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.
Simulating inter-individual contact in the inter-station passenger transfer system connecting multiple metro stations based on space–time path data

Zongchao Gu a, Toshihiro Osaragi b, Wei Lu a,⁎

a School of Architecture and Fine Art, Dalian University of Technology, Dalian, China
b School of Environment and Society, Tokyo Institute of Technology, Tokyo, Japan

Abstract

Determining passengers’ inter-individual contact in the metro station area (MSA) is an important issue to simulate and mitigate the spread of the Coronavirus disease 2019 (COVID-19) pandemic. Taking the inter-station passenger transfer system (IPTS) as an example, this study aimed to verify the passenger flows’ influence on the inter-individual contact around the MSA. Based on actual observed data, the passengers’ space–time paths (STP) in the network were obtained through an agent-based simulation. In this study, the direct contact model and the mediate contact model were used to describe the inter-individual contact in view of the passengers’ STP. The contact count and the exposure duration were defined as indicators to measure the contact degree of individual and the system. The results show that the time-varying trip distribution of the metro passengers significantly affected the inter-individual contact degree and the spatial distribution of contact risk region in the MSA. The intersection of passenger flow in different directions and the concentrated movement of passenger flow in the same direction increased the inter-individual contact and prolonged exposure in the morning. Through simulation experiments, the study verified the effects of controlling the flow direction and equalizing passenger flow generation measures aiming to reduce inter-individual contact and cumulative exposure duration.

1. Introduction

Metro system construction is an important driving factor for the underground space development of urban centre areas (Lin et al., 2021a). In the network trend of the metro system, multiple metro lines converge in the urban centre to form a metro hub area. Due to differences in the construction period, operating rights and service areas, the stations on each line are built in different locations and are connected by transfer channels to prevent conflicts between automobile and pedestrian transport in the hub area (Cui, 2021). Driven by economic and climatic factors, the surrounding commercial and business facilities open underground entrances and indoor public walkways connected to transfer passages and metro stations to provide convenient mobility for transfer passengers and facility users (Deguchi and Cho, 2003; Lin et al., 2021b). Transfer passages and underground streets have been developed into an inter-station passenger transfer system (IPTS) that covers the metro hub area in the city centres of mega-cities (Cui and Lin, 2016). The IPTS, connecting multiple metro lines, plays an essential role in improving urban transportation (Zacharias, 2014), commercial vitality (Xu and Chen, 2021; Yang et al., 2020), regional sustainability (Cui and Lin, 2016) and regional resilience (Sterling and Nelson, 2013).

Local metro users and transfer passengers in the IPTS have brought vitality to the sustainable development of the surrounding area of the metro station (Cui et al., 2021). As COVID-19 has spread globally, transmission via passengers using public transportation has become a risk dissemination channel for the further spread of the disease (Sanatan, 2020). The enclosed environment and frequent crowd movement make metro stations and surrounding underground space more vulnerable to the spread of respiratory diseases (Leng and Wen, 2021). Reducing public exposure and exposure time is considered an effective way to mitigate the spread of the outbreak (Sun et al., 2021). Although the contact between pass-by passengers is generally short, the cumulative interaction between individuals in complex passenger flows is nonnegligible. Quantitative measurement of person-to-person contact and exposure duration in the IPTS is expected to provide essential indicators for assessing spread risk and improving the safety performance of the metro system.

In the IPTS connecting several stations, the number of passengers and inter-individual interaction are complex and dynamic owing to the differences in the arrival frequency and passenger flow of connected...
stations (Gu et al., 2018; Nishimura and Osaragi, 2009). This makes the risk of personal exposure of passengers in the IPTS vary over time. Clarifying the dynamics of the exposure among passengers in the IPTS and its relationship with crowd behavior factors is an important issue in metro facility planning and public management. Based on the time series of the spatial relationship between individuals and others in crowds, this study aimed to examine the variation of inter-individual contact levels in the IPTS on the microscopic time scale.

Understanding the spatial distribution of inter-individual contact is also essential for the ventilation environment improvement of IPTS (Xu and Liu, 2018). The hourly differences in transfer passengers’ origin—destination cause their movement trajectories and spatial distribution in the IPTS network to vary over time (Gu et al., 2019). This leads to spatiotemporal heterogeneity in the occurrence of inter-individual contact events. Obtaining the risk areas and risk periods where transmission occurs is important for decision makers to take effective and targeted measures. Using the cumulative results of contact events on links of the network in different time intervals, this study aimed to visualize the spatiotemporal distribution of transmission risk areas.

This paper proposes a method of simulating the contact between the passengers in the underground space of the MSA. The proposed model considers the impact of environmental and behavioural factors in the microscopic system. Based on observed flow count data in an actual underground environment, the spatiotemporal distribution module was constructed to obtain the passengers’ space–time paths (STP) in the MSA. Direct and intermediate contacts between individuals’ STP were simulated by two models. Indicators were defined to examine the metro passenger flows’ influence on the contact degree of individuals and the system in different aspects. By quantifying the cumulative contact of individuals and the exposure duration on the link, we aimed to assess the contact risk of individuals and the system in different time intervals. We visualized the spatiotemporal distribution of contact risk to support the evaluation of non-pharmaceutical interventions.

2. Related works

2.1. Epidemic spread and metro system

A pandemic significantly changes the travel patterns of urban residents (Marra et al., 2022) and passengers’ perceptions of congestion (Kayvan et al., 2021). Existing studies have mostly focused on the impact of the pandemic on metro accessibility (Yang et al., 2022) and the travel willingness of passengers at the macro scale (Mützel and Scheiner, 2021). There have been few studies on the passenger dynamics concerning contact risk in the metro system (Lu et al., 2022). Lei et al. (2020) simulated the effect of the metro line load level on the probability of transmission between passengers during the inbound/outbound and ride phases but did not consider factors such as the plan layout of metro station facilities and the spatiotemporal distribution of passengers.

Research on inter-station transfer systems has mostly focused on traffic flow forecasting (Zacharias, 2000) and the level of service evaluation of pedestrian facilities (Lu and Han, 2016). Hu et al. (2021) applied a system dynamics model to analyze the susceptibility of epidemic transmission within a single metro station facility. However, little attention has been paid to the interaction between transfer passengers flows from several metro stations. In this study, we simulated the inter-individual contact between passengers in inter-station transfer systems connecting multiple metro stations and surrounding facilities.

2.2. Inter-individual contact model

Most existing models focus on the influence of physical contact and exposure duration among people staying in a specific spatial unit (Gu et al., 2022) or a spatial mesh (Huang et al., 2021). The interactions between individuals in motion is relatively rare. Most studies of inter-individual transmission on motion have been based on hypothetical contact probabilities (Lei et al., 2020) or simulating inter-individual contact by random movement (Xiao et al., 2021) and less often on the effect of the spatiotemporal behavior of individuals in a crowd on the degree of systematic contact.

