \cdot S(t) \cdot e^{-\gamma t}$$

where $S(t)$ is the concentration of SARS-CoV-2 in the aerosol at time $t$, $k_{\text{decay}}$ is the decay rate, $C_{\text{risk}}$ is the risk of infection, and $\gamma$ is the decay coefficient. The initial condition is $S(0) = S_0$, and the final condition is $S(T) = 0$. The solution of this equation is given by

$$S(t) = S_0 \cdot e^{-\gamma t} \cdot \left(1 - C_{\text{risk}}(D, P) \cdot \frac{1}{\gamma} \cdot e^{-\gamma t}\right).$$

Fig. 1. Geometric interpretation of exposure risk calculation of the simplest scenario based on integration. Infected and exposed people are motionless and $t_{\text{start}, e} > t_{\text{end}, i}$.
where \( k_{\text{infectivity}} \) is the decay rate of infectivity (min\(^{-1}\)), \( T \) denotes the temperature (°C), \( RH \) refers to the relative humidity (%), and \( S \) is the UVB irradiance (W/m\(^2\)).

Let \( t_o \) denote the starting time of aerosol formation. \( r(p_i, p_j, t) \) is the instantaneous risk at a given timestamp \( t \) (\( t > t_o \)), which can be calculated as follows:

\[
r(p_i, p_j, t) = C_{\text{optimal}}(p_i, p_j) \times k_{\text{infectivity}}(T, RH, S) \times (t - t_o). \tag{3}
\]

### 3.2. Integral model for the assessment of exposure risk

The exposure risk can be calculated based on Equation (3) by means of integration. The simplest scenario is shown in Fig. 1, where the two agents are motionless, and the infected person’s movement event ends before the exposed person’s starts. Consequently, the exposed person will suffer from all the virus-laden aerosols formed during the movement of the infected person. In that sense, the individual risk \( R_e \) for the exposed person can be expressed as follows:

\[
R_e = \int_{t_o}^{t_{i,1}} \int_{t_o}^{t_{j,1}} r(p_i, p_j, x) \, dx \, dy. \tag{4}
\]

However, given that locations \( p_i \) and \( p_j \) are not always static, and the moving event may start and end in different orders, using Equation (4) will lead to logic contradictions (e.g. exposure risk exists only after the action of infected people). Therefore, Equation (4) needs to be improved further to fit different cases.

Considering that the sampling interval of the WiFi-based localisation technique can be at the secondary level [24], this study assumed that the agent moves uniformly and straight between each two trace points. Thus, an agent’s trajectory can be decomposed into multiple linear segments (i.e. moving events), and the agent’s location can be possibly inferred at any timestamp. Without loss of generality, we used \( M_i \) to denote the movement between two successive sampling timestamps, namely, \( t_i \) and \( t_{i+1} \), and the agent’s location at timestamp \( t \) (\( t_i \leq t < t_{i+1} \)) can be expressed as \( M_i(t) \).

Notably, the risk exists only after the formation of aerosols. Therefore, given the movements \( M_i \) and \( M_j \) of exposed and infected people, respectively, six possible cases can be obtained when calculating the exposure risks by means of integration (Fig. 2).

**Case 1.** When \( t_j > t_{i+1} \), the exposed person gains no risk from this pair of movements given that his movement ends before the virus-laden aerosols form.

**Case 2.** When \( t_i > t_{j+1} \), the exposed person suffers from the full aerosols formed by the infected person along all movements.

**Case 3.** When \( t_i \leq t_i \leq t_{j+1} \), a given timestamp \( t_i \leq t < t_{i+1} \), the exposed person suffers only from the aerosols formed during \( t_i \) to \( t \).

**Case 4.** When \( t_i \leq t_j \leq t_{i+1} \), such case can be treated as a combination of Cases 1 and 3 by taking \( t_j \) as the cutting line of \( M_i \).

**Case 5.** When \( t_i \leq t_j \leq t_{i+1} \), such case can be treated as a combination of Cases 2 and 3 by taking \( t_{j+1} \) as the cutting line of \( M_i \).

**Case 6.** When \( t_i \leq t_j \leq t_{i+1} \), such case can be treated as a combination of Cases 1, 2, and 3 by taking \( t_j \) and \( t_{j+1} \) as the cutting lines of \( M_i \).