Previous studies have generally used the number of infections as the indicator for evaluating the risk of transmission in public space (Lei et al., 2020). As the prediction results have been significantly influenced by the probability of transmission and the proportion of asymptomatic people, many simulations have been required to determine the risk of the study area under a wide range of parameters (Johansson et al., 2012). The cumulative number of close contacts and exposure duration between individuals in indoor space have also been used as indicators to evaluate the epidemic transmission risk in the public facilities (Xiao et al., 2021). This study defined the cumulative contact degree of people in motion as an intermediate indicator to explore the influential factors on the transmission risk.

Simulation analysis, such as the cellular automata approach (Kleczkowski and Grenfell, 1999), bipartite graphs (Subank et al., 2004), and social networks, can estimate the person-to-person contact in urban and indoor facilities. Pedestrian dynamics have attracted attention in the field of mathematical epidemiology. Compared with the macro model, simulation of crowd pedestrian movement enables us to detect individuals of the same aggregate epidemic spreading patterns (Johansson et al., 2012). The effect of pedestrian dynamics on epidemic spread was proved by simulating the contact between individuals randomly walking in an indoor environment (Xiao et al., 2021).

Simulating intermediate contact between all individuals in a two-dimensional plane usually requires a large amount of computation and storage (Zhao et al., 2020). Based on a link-node network, Flurin et al. (2020) simulated passenger dynamics on board and in stations and used them for service-level assessment and congestion estimation. To rapidly assess the risk of contact and the effectiveness of non-medical interventions, this paper proposes two inter-individual contact models based on a two-dimensional space–time coordinate system.

2.3. Space–time path generation

Determining the STP of pedestrians in an underground space is the prerequisite for contact analysis in IPTS. Subway card swipe data (Liu et al., 2020) and QR code data (Wu et al., 2021) enable us to collect the personal information of pedestrians passing through ticket wickets or the entrances/exits of public facilities. However, people’s movement during passages are not included in this type of data. Owing to the enclosed environment, GPS and location based services data are unavailable in the underground space. Indoor positioning system data are limited to pedestrians carrying tagged devices (Huang and Wu, 2017). Ultra-Wideband positioning (Zhang et al., 2011) and Radio Frequency Identification techniques (Qu and Matsushita, 2010) are seldom applied to public transportation facilities, such as airports and railway stations. Computer vision techniques help to collect the flow counts and contact between pedestrians in the observed area (Asada and Osaragi, 2000). However, in the video data of a limited coverage area, the continuous motion traces of passengers around metro stations cannot be obtained. Wi-Fi localization techniques make it convenient to obtain the fine-grained trajectory data of devices with open Wi-Fi access in indoor environments (Chen et al., 2022), but hard to collect all passengers’ space–time paths in the underground pedestrian network. Hence, there have been few studies on the inter-individual contact in the IPTS through motion tracing.

If the origin and destination (OD) of pedestrians are known, it is possible to obtain the inter-individual contact by recreating their spatial motion in the MSA. Using person trip (PT) data, Osaragi (2009) estimated passengers’ time-specific positions in railroad networks. Based on the observed flow count data, Ahn and Tsukaguchi (2015) estimated passengers’ OD and motion traces in MSA. Nishimura and Osaragi
(2009) estimated railroad users’ possible destinations based on their personal trip data and simulated pedestrian traffic to evaluate the service level in front of a train station. Gu and Osaragi (2017) adopted the dynamic discrete route choice model to calibrate pedestrians’ route choice tendency in IPTS. The effects of periodic train arrivals on the uncertainty of interaction among individuals in the wicket-front plaza (Osaragi, 2004) and the transfer passages (Gu et al., 2019) have been supported by agent-based simulations. In this paper, we propose a simulation method to determine the time-specific movement of passengers in an IPTS connecting multiple metro stations.

3. Simulation modelling

Fig. 1 shows the scheme of the proposed model. Section 3.1 introduces initial data processing. Section 3.2 describes how to simulate pedestrians’ space–time paths in the IPTS and the basic setting of the simulation. Section 3.3 defines the individual contact indicators based on two contact models; the model outputs were used to examine the metro passengers’ influence on the individual contact degree and the contact risk distribution in the IPTS.

3.1. Study area and data processing

3.1.1. Study area

We selected Tenjin Underground Shopping Streets, located in the city of Fukuoka, Japan, as the study area. The underground streets connecting two metro stations and one light rail station provide passages for transfer passengers and pedestrians crossing the arterial road. As shown in Fig. 2, Tenjin Minami (TJM) Station and Nishitetsu Fukuoka (NTF) Station are the terminal stations, and Tenjin (TJ) Station is a middle station on a metro line leading to an intercity railway station. The office and commercial buildings in the study area are directly connected to the IPTS.

3.1.2. System representation

The IPTS in the study area is represented as a link-node network. As shown in Fig. 3, links represent the passages, while nodes represent intersections and exits in the underground streets. The circles in Fig. 3 represent the position of entrance/exit nodes as people enter and leave the IPTS. Dots represent the cross nodes where pedestrians choose their subsequent routes. According to the connecting facilities, this study divided the OD nodes into station nodes, building nodes, and ground nodes (Fig. 3).

Fig. 1. Research Framework.

Fig. 2. Layout of Tenjin Underground Streets.
3.1.3. Available data

Pedestrian traffic survey data: This study used the data of the pedestrian traffic survey carried out by the Fukuoka city government. Considering the inter-individual contact condition without non-medical interventions, this study chose the traffic survey data before the pandemic. This survey counted the number of people passing by each observation point (red links in Fig. 3) on a weekday (March 7, 2012). The flow counts of the two directions were collected each hour from 7 am to 8 pm.

Train arrival data: This study considered the contact between individuals in the study area to be affected by the periodic passenger flow from the metro stations. The train arrival data were obtained in view of the train schedule and travel time between stations (Gu et al., 2019).

3.1.4. Numerical estimation data

OD matrix: Based on the pedestrian survey data, a set of hour-specific OD matrices is obtained through a numerical estimation process until they have reached their destination. The time-space path simulation approach was operated by the second in a time interval of 1 h. Agents were assumed to appear at the OD node when the simulation reached their generation time. Each agent was assumed to continuously update their position each second according to dynamic interaction between individuals, such as the position bias resulting from conflict-avoidance behaviours and information exchange (Zhao et al., 2022).

Route choice probability: In our previous work (Gu and Osaragi, 2016), we proposed a dynamic discrete choice model to describe the path decision behavior of the agent moving to the end of the road segment in the study area. The free walking speed of each agent is randomly assigned, with variables satisfying a normal distribution (Willis et al., 2004) based on their gender attributes.

As this study focused on the contact between individuals in the transfer system, the movements and activities of people staying in underground shops was not considered. Considering the computation load of simulation, this study did not consider the effect of dynamic interaction between individuals, such as the position bias resulting from conflict-avoidance behaviours and information exchange (Zhao et al., 2022).

3.2. Space-time path simulation

An agent-based simulation approach is adopted to recreate agents’ space–time paths in the IPTS network. The simulation programming language was C#, and the computing platform was Visual Studio 2017. The IPTS network of study area was input as the classes of links and nodes, which contain the geographic information of actual environment. The OD matrix and route choice probability matrix were input as the initial data of the simulation.