Based on the above discussion, the risk of viral transmission produced by moving events \( M_i \) and \( M_j \) can be expressed as follows:

\[
TR(M_i, M_j) = \begin{cases} 
0 & \text{if } t_j \geq t_{i+1} \\
\int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} r(M_i(y), M_j(x), x) \, dx \, dy & \text{if } t_i \leq t < t_{i+1} \leq t_{j+1} \\
\int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} r(M_i(y), M_j(x), x) \, dx \, dy & \text{if } t_i \leq t < t_{i+1} \leq t_{j+1} \\
\int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} r(M_i(y), M_j(x), x) \, dx \, dy + \int_{t_{i+1}}^{t_{j+1}} \int_{t_j}^{t_{j+1}} r(M_i(y), M_j(x), x) \, dx \, dy & \text{if } t_i \leq t < t_{i+1} < t_{j+1} \\\n\int_{t_i}^{t_{i+1}} \int_{t_j}^{t_{j+1}} r(M_i(y), M_j(x), x) \, dx \, dy & \text{if } t_i \leq t < t_{i+1} < t_{j+1} \end{cases} \tag{5}
\]
Finally, given the trajectory of infected person \( j \), the risk \( R = (i,j) \) of exposed person \( i \) is computed by accumulating the TR values, which can be expressed as follows:

\[
R(i,j) = \sum_{i=1}^{n_{\text{infect}}(j)} \sum_{j=1}^{n_{\text{exposed}}(i)} TR(M_i, M_j)
\]  

(6)

where \( n_{\text{infect}} \) and \( n_{\text{exposed}} \) are the number of segments in the trajectory of the infected and exposed persons, respectively.

### 3.3. Weight graph accounting for the effect of obstacles

The premise of airborne transmission between two places is that virus-laden aerosols can float from one to the other. Therefore, given the complex topological structure of public indoor spaces, this study assumed that airborne transmission will only occur when people’s locations are mutually visible. However, the computation of the visibility between each pair of trace points is enormous, especially when the results need to be applied to an integral function, i.e. Equation (6).

To deal with the effect of obstacle on transmission risk, this study introduced a simplified algorithm based on the concept of ‘convex space’ (CS) from the space syntax analysis. A CS is defined as a space where no cross exists between the perimeter and the line between any two of its points [52], such as spaces A, B and C in Fig. 3. Notably, any two points in a CS are mutually visible. Therefore, the exposure risk can be directly computed based on Equation (5) for any moving events that occur in the same CS. As for moving events that occur in two different CSs, a constant weight is firstly calculated based on the statistics of inter-visible point pairs between the CSs. With CSs A and B in Fig. 3 as the example, a predefined size of grid cells, the weight from the transmission from A to B is defined as follows:

\[
W_{A,B} = \frac{\sum_{M_A} \sum_{M_B} TR(M_A, M_B) \ast W_{\text{FindID}(M_A, M_B)}}{n_A \ast n_B},
\]

(7)

where \( n_A \) and \( n_B \) are the total number of grids in spaces A and B, respectively, and \( q_i \) is the number of grids in space A that can be observed from grid \( i \) in space B. The weight \( w \) ranges from 0 to 1. The lower value of \( w \), the less inter-visible are the CSs. Notably, the weight graph is a symmetric matrix given that \( W_{A,B} \) is always equal to \( W_{B,A} \).

Finally, a weight matrix \( W \) can be constructed for any indoor space. Thus, Equation (6) is rewritten as follows:

\[
R(i,j) = \sum_{i=1}^{n_{\text{infect}}(j)} \sum_{j=1}^{n_{\text{exposed}}(i)} TR(M_i, M_j) \ast W_{\text{FindID}(M_i, M_j)}
\]  

(8)

where \( \text{FindID}() \) refers to the function that assigns the space ID to a moving event. The weighting term will constrain the magnitude of risks if the CSs are relatively less inter-visible. Notably, several indoor objects, such as tables and cabinets, may influence the airborne transmission. The effect of these objects can be modelled by assigning a fuzzy value, which corresponds the degree of obstacle, to each pair of grids during the visibility analysis. However, such an improvement requires more empirical information about the transport of airborne particles, which is beyond the scope of this study.

### 4. Simulation experiments

#### 4.1. Experimental setup

Given the rarely available open WiFi positioning datasets, this study adopted a dataset simulated by the PedSim plug-in in Rhino7 [53]. In PedSim, we can easily control the entrance, exit, visit interest and velocity of agents. As shown in Fig. 4, an indoor environment of a branching linear structure was created to represent a scenario in an underground shopping street, which is common in tunnels connecting public transportation systems (e.g., metro systems) and commercial centres. The indoor space had an accessible area of 895 m² with two exits, five restaurants and five stores.

We designed three types of agents of different visiting interests in PedSim, that is, eating, shopping and passerby. As a result, we obtained 179 trajectories with a sampling interval of 1 s and a total duration of over 20 min. Table 2 summarises the detailed configuration for these agents. Fig. 5 displays the whole dataset.

Without the loss of generality, a specific indicator called ‘aggregate risk’ (AR) was defined to test the performance of our method in different scenarios. To calculate AR, we firstly sampled one of the agents as an infected person without replacement and simulated the corresponding transmission scenarios. Then, by repeating the sampling and simulation processes for \( m \) times (no duplicated samples were allowed), AR can be expressed as follows:

\[
AR = \frac{\sum_{i=1}^{m} \sum_{j=1}^{N} R(i,j)}{m \ast N},
\]

(9)

where \( N \) is the number of agents.

In addition, from the perspective of an individual, a personal aggregate risk score (PAR) is defined as follows:

\[
\text{PAR} = \frac{\sum_{i=1}^{m} R(i,j)}{m}.
\]

(10)

To reduce the computational cost, we selected \( m = 30 \) to demonstrate our method in the following experiments.
4.2. Construction of the weight graph

According to the ‘least and fattest’ principle for the construction of CSs [54], we obtained 11 CSs from the experimental region. Fig. 6 shows the CSs and the heatmap representing their weight matrix. Notably, we set the cell size of 0.25 m$^2$ for the visibility analysis when calculating the weight matrix. The distant pair, such as CS5 and CS8, presented a significantly low weight, which meant that the transmission between these spaces was ignorable given that the diffusion of virus-laden aerosols was assumed to be largely blocked.

![Fig. 5. Overview of the simulated dataset.](image)

![Fig. 6. Construction of weight graph. (a) CSs of the made-up indoor space; (b) heatmap of the weight matrix.](image)

Notably, in the following experiments, each moving event was assigned with a CS if the start point was within the space.

4.3. Performance in capturing spatiotemporal decay effect

To exemplify the performance of the proposed method in dealing with the spatiotemporal decay effect, we selected one passerby instance as the patient in this simulation. Fig. 7(a) shows the estimated risks for each moving event of exposed people. The main risky areas were primarily located in the corridor along the trajectory of the patient, whilst...
the moving events within branching rooms were relatively safer given that they were distant from the patient.

The corridor area was used as an example to show the temporal decay effect. As shown in Fig. 7(b), no exposure risk was observed before the patient entered. Therefore, the agents who entered the space before the patient were not recognised as close contacts although they have visited the same place within a certain period. Then, the risk rapidly increased as the patient remained longer in this space, consistent with the notion that the longer a patient stays in a location, the more virus-laden aerosols are produced. After the patient left the shop, the space was still unsafe, whilst the average risk gradually decreased over time due to the reduction and dissipation of virus-laden aerosols.

### 4.4. Effectiveness of weight graph

As one of the main innovations of this study, the weight graph played an essential role in dealing with the obstacle effect in indoor spaces. To demonstrate the effectiveness of the weight graph, we compared the AR values estimated by Equations (6) and (8). In the following discussion, we termed the model of Equation (6) as ‘Unweighted’ and that of Equation (8) as ‘CS-based weights’. In addition, to test the validity of using CS instead of finer coordinates (i.e., dense grid) to quantify the obstacle effect, we added the estimation based on a dense grid with a cell size of 0.25 m$^2$ as the baseline model, termed as ‘Grid-based weights’.

Notably, we selected a grid cell of 0.25 m$^2$ to compute the ground truth because the maximum movement during a sampling interval (i.e., 1 s) of the agents in our simulation was approximately 0.5 m. Thus, each of the agent’s movements can be mostly represented by one grid.

As shown in Table 3, compared with the baseline, the AR was overestimated by 7.81% without weighting, whilst using CS-based weights underestimated the AR with a relative difference of 2.59% (100%–97.41%). From the individual perspective, the relative error of...
PAR corresponding to CS-based weights was 4.41% on average. By contrast, this value increased to 11.32%, and the maximum was over 36.50% without weighting. These findings indicated that exposure risk may be largely overestimated if the obstacle effect is ignored, and using the proposed weight graph can significantly reduce the relative error.