3.2.1. Basic setting

Passengers in the study area were denoted by the class of agent with personal attributes. The number of agents generated at each entrance/exit (OD node) was determined according to the estimated OD matrix. The destination information of agents generated at each entrance was proportionally determined based on the estimated OD matrix. Each agent only enters the network once, and their destination does not change during the simulation process.

We determined the generation time of the agent according to the type of OD nodes. At the non-station nodes, the time when an agent enters the system was randomly assigned. The time of agents generating from a station node was determined by the train arrivals of the connected station in the simulation (Gu et al., 2019). The setting methods are explained in detail in Appendix C.

Agents’ gender attributes were randomly set according to the proportion shown in Table 1. The gender proportions were assumed according to the empirical settings in previous research (Osaragi, 2004). The free walking speed of each agent is randomly assigned, with variables satisfying a normal distribution (Willis et al., 2004) based on their gender attributes.

As this study focused on the contact between individuals in the transfer system, the movements and activities of people staying in underground shops was not considered. Considering the computation load of simulation, this study did not consider the effect of dynamic interaction between individuals, such as the position bias resulting from conflict-avoidance behaviours and information exchange (Zhao et al., 2022).

3.2.2. Simulation process

The passengers’ movement in the IPTS was simplified to the motion along the link-node network. As shown in Fig. 4, the simulation approach was operated by the second in a time interval of 1 h. Agents were assumed to appear at the OD node when the simulation reached their generation time. Each agent was assumed to continuously update their position each second according to dynamic crowd density of the located link (Weidmann, 1992). When agents moved to a cross node, they would randomly determine their subsequent routes in view of the predetermined route choice probability. All agents were set to repeat the above process until they have reached their destination. The time-specific positions, $\{i, d\}$, of each agent were recorded by its position on the located link by the second. By sequentially connecting the time-series position of each agent, we could obtain the passengers’ STP in the network.

3.3. Contact model

3.3.1. Direct contact module

A link is taken as the unit of analyzing the contact between
individuals in the IPTS network. For each link \( l \) in the network, the trajectories of all agents through \( l \) during each second \( t \) can be represented by a set of space–time vectors in a two-dimensional coordinate system (Fig. 5(a)). The vertical axis represents time, and the horizontal axis represents the location of the road segment. The time and location of the pedestrians’ meeting on the link can be determined by the intersection of the space–time vectors (Fig. 5(b)). The event is represented by a space–time intersection without time duration and spatial scope. The contact relation between two agents is undirected.

We define a logical variable, \( B_{ghl} \), to represent the time-specific contact event between two agents, \( g \) and \( h \), on the link \( l \) as follows:
### Table 2

Simulation results in different time intervals.

| Time Interval       | Total       | Mean   |
|---------------------|-------------|--------|
| 8:00–9:00           | 1,768,748   | 154.0185 |
| Intermediate contact counts | 2,430,340 | 211.6284 |
| Cumulative exposure duration | 1.05E + 07 | 913.9747 |
| Agent number        | 11,484      | 17,263 |

| Time Interval       | Total       | Mean   |
|---------------------|-------------|--------|
| 12:00–13:00         | 1,880,074   | 105.4321 |
| Intermediate contact counts | 2,364,378 | 136.9738 |
| Cumulative exposure duration | 9.39E + 06 | 543.7284 |
| Agent number        | 17,263      | 19,741 |

| Time Interval       | Total       | Mean   |
|---------------------|-------------|--------|
| 18:00–19:00         | 2,742,638   | 128.9311 |
| Intermediate contact counts | 3,596,493 | 182.1839 |
| Cumulative exposure duration | 1.42E + 07 | 719.2135 |
| Agent number        | 19,741      |        |

The exposure duration, approximately assumed that the breathing dissemination lasts 5 s in this study. The contact count between each agent pair g-h in the IPTS network can be obtained by.

\[
B_{ghl} = \begin{cases} 
1, & g \text{-does not cross } h \\
0, & g \text{-crosses } h
\end{cases}
\]

where \(g\) and \(h\) are the space–time vectors of agent \(g\) and agent \(h\), respectively, at second \(t\). The cumulative direct contact counts (DCCs) of agent \(g\) in the IPTS network are obtained by.

\[
C_{gh} = \sum_{j} \sum_{l} B_{ghl}
\]

(2)

The cumulative direct contact counts (DCCs) of agent \(g\) in the IPTS network are obtained by.

\[
DCC_g = \sum_{h} C_{gh}
\]

As the contact event between two agents would be counted separately in each agent’s loop, the cumulative DCCs on a specific link \(l\) are obtained by.

\[
LC_l = \frac{1}{2} \sum_{j} \sum_{g} \sum_{h} B_{ghl}
\]

(4)

Accordingly, the time-specific DCCs in the system are obtained by.

\[
SC_l = \frac{1}{2} \sum_{j} \sum_{g} \sum_{h} B_{ghl}
\]

(5)

\(SC_l\) stands for the total counts of contacts occurring in the system at second \(t\). These variables are defined to indicate the spatiotemporal contact degree of individuals and the IPTS.

#### 3.3.2. Intermediate contact module

This module aims to model the indirect contact between individuals. Agents are assumed to continuously leave dissemination, for example, droplets on their motion trace, which generates a space–time impact area in the spatial dimension. If agent \(h\)’s time–space vector goes through the space–time area of individual \(g\), it is considered that \(h\) has been indirectly affected by \(g\) through dissemination media. This event is defined as the intermediate contact event in this study. The intermediate contact event has a duration in the time dimension and a scope in the spatial dimension.

Exposure duration is defined to indicate the contact degree between two agents. For agent \(g\) on the link \(l\), the space–time area at the second \(t\) is constructed by moving up the space–time vector in the direction of time axis. According to dissemination duration of coaching (15 s indoor and 5 s outdoor) in an existing study (Xiao et al., 2021), we approximately assumed that the breathing dissemination lasts 5 s in this study. The exposure duration, \(D_{ghl}\), is obtained by the time difference between the points where vector \(h\) intersects the area edges (Fig. 6(b)). The relation between two pedestrians is directed. For each agent pair \(g-h\), the duration of \(h\)’s exposure to \(g\)’s impact area is obtained by the summation of the contact duration on each link, as follows:

\[
D_{ghl} = \sum_{j} \sum_{h} D_{ghl}
\]

(6)

The cumulative exposure duration (CED) of agent \(h\) in other agents’ impact area is obtained by.

\[
CED_h = \sum_{l} \sum_{g} \sum_{j} D_{ghl}
\]

(7)

We denote the intermediate contact result of each agent pair \(g-h\) by a logical matrix, \(M_{ghl}\), with the element of 0 and 1 based on the agent-to-agent exposure duration, as follows:

\[
M_{ghl} = \begin{cases} 
1, & D_{ghl} > 0 \\
0, & D_{ghl} = 0
\end{cases}
\]

(8)

The number of agents contacted by \(h\) is obtained by.

\[
ICC_h = \sum_{g} M_{ghl}
\]

(9)

Accordingly, the cumulative exposure duration, \(LD_l\), on link \(l\) is obtained by.

\[
LD_l = \sum_{j} \sum_{h} \sum_{g} D_{ghl}
\]

(10)

where \(D_{ghl}\) is the duration of agent \(g\)’s exposure to the impact area of agent \(h\) on link \(l\) at second \(t\).