Notably, although the use of a dense grid theoretically provides a more accurate result, the weight graph will be more efficient in practice because the spatial association of moving events to grid cells can be very costly, especially for large indoor spaces. In addition, overestimation from the obstacle effect will be exaggerated for environments with enclosed subspaces. Hence, finer-grained CSs will be an appropriate solution in practice as a balance between accuracy and computational cost.

4.5. Model validation

4.5.1. Validation with computational fluid dynamics (CFD) simulation

Computational fluid dynamics (CFD) simulation is a powerful tool in modelling the aerosol transport and has recently applied to the evaluation of COVID-19 transmission [55]. In this section, a CFD simulation was conducted using Autodesk CFD [56] to further validate the proposed method.

According to the conventional Wells-Riley model [57], which is widely adopted in assessing the risk of airborne transmissible diseases, the infection probability $P$ could be expressed as:

$$P = 1 - \exp\left(-C_{\text{quantum}} \rho \Delta t_e\right)$$

where $C_{\text{quantum}}$ is the concentration of virus-laden aerosols, $\rho$ is volumetric flow rate of susceptible person’s breath, and $\Delta t_e$ is the exposure duration. Without loss of generality, $\rho$ was set to 1 to factor out its effect, and the concentration $C_{\text{quantum}}$ is represented by the concentration of carbon dioxide [58,59].

As shown in Fig. 8, the spatial distribution of transmission risk estimated by CFD and our method are highly correlated with a Pearson correlation over 0.7 in different scenarios. For both methods, the transmission risk radiates out from the patient’s location, and the strength generally declines as the distance increases. Also, both the methods are capable of capturing the obstacle effect by indicating a sharp change of risk close to the wall (most significant Scenario 2). One difference observed is that the dispersion of risks at the entrance of rooms is less significant in our method. This is mainly caused by the simplification of topology using the weight graph. However, such dispersion can be better modelled by using finer-grained convex spaces.

According to our analysis, different evaluation standards between the two kinds of risks are the main factor affecting the specific value of correlation coefficient. As we currently focus on justifying the effectiveness of our method rather than its numeric consistency with CFD simulations, the conversion of these risks is beyond the scope of this study.

4.5.2. Visualization of the risks caused by moving object

To further verify our method in dealing with moving objects, we selected two agents as the patients, respectively, and plotted the spatial distribution of transmission risks at different timestamp during their movement.

As shown in Fig. 9 (a), the risk is radiatively distributed along the patient’s path, and where the patient has not visited are relatively safe. By comparing Fig. 9 (b) and (c), a decay of risks can be observed after the patient left the indoor space. Similarly, in case 2, there is an increase of risk as the patient stayed longer in the bottom-right room (Fig. 9 (d)–(e)), and a decay after the patient left the space (Fig. 9 (e)–(f)). Different from the relatively uniform risk distribution along the patient’s path in case 1, the bottom-right room is dominant in transmission risk because the patient spent most time on ‘eating’ in this room. The visual results are generally consistent with our common perception about the diffusion and decay of transmissible virus, which to some extent justifies our approach.

Fig. 8. Comparison between the maps of transmission risk produced by CFD and our method in different scenarios. Each scenario represents the moment when a motionless patient stays at his location for 30 s. The risk of each grid cell is evaluated by assuming a motionless person at the cell’s centroid with a duration of 1 s. Maps are colored based on Natural breaks.
4.6. Influence of data uncertainty

4.6.1. Positioning error

Positioning error is inevitable in WiFi localisation due to obstacles and signal interference. In this experiment, we applied errors obeying a normal distribution $N(0, \sigma)$ to the agents’ coordinates $\sigma$ from 0.3 m to 1.8 m to investigate the effect of different magnitudes of positioning errors on PAR. Notably, to reduce the randomness effect, we ran each simulation with a specific $\sigma$ for 20 times to obtain stable assessments.

The experimental results are shown in Fig. 10, in which agents were ordered by their actual PAR for a better comparison. The exposure risks were generally underestimated with positioning errors. In addition, such underestimation became more significant as the value of $\sigma$ increased, especially when $\sigma > 1$. This finding is reasonable because according to the error propagation theorem, the expected distance between each pair of coordinates will be enlarged when normal errors are applied, especially when the error is larger than the actual distance measurement. As a result, the distance between agents will be generally overestimated, and the risk will drop due to the spatial decay effect.

Kendall’s $\tau$, which measures the correlation between two rankings [60], was calculated for the mean observation and ground truth. The value of Kendall’s $\tau$ was above 0.9 for all cases. Thus, despite the positioning errors at a certain level of uncertainty on risk evaluation, the estimations remained useful to distinguish risky contacts. Furthermore, the confidence interval was narrow for all cases, which proved the robustness of the proposed method for detecting randomness in people’s movements.

4.6.2. Omission of positioning data

Data omission is another source of uncertainty in positioning data due to the loss of WiFi signal and shutdown of WiFi access. Intuitively, if a user’s trace points are completely missed, we cannot evaluate neither his exposure risk nor his likelihood of infecting others. Hence, the complete omission of data is beyond the scope of this study.

In this experiment, we focused on a general scenario, in which people’s trace points were partially lost. We randomly discarded the
Fig. 10. Changes in PAR with different sizes of positioning error.

Fig. 11. Effect of data omission on PAR.

Fig. 12. Changes in risks with respect to different velocities. (a) Change ratio of PAR. (b) Change ratio of AR.
trace points with an omission ratio $p$ ranging from 0.2 to 0.8, and the effect of omission data was reflected by the changes in PAR. Notably, for each $p$ value, the simulation was run for 20 times to reduce the randomness effect.

As shown in Fig. 11, the PAR values were generally lowered because several risky exposures had not been recorded due to omission, and the underestimation exacerbated with a higher omission ratio. However, in terms of the Kendall’s $\tau$ over 0.97, the ranks of agents’ PAR were almost unchanged, which indicated that our method remained effective in distinguishing highly risky exposure. These findings further verified the robustness of our method for uncertainties from positioning data.

4.7. Influence of people’s velocity on exposure risks

In addition to the evaluation of personal infection risk, our method can be applied to support better decision making regarding epidemic prevention and control. Reduction of eat-in services and rapid departure from public areas are common policies for preventing the transmission of COVID-19 [61]. Under these measures, people generally spend less time for shopping and eating. In this experiment, we simulated the effects of these policies by adjusting the people’s velocity (equivalent to changing the stay duration). Firstly, a speeding-up ratio was applied to scale the time of the raw data to simulate the changes in velocities. Secondly, with the estimations corresponding to the original speed as the base, the change ratios of PAR and AR were calculated for further analysis.

Considering the limitation of people’s walking speed, the speeding-up ratio ranged from 0.5 to 2 with a step of 0.1 in this experiment, where a ratio larger than 1 represents the scenario of speeding up, and that with a value lower than 1 indicates the scenario of slowing down. As shown in Fig. 12, a high velocity led to relatively low exposure risks because people’s interactions were shortened. From the individual respective, as labelled in Fig. 12(a), PAR generally underwent a logarithmic reduction as people’s velocity increased. Thus, the exposure risk increased rapidly when people slowed down and spent more time in indoor spaces. A similar exponential trend can be observed in the change ratio of AR (Fig. 12(b)). These findings can be valuable for directing rational policies, such as the degree of eat-in limitation, and prompting pedestrians to reduce their duration in crowded areas reasonably.

5. Limitations of this study

The proposed method was currently demonstrated with simulated datasets: its validity was justified by an external comparison to the results of CFD simulations, and its robustness was examined by a series of sensitivity analyses on data uncertainty. Even though the simulation results have proved the effectiveness of our method, it is worth conducting further validation with real-life data in our future research. It is worth mentioning that access to real trajectory data would be limited due to privacy concerns. Compared with Bluetooth and GPS, WiFi receives relative low privacy concerns and the highest uptake because most people are likely to enjoy the free wireless services in public spaces [25,34]. Existing studies also demonstrated the prospect of some encryption techniques in WiFi localisation preservation [62]. Thus, we believe that the privacy issue would not seriously hamper the application of our method in practice.