#### 3.4. Model output

The model outputs the following result as the initial data for further analysis.

**Agent-specific data:** The accumulative contact results of each agent in the simulation are recorded by a list with the length of agent number. The output data contain the direct contact count, cumulative exposure duration, and the intermediate contact count (ICC). These variables indicate the contact degree in the aspect of individual.

**Agent-to-agent data:** The interaction between individuals is output by two-dimensional matrices with the size of the agent number. These data are expected to support assessing the risk of epidemic spreading, in future work.

**Link-specific contact event:** The cumulative contact counts and exposure duration on a specific link indicate the spatial distribution tendency of the contact risk in the IPTS. Visualization based on the link-specific contact indicators is expected to clarify the influential factors of the contact distribution and assist in the risk assessment in the MSA.

### 4. Results

#### 4.1. Global contact degrees in different hours

This study selected 1-h data in the morning (8:00–9:00), noon (12:00–13:00), and evening (19:00–20:00) on a weekday as an example and simulated the direct and intermediate contact between individuals in the IPTS. As shown in Table 2, the total values of the three indicators of the evening group were higher than those at other hours, which indicates a high global contact degree of the IPTS in the evening rush hour. The number of agents at noon was slightly less than the number in the morning, but the total contact indicators were close to those of the evening group. This indicates that the individual number is not a decisive factor of the global contact degree of the IPTS.
contact in the aspect of individual. The average value of the contact indicators in the noon period was significantly lower than those of the other two periods, while its agent number was close to that in the evening rush hour. This also indicates that the individual contact degree is not determined by the number of individuals. The effect of passengers’ OD information and spatiotemporal distribution on the contact between individual will be examined in subsequent sections.

As shown in Fig. 7, the number of direct contacts occurring in the system fluctuated with time; so, it was not possible to estimate the direct contact degree of the system over the whole time period by time sampling. The number of contacts in the system fluctuated most frequently during the morning hours (blue curves), which may be related to the intersection of different directional interchange flows during commuting hours. In the evening hours (green curves), the distribution of contacts was consistently high for a longer period of time. As mentioned above, although the total number of agents in the noon hour was close to that in the evening, the degree of direct contact (red curves) was indeed lower than that in the evening. This suggests that the difference in contact degrees is due to temporal variation factors such as OD distribution rather than occasional factors such as train arrival.

Fig. 7. Time-specific direct contact counts of IPTS in different time intervals.

Fig. 8. Spatiotemporal distribution of direct contact counts: (a) morning; (b) noon; (c) evening.

Fig. 9. Relation between traffic volume and direct contact counts at each link: (a) morning; (b) noon; (c) evening.
issue will be examined in subsequent sections.

4.2. Spatiotemporal distribution of inter-individual contacts

4.2.1. Result of direct contact event

Fig. 8 shows the time-varying spatial distribution of direct contact positions in the IPTS network. The height of the slices represents the contact degree of links. The color of the slice represents the ranking of the contact level of the link among all the links in the system. Although the total number of contacts was similar during the morning (1.76 million) and noon (1.82 million) hour, their spatial distribution tendencies were significantly different. The high-risk region during the morning hour was concentrated on the links close to the stations (Fig. 8 (a)). The contact events during the noon hour were more distributed in the passages connecting the commercial facilities in the central part (Fig. 8 (b)). This indicates that the periodic passenger flows generated by the station are a possible cause of the high risk of contact in front of the station. The comparison between Fig. 8 (b) and (c) reveals that the frequency of contacts in the network significantly increased in the evening hour, while the high-risk region distribution of direct contact was similar.

The linear correlation between link traffic volume and link contact counts cannot be seen in Fig. 9. Links with a traffic volume less than 1,000 people/hour seldom enabled direct contact between individuals. With growth of the link traffic volume, the upper limit of link contact counts increased. The effect of traffic volume on the upper limit of contact was higher at the noon hour and lower during the morning hour. This also indicates that the traffic volume is a necessary condition for the direct contacts, rather than the dominant factor.

4.2.2. Results of cumulative exposure duration

Fig. 10 shows the spatiotemporal distribution of cumulative exposure duration at each link.

Fig. 11. Correlation between direct contact counts and intermediate contact duration at each link.
between individuals in two types of motion relation. During the noon (Fig. 12(b) and (e)) and the evening (Fig. 12(c) and (f)) hours, motion relation did not have a significant effect on the spatial distribution tendency of exposure duration. In the morning (Fig. 12(a) and (d)), the exposure duration between agents in the same direction showed a higher distribution tendency on the links close to the stations (dotted circles in Fig. 12(a)). The effect of concentrated movement of passenger flows in the same direction will be examined in the next section.

4.3. Individual contact results

4.3.1. Frequency distribution of individual DCC and CED

Individual DCCs of agents in the IPTS were studied in view of their frequency distribution results. As shown in Fig. 13, the results at noon
(Fig. 13(b)) and evening (Fig. 13(c)) had a Poisson distribution tendency, with the highest frequency at 60 times per people. The morning result had a distribution tendency, with two peaks at the columns of 60 and 300 times, respectively (Fig. 13(a)). The maximum of DCC at noon (360 times) was smaller than that in the morning (450 times) and evening (480 times). This suggests that individual contact counts are affected by the passenger flows at commuting hours.

Fig. 14 shows the time-specific frequency distribution of the

![Fig. 14](image1)

**Fig. 14.** Frequency of contact duration in different hours: (a) morning hour; (b) noon hour; (c) evening hour.

![Fig. 15](image2)

**Fig. 15.** Direction specific frequency of contact duration: (a) morning; (b) noon; (c) evening.

![Fig. 16](image3)

**Fig. 16.** Visualizing the OD distribution of high-contact population in the IPTS network: (a) upwards trip; (b) downwards trip.
cumulative exposure duration of the agents’ intermediate contact. The agents were more frequently distributed in the range of 500–750 s. The maximum values of the exposure were higher during the morning hour (50 min) and lower during the noon hour (37.5 min). Similar to the result of DCCs (Fig. 13(a)), the proportion of the population in the long-duration range (1500–2500 s) was higher during the morning hour (dashed rectangle in Fig. 14(a)).

We classified the exposure duration of intermediate contact between agents according to their motion relation. The frequency distribution of agents’ exposure duration to the impact area of the agents moving in the same and different directions is shown in Fig. 15. A comparison between two classes in the same hour showed that agents were exposed longer in the impact area of the same direction. The frequency distribution of the same direction class at different time intervals showed that agents’ exposure duration was longer during the morning hour (Fig. 15(a)). The impulsive transfer flow during the morning hour is a possible reason for the long duration contact. We will examine this issue in Section 5.2.

4.3.2. Spatial distribution of high-risk population of direct contact in the morning

A significant high-frequency tendency appeared in the range from 270 to 390 (Fig. 13(a)). This indicates that a group of agents suffer from high risk of trajectory contact in the morning rush hour. The growing agent proportion at the high contact range (red rectangle in Fig. 13(a)) is a possible reason for the mean value of the direction contact count of IPTS rising higher than for other hours. To clarify this issue, we abstracted the agents with a contact count in the range of 270–390 to visualize their OD distribution. As shown in Fig. 16, the trip number between OD pairs was visualized by a spatial quadrangle representing the generation numbers by the color of the quadrangle. Trips with different directions between each OD pair are plotted separately. OD pairs with trip counts less than 50 people were excluded.

As shown in Fig. 16(a) and (b), abstracted agents are mainly distributed between the TJM station and the entrances/exits at the upper end of the IPTS and between the metro stations. This indicates that the high contact risk group is composed of the local passengers commuting by subway and the transfer passengers between stations during the morning rush hour.

4.4. Influential factors on individual contact indicators

4.4.1. Correlation between individual contact indicators

The DCCs and CED of the individual showed different correlations during the three time intervals (Fig. 17). The linear positive correlations were obvious in different hours. There were samples located in the area close to the exposure duration axis during the morning and the evening hour (dotted red circle in Fig. 17(a)). This indicates that there are a certain number of passengers with a long exposure duration of intermediate contact without being conscious of it during the commuting hours. This issue will be discussed in Section 4.4.2 in detail.
Fig. 18 shows the significant correlations between the number of people with the direct and intermediate contact with each agent. The vertical distance between points and the diagonal (red arrow in Fig. 18 (a)) means the cumulative number of intermediate contact without track intersection for each agent. This relation was defined as the unperceived contact in this study. The trend lines of the sample distribution show that the range of unperceived contact counts increased with the growth of direct contact. Samples away from the diagonal were present in the region near the vertical axis in Fig. 18 (a). These samples with little direct contact had a high possibility of receiving unperceived contact in the morning hour. These samples were less likely to cross over with other individuals’ STPs; thus, they were less aware of the risk of exposure within the system and may have been less willing to take protective measures. Thus, identifying the spatial movements and spatiotemporal distribution of samples with unperceived contact counts (UCCs) plays an important role in arousing the public’s attention to the potential contact risk in IPTS.

4.4.2. Relation between travel distance and individual contact indicators

To further clarify the issue of unperceived contact, we examined the effect of travel distance on individuals’ UCCs in the three time intervals. There was no obvious linear correlation between travel distance unperceived contact counts in view of the R-square values (Fig. 19(a–c)). The positive linear correlation between distance and direct/intermediate (Fig. 19(d–f) and Fig. 19(g–i)) contact count was significant. Fig. 19(a) shows that the group of samples with distance of 600 m
had a larger unperceived contact number (up to 250–300 counts) during the morning hour (dashed red rectangle in Fig. 19(a)). As shown in Fig. 19(b), the number of samples with a high UCC in the 600 m interval decreased significantly. As the transfer distances between stations are around 500–600 m, these samples with high UCC in the morning hour were considered to be related to metro passengers. The tendency of the sample distribution in Fig. 19(c) approximated to the morning hour (Fig. 19(a)), but few samples had UCC values above 200. Because the maximum travel distance of the sample of individuals in the evening (600 m) was less than the value of morning hour (800 m), passengers with long travel distance were likely to have indirect contact with more individuals without being aware of it.

The positive linear correlation between distance and DCC was significant (Fig. 19(d–f)). There is almost no sample distribution on the upper part of the trend line, which indicates that a long travel distance is a necessary condition for direct contact. The slopes of the trend lines in three time intervals are similar, which indicates that the OD distribution and the agent number do not significantly affect the linear relationship between individual distance and the DCC. Taking the samples around 600 m as an example (Fig. 19(d)), the DCC values for the morning hour were distributed in the intervals of 0–100 and 150–400, which are significantly lower than those for the noon hour (250–350) and the evening hour (250–400). This may have been caused by the interchange traffic moving in the same direction during the morning hour, as we will verify in Section 5.1. In the morning hour, there was a concentrate distribution of samples close to the horizontal axis in the range of 600 m (dashed red rectangle in Fig. 19(d)). In view of the results in Fig. 18(a) and Fig. 19(a), the unperceived contact in this range was considered to be received by the metro passengers with little inter-individual direct contact. This issue will be examined in the following section.

Linear correlation between distance and ICC is shown in Fig. 19(g–i). Compared to the result of direct contact (Fig. 19(d–f)), there were few samples distributed close to the horizontal axis, which indicates that travel distance is a sufficient and necessary condition for intermediate inter-individual contacts in the IPTS. The slope of the linear regression trend line increases with the total number of individuals, suggesting that the number of agents in the system has an effect on the risk of individual indirect exposure. Besides that, samples in the range of 600 m in Fig. 19(d) and (g) had a discrete distribution trend in the morning, which means that the intermediate contact counts of these samples do not correspond to their direct contact results. This issue will also be discussed in the section 5.1 and section 5.2.

To figure out the reason of unperceived contact, we abstracted the samples in the dashed rectangles in Fig. 19(a) and (d) and plotted their OD distribution in Fig. 20. Fig. 20. OD distribution of abstracted agents: (a) abstracted agents in Fig. 19(a); (b) abstracted agents in Fig. 19(d).
agents from Fig. 20(d) were distributed between the TJM station and TJ station or exits in the north (Fig. 20(b)). This is because that the simulation ignored the agents existing in the study area before the start time. The passengers from the first arrival train at the TJM station arrived at their destination without tracking contact with other agents. In addition, the number abstracted from Fig. 19(d) was significantly less than that from Fig. 19(a). This means that the large proportion of unperceived contact in the morning is not caused by the agents with seldom direct contact but by the frequent intermediate contact of the metro passengers.

Fig. 21 shows the correlation between agents’ travel distance and cumulative exposure duration. It was found that the exposure duration of agents with distance from 450 to 750 m showed extension tendency during the morning rush hour. As the range is consistent with transfer distance between stations, it was considered that the duration extension was related to periodic transfer flows between stations. This issue will be examined in the next section.

5. Applications
This study took the morning rush hour as an example to simulate the effect of two possible non-pharmaceutical interventions on the inter-individual contact between pedestrians in the study area. We also verified the influence of the directionality and agglomeration of subway passenger flows on the proposed contact indicators through simulation experiments.

Table 3
Simulation results under different flow control plans.

|                          | No control | Plan 1          | Plan 2          |
|--------------------------|------------|-----------------|-----------------|
|                          | Total      | Mean            | Total           | Mean            | Total           | Mean            |
| Direct contact count     | 1,768,748  | 154.0185        | 1,308,178       | 113.9131 (-26 %)| 1,279,904       | 111.4511 (-27 %)|
| Intermediate contact     | 2,430,340  | 211.6284        | 2,027,409       | 176.5421 (-16 %)| 1,982,271       | 172.6116 (-18 %)|
| Cumulative exposure      | 1.05E + 07 | 913.9747        | 11,322,956      | 985.9767 (+7 %) | 11,470,249      | 998.8026 (+8 %) |

Fig. 22. Passenger flow control plans.