Since the scope of this study is limited to building a spatial-temporal model for fast evaluation and identification of risky exposure to COVID-19, some perturbing factors from the environment as well as human behavior are not fully considered. The current model assumes that the transmission risk is governed by five factors, namely, temperature, humidity, sunlight, PD and duration, which are insufficient to cover all the possible cases due to the complexity of the airborne transmission of COVID-19 in reality. The effect of several preventative measures that may expedite the decay of viral infectiousness, such as wearing of masks, ventilation and disinfection, if quantifiable, can be further incorporated into our model via modifications of Equations (1) and (2). Furthermore, given that people’s behaviour, such as coughing and touching, cannot be recorded by the positioning data, the risks raised by large droplets and fomite transmission have not been considered in the current model. Regardless, these transmissions are markedly less efficient than airborne transmission and can be significantly reduced by wearing of masks and self-disinfection [63].

Moreover, considering the aerosol diffusion microscopically, the height of a person, human thermal plume [64] and the change in airflow pattern caused by the movement of people [59] will have a certain influence on the transmission risk. One possible solution is to convert these complex fluid dynamic processes into empirical coefficients to incorporate them into our model. For example, given that thermal plume is related to the type of clothing [64], we may adjust the temporal decay rate in Equation (2) appropriately based on what people are wearing. Similarly, the effect of varying infectiousness of virus at different infection stages [65], if more clinical data are available, should be further integrated into our method by setting specific initial parameters for each patient differently regarding their infection status.

Lastly, this work did not consider secondary transmission, which may lead to a certain level of risk underestimation. In future work, epidemiological approaches, such as the susceptible-infected model [66], can be used together with the proposed framework for a more precise evaluation of exposure risk.

6. Discussion and conclusion

Grounded in the ‘1 m and 15 min’ guideline for close exposure of COVID-19 from WHO, various smartphone-based approaches have been established using spatial–temporal thresholds for fast contact tracing. However, these methods assumed a crisp boundary to determine risky interactions whilst ignoring that the exposure risk will increase with a shortened distance, prolonged duration and under the influence of environmental factors. Given the airborne transmission of SARS-CoV-2, people will be at a certain level of risk despite the absence of an evident space–time intersection with infected individuals.

This study introduced an extensible framework to evaluate the transmission risk of COVID-19 based on fine-grained individual trajectories in indoor environments. By means of integration, the proposed method can well capture the spatial–temporal decay of airborne transmission based on the epidemiological characteristics of SARS-CoV-2. A weight graph was also implemented based on the space syntax technique to make our method adaptive to open and obstructed spaces. The experimental results showed that our method is insensitive to the either omission or deviation in positioning data, and can presents stable result in distinguishing relatively high-risky people from a crowd. It was also found that the proposed weight graph can significantly reduce the overestimation of risk by considering the obstacle effect and theoretically more efficient than the division by a dense grid, which might lead to intensive computation during spatial indexing of the positioning points.

Of interest, the proposed method can be easily adapted to simulate how different environmental factors, people’s mobility patterns and indoor layout will influence the infection risk [67,68]. Compared to the CFD simulation that focuses on the physical process of airborne transmission, our method offers a fast response to the spatial-temporal dynamics in people’s interactions, which makes it suitable in modelling individual-level exposure risk among large-scale population in complex environments. This framework is also applicable to outdoor spaces when people’s fine-grained trajectory is available. With the reopening of society, our method will contribute to efficient contact tracing and serve as a useful analytic tool for planners and policymakers in decision-making in post-pandemic societies.

CRediT authorship contribution statement

Pengfei Chen: Methodology, Visualization, Writing – original draft,
Funding acquisition, Conceptualization. Dongchu Zhang: Software, Methodology, Conceptualization. Jianxiao Liu: Conceptualization, Investigation. Izzy Yi Jian: Writing – review & editing, Validation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Natural Science Foundation of China [grant number: 42101351] and the Hong Kong Polytechnic University [grant number:1-W19P].

References

[1] T.P. Velavan, C.G. Meyer, The COVID-19 epidemic, Trop. Med. Int. Health 25 (2020) 228.

[2] T.P. Velavan, C.G. Meyer, The COVID-19 epidemic, Trop. Med. Int. Health 25 (2020) 861-866, https://doi.org/10.1111/tmi.13647.

[3] T.C. Bafile, M. Malekejnad, G.W. Rutherford, N. Razani, Outdoor transmission of SARS-CoV-2 and other respiratory viruses: a systematic review, J. Infect. Dis. 223 (2021) 1059–561.