Fig. 23. Time-specific direct contact counts of IPTS in different flow control plans.
5.1. Directional control

5.1.1. Simulation experiment setting and results

The agents moving in different directions were administrated to take separate passages in the morning. We proposed two control plans for flow control in the morning hour when most shops were closed (Fig. 22). The utility of motion in the controlled direction was set to infinity to prevent the agent from passing through. We assumed that all agents knew the control plan beforehand, and they would remake their choice at each cross node related with the controlled link according to the updated probability matrix. Contact indicators were obtained through a simulation experiment based on different flow control plans.

As shown in Table 3, the contact counts of the direct and intermediate module decreased in the simulation experiment. Two flow control plans played similar roles in mitigating the contact counts. The direct and intermediate contact counts were reduced by 26–27% and 16–18%, respectively. However, the cumulative exposure duration in the study area was extended by 7–9% in the proposed plans. As mentioned in section 4.2.2, there was a significant positive correlation between CEDs and DCCs on the links. The different variation of DCC and CED of the study area also indicates that the flow control changes the linear correlation between DCC and CED on the link. We will further study this issue in the next paper.

Fig. 23 shows the number of direct contacts of IPTS in the time sequence for the two flow control plans. The fluctuation patterns in both scenarios remained the same as when there was no control, which indicates that the dynamic fluctuations in the DCC of the system are independent of passenger crossings. The downward shift of the curves in both scenarios indicates that the control plans improved the level of direct contact on average for all moments during the time interval.

5.1.2. Spatial distribution tendency of direct contact counts

As shown in Fig. 24, the distribution of direct contact events in the upper part of parallel corridors (dashed rectangle in Fig. 24(a)) was significantly reduced. The influence of the proposed plans on the spatial distribution of risk region was not consistent. The effect of plan 1 on the lower part of the parallel corridors was not significant. The link close to the TJM station (dotted circle in Fig. 24(c)) had more contact events in plan 2. In plan 1, the contact events were concentrated on a link close to the NTF station (dotted circle in Fig. 24(b)), while fewer events were located on the other parts of parallel corridors. Thus, it is more convenient to adopt additional measures, such as ventilation and air disinfection, to reduce the direct contact risk between individuals on this section.

5.1.3. Assessing the performance of the proposed plans in mitigating individual contact

Fig. 25 shows the influence of the two control plans on the frequency distribution of individual contact counts. Both plans effectively reduced the frequency distribution of high contact counts. The peak of the high contact range existing in the original result moves to the left side in the condition of flow control (yellow arrow in Fig. 25). This indicates that the intersection between contouring passenger flows is not a decisive factor leading to a high risk of direct contact in the morning, but it increases the cumulative contact degree of individuals. A comparison between the two plans showed that plan 1 worked better in increasing the proportion of agents in the low-risk range (30–60 counts). The effect
of plan 2 in the high-risk range (360–480 counts) was also more obvious. As shown in Fig. 26, the influence of the proposed plans in mitigating the intermediate contact was also obvious (yellow arrow in Fig. 26). The proposed plans showed less of an effect in the low-risk range. The comparison between Figs. 25 and 26 showed that although plan 1 increased the proportion of agents in the low-risk range to 20% (dashed orange rectangle in Fig. 25), the proportion of intermediate contact was still less than 4% (dashed orange rectangle in Fig. 26). This means that the flow control plans only make people in this range perceive fewer track contacts; their actual intermediate contact degree is not effectively improved.

Figs. 27 and 28 show the relationship between the travel distance and direct/indirect contact count in different control plans. The slopes of the trend lines of the samples in controlled flow have a downward trend, which shows an effective role in mitigating the contact risk of samples with long travel distances. Besides that, the dispersion degree of the sample distribution increased in the proposed plans. This means that the effect of the proposed plans on the equal-distance agents was different. These differences were considered by the periodic generation pattern of passenger flow from the metro station.

Fig. 29 (a) shows the effect of the proposed plans in increasing the exposure duration between individuals. The frequencies of long duration range (>2500 s) significantly increased in the controlled plans (Fig. 29(a)). Comparing the direction-specific results, we found that although the proposed plan could significantly improve the contact duration of moving in opposite directions (Fig. 29(c)), it significantly increased the contact duration between individuals moving in the same direction (Fig. 29(b)). In general, the proposed plans could not play a role in reducing the exposure duration.

Fig. 30 reflects the relationship between the DCCs and CED of individuals under different flow control conditions. We found that the cumulative exposure duration of samples with the number of direct contacts around 200 times increased significantly in the proposed plan (Fig. 30(b) and (c)). Comparing the result in Fig. 25, these samples were considered to correspond with the agents distributed in the peak ranges of plan 1 and plan 2. This indicates that the passengers who received less direct contact in controlled flows may be exposed for a longer duration to others in the impact area.

Fig. 31 shows the relationship between individual travel distance and CED in the proposed plans. The exposure duration of samples with 500–700 m distance (dashed red rectangle in Fig. 31) increased significantly by the proposed plans. In summary, the flow direction control measures reduced the contact risk between individuals in different directions, while they also increased the intermediate contact duration by converging the flows to the same direction.

5.2. Average arrival schedule

5.2.1. Simulation setting and result

This simulation experiment aimed to examine the influence of periodic passenger flows on the contact degree of individuals and the system. In a random generation mode, all agents’ generation time was randomly assigned. Table 4 shows the simulation results in the random generation mode. The reduction in the number of direct contact and
intermedia contact was not significant (1.6 % and 5 %, respectively). The cumulative exposure duration of the system was reduced by 13 %. Hence, equalizing the agent generation velocity plays a positive role in reducing the contact risk in the study area.

As shown in Fig. 32, the amplitude fluctuations in the curve significantly decrease in the random generation mode. The period of direct contact fluctuations grows significantly after averaging the passengers’ arrival time, while the curve converges to the median region of fluctuations in the schedule mode. This indicates that agglomerative passenger flows from metro stations do not significantly change the total system contact degree but do increase the uncertainty in the time of direct contact occurrence.

5.2.2. Frequency distribution analysis

Fig. 33 shows that the frequency distribution of the individual DCC did not change obviously in the average arrival condition. Averaging the generation time of passengers from metro stations played a positive role in reducing the exposure duration of long-duration population (>2000 s) in Fig. 34(a). Comparison between the direction-specific results showed that the random generation mode mainly reduced the duration of individuals exposed to the impact area of others in the same direction (Fig. 34(b)). Its effect on the exposure in the opposing direction was not obvious (Fig. 34(c)).

Fig. 35 shows the CED of agents with different DCCs in the two generation modes. The random mode significantly reduced the cumulative exposure duration of the agents with 300 direct contacts (orange arrow in Fig. 35(b)), corresponding to the peak ranges in Fig. 33. The downward shift of the samples near the y-axis area (green arrow in Fig. 35(b)) indicates that the random mode reduced the exposure risk of the population with less direct contact. In summary, although the
random generation mode could not reduce the inter-individual contact perceived by passengers, it played a potential role in mitigating the exposure duration between individuals.