[4] C. Yang, H. Shao, WiFi-based indoor positioning, IEEE Commun. Mag. 53 (2015) 150–157, https://doi.org/10.1109/MCOM.2015.7060497.

[5] C. Chen, Y. Chen, Y. Han, H.-Q. Lai, K.R. Liu, Achieving centimeter-accuracy indoor localization on WiFi platform: a frequency hopping approach, IEEE Internet Things J. 4 (2016) 111–121.

[6] F. Zafari, A. Gkelias, K.K. Leung, A survey of indoor localization systems and technologies, IEEE Commun. Surv. Tut. 21 (2019) 2568–2599, https://doi.org/10.1109/COMST.2019.2941155.

[7] B. Lo, I. Sim, Ethical framework for assessing manual and digital contact tracing for COVID-19, Ann. Intern. Med. 174 (2021) 395–400, https://doi.org/10.7326/M20-5814.

[8] V. Huang, Q. Wu, P. Wang, Y. Xu, L. Wang, Y. Zhao, D. Yao, Y. Xu, Q. Li, S. Xu, Measures undertaken in China to avoid COVID-19 infection: internet-based, cross-sectional survey study, J. Med. Internet Res. 22 (2020), e18718.

[9] K.N. Althoff, S.B. Coburn, D. Nash, Contact tracing: essential to the public health response and our understanding of the epidemiology of coronavirus disease 2019, Clin. Infect. Dis. 71 (2020) 1960–1961.

[10] D. Bresic, The German Corona-Data-Donation-App as an Example of the Concept of Data Donation, CITIP Blog, 2021, https://www.law.kuleuven.be/citip/blog/the-german-corona-data-donation-app-as-an-example-of-the-concept-of-data-donation/.

[11] J.W. Braddock, E.C. Alley, J.H. Huggins, A.L. Lloyd, K.M. Eveldt, Bidirectional contact tracing could dramatically improve COVID-19 control, Nat. Commun. 12 (2021) 232, https://doi.org/10.1038/s41467-020-20225-7.

[12] W. Hu, H. Janicke, S.K. Jha, A survey of COVID-19 contact tracing apps, IEEE Access 8 (2020) 15193–15196, https://doi.org/10.1109/ACCESS.2020.2952804.

[13] K.N. Althoff, S.B. Coburn, D. Nash, Contact tracing: essential to the public health response and our understanding of the epidemiology of coronavirus disease 2019, Clin. Infect. Dis. 71 (2020) 1960–1961.

[14] F. Ying, N. O’Clergy, Modelling COVID-19 transmission in supermarkets using an agent-based model, PLoS One 16 (2021), e0248821, https://doi.org/10.1371/journal.pone.0248821.

[15] X. Yang, J. Zuo, J. Zhou, P. Wang, S. Xu, C. Ai, Y. Zhao, D. Yao, Y. Xu, Q. Li, S. Xu, Measures undertaken in China to avoid COVID-19 infection: internet-based, cross-sectional survey study, J. Med. Internet Res. 22 (2020), e18718.

[16] C. Jenkins, A. Sunjaya, Social distancing as a strategy to prevent respiratory virus transmission, J. Infect. 82 (2021) 159–162, https://doi.org/10.1016/j.jinf.2021.01.009.

[17] K.A. Prather, L.C. Marr, R.T. Schooley, M.A. McDiarmid, M.E. Wilson, D.K. Milton, The efficacy of social distance and ventilation effectiveness in preventing airborne transmission of SARS-CoV-2 infection to inform risk-cost-benefit analyses of activities during the SARS-CoV-2 pandemic, PLoS One 16 (2021), e0245381.
[50] W. Chen, N. Zhang, J. Wei, H.-L. Yen, Y. Li, Short-range airborne route dominates exposure of respiratory infection during close contact, Build. Environ. 176 (2020), 106859, https://doi.org/10.1016/j.buildenv.2020.106859.

[51] P. Dabisch, M. Schuit, A. Herzog, K. Beck, S. Wood, M. Krause, D. Miller, W. Weaver, D. Freeburger, I. Hooper, B. Green, G. Williams, B. Holland, J. Bohannon, V. Wahl, J. Yolitz, M. Hevey, S. Ratnesar-Shumate, The influence of temperature, humidity, and simulated sunlight on the infectivity of SARS-CoV-2 in aerosols, Aerosol Sci. Technol. 55 (2021) 142-153, https://doi.org/10.1080/02786826.2020.1829536.