Fig. 36 shows the CED difference of agents with different travel distances in the two generation modes. The CED of transfer passenger flow (around 600 m) significantly improved in the random generation mode (orange arrow in Fig. 36 (b)). This also proved the influence of the concentrated flow of passengers from the metro station on the long exposure duration in the MSA in the morning.

5.2.3. Spatial distribution result of intermediate contact

As the effect of random generation mode on direct contact was not obvious, we mainly investigated its influence on the spatial distribution of CED in the study area. As shown in Fig. 37(b), the exposure duration in front of the TJM station during the morning hour effectively improved (dotted circles in Fig. 37). For the surrounding areas of the TJ station and the NF station, the improvement was not obvious. In view of the generation velocities in Fig. C.1, the average arrival measure worked better in reducing the exposure risk around the terminal station without frequent train arrivals. The influence of concentrated and periodic passenger flows on exposure risk around the metro station was also verified.

6. Conclusion

This study estimated the spatiotemporal distribution of direct and intermediate contact in the IPTS connecting several metro stations based on observation data. We gathered pedestrians’ STPs in the network by using numerical models and a simulation approach. Contact count and exposure duration between individuals were calculated by space–time vector analysis. Through the quantification of cumulative contact count...
and duration, the spatiotemporal distribution of contact risk in the network was visualized. Based on the results of estimation and application, the following conclusions were drawn:

1. The contact degrees of individuals and the system are time varying. Both the direct contact degree at a specific link and global degree of IPTS are not determined by the total individual number, but rather by the OD distribution. The spatial distribution tendency of intermediate contact duration at links is significantly correlated with the distribution of the direct contact event.

2. Passengers are exposed to more individuals through indirect means than they are aware of. Transfer passengers between stations have a high possibility of contact with others without consciousness of it. Metro commuters destined to the local facilities in the MSA are the high-risk population for direct contact in the morning. Unperceived contact in the morning is not caused by the agents with seldom direct contact but rather by the frequent intermediate contact of the metro passengers.

3. The intersection between passenger flows in the opposite direction is not a decisive factor leading to high risk of direct contact in the morning, but it increases the contact degree of individuals. Directional flow control plays effective roles in mitigating the
direct and indirect contact counts of individuals with high contact risk. Flow control makes low-risk individuals perceive fewer track contacts. The individuals with less direct contact in controlled flows might be exposed for a longer duration to others in the impact region.

Fig. 36. Cumulative exposure duration difference of agents with different travel distances in two generation modes: (a) scheduled mode; (b) random mode.

Fig. 37. Spatial distribution result of cumulative exposure duration: (a) scheduled mode; (b) random mode.

Fig. B.1. Route choice utility for pedestrians.
(4) Flow control plans enhance the exposure duration degree of the system. However, there is an effective role in mitigating the contact risk of samples with long travel distances. The influence of different directional plans on spatial distribution of risk region is not consistent.

(5) The concentrated flow of passengers from the metro station causes the long exposure duration in the MSA in the morning. Equalizing the passenger generation plays a positive role in reducing the duration of individuals exposed to the impact area of others in the same direction. Although average generation cannot reduce the direct inter-individual contact perceived by passengers, it plays a potential role in mitigating the exposure duration between individuals. The average arrival measure works better in reducing the exposure risk around the terminal station without frequent train arrivals.

The main aim of the proposed framework is to examine the effect of the passenger dynamics on the inter-individual contact in the system. The contribution of this framework is its recreation of interaction between individuals in a complex system with limited actual observational data, which provides quick feedback on the performance of different intervention scenarios in transmission mitigation. The simulation results do not accurately correspond to the actual situation, which is a shortcoming compared to digital contact tracing. The non-medical interventions proposed in the application aimed to demonstrate the impact of passenger flow factors. Actual policy development should more comprehensively consider the impact of decisions in terms of business dynamics and individual travel efficiency.

For computational purposes, two simple contact models were proposed in this paper to describe the roles of pathway crossings and breathing dissemination in individual exposure. The proposed models could not be used directly to predict the transmission probability because the transmission of viruses requires consideration of the influence of complex environmental factors (e.g., air flow and diffusion of respiratory fugitives). This was the limitation of this study. In addition, this study did not consider the individual variation in infectiousness and susceptibility (He et al., 2020). The effect of passengers’ infectious period and virus shedding amount are expected to be studied in future work. In this paper, the contact risk degree was proposed as an intermediate indicator validating the impact of the spatiotemporal dynamics of passenger clusters on transmission. This intermediate indicator is expected to provide a reference for contact risk assessment and policy decision.

The duration of personal breathing dissemination is set based on assumption. This study had a limitation in the setting of the duration because the droplets are still infectious after evaporating into droplet nuclei. The duration estimation of the pedestrians exposed to the impact region of the droplet nucleus is more complex, requiring the consideration of physical environmental factors such as temperature, air flow, and objects in two-dimensional space. These factors were ignored in this study because of the large calculation load. Moreover, the effect of dissemination duration on the correlation between contact indicators at the system, link, and individual levels were not included in this study. These issues will be further studied in the next paper.

Funding

This study was supported by the National Natural Science Foundation of China (Grant No. 51808094).

CRediT authorship contribution statement

Zongchao Gu: Conceptualization, Data curation, Software, Visualization, Writing – original draft, Project administration, Funding acquisition. Toshihiro Osaragi: Conceptualization, Investigation, Data curation, Writing – review & editing. Wei Lu: Conceptualization, Writing – review & editing.

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.

Data availability

The authors do not have permission to share data.

Acknowledgment

The pedestrian traffic survey data used in this paper were obtained from by Fukuoka City Government.

Appendix A. Abbreviations
Appendix B. Route choice model

This study adopted a discrete dynamic route choice model (Gu and Osaragi, 2017) to determine passengers’ route choice tendency in IPTS network. This model defines the movement through a link/entrance/exit in a specific direction as the “state.” For a pedestrian who is destined for destination $j$ in Fig. B.1, the utility of choosing alternative state $b$ at the current state $a$, and $U_{ab}$, is composed of two parts:

$$U_{ab} = u_{ab} + \beta E_{ab}$$

(b1)

where $u_{ab}$ is the utility of the transition process from state $a$ to state $b$, and $E_{ab}$ is the expected utility of subsequent transition processes to the destination stage $j$. $\beta$ here is the discount parameter.