[52] B. Klarqvist, A space syntax glossary, NA 6 (2015).

[53] P. Wang, PedSim, Food4Rhino, 2017. https://www.food4rhino.com/en/app/pedsim. (Accessed 8 January 2022).

[54] B. Hillier, J. Hanson, The Social Logic of Space, Cambridge university press, 1989.

[55] X. Li, D. Lester, G. Rosengarten, C. Aboltins, M. Patel, I. Cole, A spatiotemporally resolved infection risk model for airborne transmission of COVID-19 variants in indoor spaces, Sci. Total Environ. 812 (2022) 152592, https://doi.org/10.1016/j.scitotenv.2021.152592.

[56] Autodesk CFD, n.d. https://www.autodesk.com/campaigns/cfd2020. (Accessed 13 April 2022).

[57] E.C. Riley, G. Murphy, R.L. Riley, Airborne spread of measles in a suburban elementary school, Am. J. Epidemiol. 107 (1978) 421–432, https://doi.org/10.1093/oxfordjournals.aje.a112560.

[58] C.Y. Vouriot, H.C. Barridge, C.J. Noakes, P.F. Linden, Seasonal variation in airborne infection risk in schools due to changes in ventilation inferred from monitored carbon dioxide, Indoor Air 31 (2021) 1154–1163.

[59] C.C. Wang, K.A. Prather, J. Smitsman, J.L. Jimenez, S.S. Lakkadwala, Z. Tufekci, L. C. Marr, Airborne transmission of respiratory viruses, Science 373 (2021), eaab9149, https://doi.org/10.1126/science.abab9149.

[60] J.I. Marden, Rank correlation methods, J. Am. Stat. Assoc. 87 (1992) 249–251.

[61] D. Wang, J. Yao, B.A. Martin, The effects of crowdedness and safety measures on restaurant patronage choices and perceptions in the COVID-19 pandemic, Int. J. Hospit. Manag. 95 (2021), 102915.

[62] P. Zhao, W. Liu, G. Zhang, Z. Li, L. Wang, Preserving privacy in WiFi localization with plausible dummy locations, IEEE Trans. Veh. Technol. 69 (2020) 11909–11925, https://doi.org/10.1109/TVT.2020.3006363.

[63] J.R. Port, C.K. Vinda, I.O. Owanu, M. Holbrook, R. Fischer, T. Busbmaker, V. A. Avanzato, J.E. Schulz, C. Martens, N. van Doremalen, SARS-CoV-2 disease severity and transmission efficiency is increased for airborne compared to fomite exposure in Syrian hamsters, Nat. Commun. 12 (2021) 1–15.

[64] S. Sun, J. Li, J. Han, How human thermal plume influences near-human transport of respiratory droplets and airborne particles: a review, Environ. Chem. Lett. 19 (2021) 1971–1982, https://doi.org/10.1007/s10311-020-01178-4.

[65] J.A. Backer, D. Eggink, S.P. Andeweg, I.K. Veldhuijzen, N. van Maarseveen, K. Vermaas, B. Vlaemynck, R. Scheper, S. van den Hof, C.B. Reusken, J. Wallinga, Shorter serial intervals in SARS-CoV-2 cases with Omicron BA.1 variant compared with Delta variant, The Netherlands, 13 to 26 December 2021, Euro Surveill. 27 (2022), https://doi.org/10.2807/1560-7917.ES.2022.27.6.2200042.

[66] A.E. Matouk, Complex dynamics in susceptible-infected models for COVID-19 with multi-drug resistance, Chaos, Solit. Fractals 140 (2020) 110257.

[67] H. Ugail, R. Aggarwal, A. Iglesias, N. Howard, A. Campuzano, P. Suiñez, M. Magsood, F. Azdil, I. Mehmood, S. Gleghorn, K. Taif, S. Kadry, K. Muhammad, Social distancing enhanced automated optimal design of physical spaces in the wake of the COVID-19 pandemic, Sustain. Cities Soc. 68 (2021), 102791, https://doi.org/10.1016/j.scs.2021.102791.

[68] J. Liu, W. Shi, P. Chen, Exploring travel patterns during the holiday season—A case study of Shenzhen Metro system during the Chinese Spring festival, ISPRS Int. J. Geo-Inf. 9 (11) (2020) 651.