As shown in Fig. B.1, $E_{bj}$ is defined by the maximum expected utility of subsequent transition processes until the destination $j$. According to previous research (Itoh and Hato, 2013), $E_{bj}$ is denoted by the value function of state $b$ in a dynamic programming problem. The value function of state $b$ of the optimal solution is given by:

$$V_b(u,j) = E_{bj} = \max_{d \in D_b} \{u_{ad} + \beta V_d(u,j)\}$$

(b2)

where $d \in D_b$ is a set of alternative states available at state $b$. $u$ means the set of transition utilities in the network. Value function $V_b(u,j)$ stands for the maximum expected utility of subsequent transition processes from state $b$ to destination $j$. The value function $V_b(u,j)$ can be obtained by estimating the unique fixed point of the Bellman function in dynamic programming (Itoh and Hato, 2013). As Eq. (b2) is represented by a recursive function, $V_b(u,j)$ is decided by the set of transition utilities, $u$, in the network. The route choice utility, $U_{ab}^f$, is:

$$U_{ab}^f = u_{ab} + \beta V_b(u,j)$$

(b3)

Based on random utility theory, the transition probability from state $a$ to state $b$ can be obtained using a logit model:

$$P_{ab}(u,j) = \frac{\exp(u_{ab} + \beta V_b(u,j))}{\sum_{d \in D_b} \exp(u_{ad} + \beta V_d(u,j))}$$

(b4)

where $d \in D_b$ is a set of alternative states available at state $a$. Both the route choice utility, $U_{ab}^f$, and value function $V_b(u,j)$ are destination-specific values. $P_{ab}(u,j)$ is determined by the set of transition utilities, $u$, in the network. Based on observed link flow count data and estimated OD matrix, $u$ can be calibrated by minimizing the difference between estimated flow counts and observed link flow counts. The detailed calibration approach and results were reported in another published study (Gu and Osaragi, 2017).

Appendix C. Passenger generation time assignment

This study assumed that the passenger generation velocities after a train arrival satisfy a Poisson distribution (Gu et al., 2019). Using a 10-s interval, we first collected the number of passenger generations in the time duration $K$ (240 s) after a train’s arrival. We translated the observed data into the probability mass rate and employed Poisson regression analysis to calibrate the unknown parameter ($\lambda = 8.25$). The cumulative distribution function of the passenger generation after a single train arrival was obtained by:

$$C_k = \exp(-\lambda) \sum_{l=0}^{k} \frac{\lambda^l}{l!}, \forall k \in K$$

(c1)

As the simulation operated with 1 s intervals, we estimated the passenger generation velocity with a time scale of one-second using the spline interpolation method. The details of calibration approach are given in the published paper (Gu et al., 2019).

This study assumed that the train arrivals at the same station carries the same number of passengers. We could obtain the proportion of agents at each second after a single train arrival accordingly. For each station, the time distribution of passengers’ proportion could be obtained by the summation of all train arrivals during the simulation interval (Fig. C.1). For each agent from the metro station, its generation time was randomly assigned according to the time distribution of the generation proportion.
References:

Ahn, Y., Tsukaguchi, H., 2015. A method to estimate pedestrian flow in large transportation terminal. JST J. Traffic Eng. 1 (6), 1–9.

Asada, N., Osaragi, T., 2000. Avoidance behavior of walkers and bicycles. Papers Proc. Geogr. Inform. Syst. Assoc. 9, 281–284.

Chen, P., Zhang, D., Liu, J., Jian, I.Y., 2022. Assessing personal exposure to COVID-19 viral shedding and transmissibility of COVID-19. Nat Med 26 (5), 672–675.

Gu, Z., Osaragi, T., 2004. Modeling of Pedestrian Behavior and Its Applications to Spatial–Temporal Distribution of Pedestrians in Large Underground Space based on OD matrix and flow count data. J. Archit. Plann. (Transactions of AIJ) 76 (635), 137–155.

Huang, B., Wang, J., Cai, J., Yao, S., Chen, F., Tan, T., Hong, Y.Y., Ruktanonchai, C., Caroli, A., Floyd, J., Ruktanonchai, N., Yang, W., Li, Z., Tatem, A., Lai, S., 2021. Integrated vaccination and physical distancing interventions to prevent future COVID-19 waves in Chinese cities. Nat. Hum. Behav. 5, 1–11.

Itoh, S., Hato, E., 2013. Absorbing markov chain traffic flow assignment using a dynamic architecture. Transp. Res. Part A Policy Pract. 49, 966–982.

Johansson, A., Batty, M., Hayashi, K., Al Bar, O., Marcozzi, D., Memish, Z.A., 2012. Integrating vaccination and physical distancing interventions to prevent future COVID-19 waves in Chinese cities. Nat. Hum. Behav. 5, 1–11.

Liu, K., Yin, L., Ma, Z., Zhang, F., Zhao, J., 2020. Investigating physical encounters of individuals in urban metro systems with large-scale smart card data. Physica A 545, 123998. https://doi.org/10.1016/j.physa.2019.123998.

Lu, K., Han, B., 2016. Congestion Risk Evaluation and Precaution of Passenger Flow in Metro Stations. Open Civ. Eng. J. 10 (1), 93–104.

Liu, X., Yang, L., Yang, K., Gao, Z., Zhou, H., Meng, F., Qi, J., 2022. A distributionally robust optimization method for pedestrian flow control strategy and train scheduling on an urban rail transit line. Engineering 12, 202–220.

Marr, A.D., Sun, L., Corman, F., 2022. The impact of COVID-19 pandemic on public transport usage and route choice: Evidences from a long-term tracking study in urban area. Transp. Policy 116, 258–266.

Müttel, C.M., Scheiner, J., 2021. Investigating spatio-temporal mobility patterns and changes in metro usage under the impact of COVID-19 using Taipei Metro smart card data. Public Trans. 14 (2), 345–366.

Nishimura, K., Osaragi, T., 2009. Spatio-temporal Distribution of Pedestrians in District around Railroad Terminal Station, Summaries of technical papers of Annual Meeting Architectural Institute of Japan, Architectural Institute of Japan, 721–722.

Otsuki, T., 2020. Modeling of Pedestrian Behavior and Its Applications to Spatial Evaluation, IEEE Conference, Autonomous Agents & Multi Agent Systems (AAMAS’04), 2, pp. 834–841.

Zhao, H., Thrash, T., Schlafli, F., Kapadia, M., Aguilar, L., Helbing, D., Holscher, C., Zacharias, J., 2014. Underground pedestrian trips combats COVID-19 in China. J. Infect. 82 (2021), 159–198.

Zhao, H., Thrash, T., Kapadia, M., Wolff, K., Holscher, C., Helbing, D., Schinazi, V.R., Zacharias, J., 2020. Assessing crowd management strategies for the 2010 Love Parade disaster in Germany using a spatio-temporal simulation model. Phys. Stat. Mech. Appl. 274 (1–2), 150–156.

Yu, H., 2021. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. Science 371 (6562), 675–679.

Zacharias, J., 2000. Modeling Pedestrian Dynamics in Montreal – A new perspective for environmental behavior research. Time Geography 2001, 50–55.

Zhalong, E., 2018. Personalized ventilation: One possible solution for airborne infection control in highly occupied space? Indoor Built Environ. 27 (7), 873–876.

Zhao, Y., Yang, M., Zh, H., Yang, Z., H., Zhang, L., Glader, S., 2021. Modeling indoor-level non-pharmaceutical interventions during the COVID-19 pandemic: A pedestrian dynamics-based microscopic simulation approach. Transp. Policy 109, 12–23.

Xu, Y., Chen, X., 2021. Quantitative analysis of spatial vitality and spatial characteristics of underground urban space (US) in metro area. Tunn. Undergr. Space Technol. 111, 103875. https://doi.org/10.1016/j.tust.2021.103875.

Xu, C.W., Liu, L., 2018. Personalized ventilation: One possible solution for airborne infection control in highly occupied space? Indoor Built Environ. 27 (7), 873–876.