Electronic Research Archive

ERA, 32(7): 4472–4494.
DOI: 10.3934/era.2024202
Received: 03 February 2024
Revised: 16 May 2024
Accepted: 03 June 2024
Published: 16 July 2024

http://www.aimspress.com/journal/ERA

Research article

Incorporating mobile phone data-based travel mobility analysis of metro ridership in aboveground and underground layers

Jiping Xing[1], Xiaohong Jiang[1,*], Yu Yuan[2] and Wei Liu[3,4]

[1] College of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing 210037, China
[2] School of Transportation, Southeast University, Nanjing 211189, China
[3] WuHan INTEST Electronic Technology Co., LTD, Wuhan 430010, China
[4] School of Management, Hefei University of Technology, Hefei 230009, China

* **Correspondence:** Email: jiangxh@njfu.edu.cn.

Abstract: Metro transit is the core of urban transportation, and the mobility analysis of metro ridership can contribute to enhance the overall service level of the metro transit. Researchers studying metro ridership are focused on the spatiotemporal distribution characteristics of the ridership in the underground system of metro station by metro smart card data. However, limited by lack of travel mobility chain of ridership integrity, their activity patterns cannot be used to identify the heterogeneity of metro ridership's origin and transfer travel mode. In our research, we applied full spatiotemporal coverage of mobile phone data to identify the complete travel mobility of metro ridership in the perspective of ground and underground transit. First, the mobility of the boarding and alighting stations was extracted and the order of the transfer station was then extracted. Second, relying on the ridership flow identification method, the aboveground origin and destination of the ridership outside the metro system were extracted, and their transferred traffic mode was identified. The empirical results have shown that our proposed framework can accurately analyze the mobility patterns of metro ridership in an aboveground area and underground station.

Keywords: mobile phone data; metro ridership flow; transfer traffic mode identification; travel mobility; aboveground ridership flow

1. Introduction

Metro transit acts as a vital mode of public transit travel that can efficiently undertake most long-distance trips in metropolitan areas [1,2]. It has been widely accepted by commuters and has effectively reduced traffic congestion [3,4]. However, compared to other public traffic modes (such as buses and taxis), the expensive underground construction budget of metro restricts the number of setting of metro stations, which are deployed only in high demand locations [5,6]. This sparse deployment of metro station results in ridership usually having a long distance to travel or transfer before boarding the metro [7,8]. This size of connection distance between leaving the house and taking the metro can directly affect the satisfaction and service level of metro ridership. Therefore, mobility of metro ridership can not only analysis the underground trips of metro mobility, both aboveground and underground mobility are critical in affecting the service level of metro ridership.

The research on the mobility of metro ridership can be divided into ridership recognition and prediction [7,9], ridership choice behavior analysis [10,11], and analysis of influencing factors of the built environment surrounding metro station [12,13]. Furthermore, the research dataset in these studies mostly includes metro smart card data and questionnaire data. For example, Luo, et al. [14] analyzed spatiotemporal variation of ridership and commuting mobility of residents. Chen, et al. [15] proposed the geographically weighted regression (GWR) model to address the spatial autocorrelation and nonstationarity of metro ridership. However, these applied metro smart card data are missing the description of over-ground travel behavior, leading to the inability of a complete travel trajectory of metro ridership to be recorded. On the other hand, Zhang, et al. [16] fused questionnaire survey data and metro smart card data to infer metro ridership' transfer behavior. However, the representation of questionnaires are usually doubtful and concentrated near metro stations, particularly when residents live at a certain distance from the metro stations; thus, these tasks cannot be completed well.

With the emerging technologies on mobile phone data in the field of traffic information and with the benefit of full spatiotemporal coverage, it can adequately reflect the market penetration rate of travelers [17,18]. When mobile phone data is applied for the metro scenario, the mobile base station on the aboveground and underground of metro network can be technically distinguished, and the operation of base stations in underground layers will not be affected by the aboveground layer of the base station [19–22]. Thus, it can individually record the travel trajectory of metro ridership in the internal and external of metro networks. By matching the ID (identity document) of the travelers, the origin and destination of the ridership can be inferred with the difference of underground and aboveground base station. Compared with the traditional application of metro smart data, mobile phone data can be used to analyze the mobility of ridership with better performance [23,24]. For example, it can be used for evaluating the level of service of metro stations at different distances of the surrounding built environment and the transfer capacity of metro stations [25–29].

Motivated by the aforementioned research gap, we present a framework for analyzing the mobility of metro ridership traveling aboveground and underground by mobile phone data. We first identified the metro ridership boarding stations, alighting stations, and transfer stations. Then, we analyzed the boarding ridership at the same metro station from different distances and determined whether they had made transfers to other traffic modes. In summary, our innovative contributions to this work are the following:

1) To overcome the lack of mobility of metro ridership on the aboveground layer, we utilized mobile phone data to simultaneously profile metro ridership aboveground and underground travel

mobility and to effectively identify the metro ridership underground transfer station and the transfers to other aboveground traffic modes.

2) We not only identified riderships entering and exiting metro stations but also estimated riderships from different origin distances under the same metro station, thus indirectly assessing the service level of metro riderships at different metro stations.

The remainder of this paper is organized as follows. The literature review is provided in Section 2. In Section 3, we present the methodology. A case study of a real-world metro is presented to verify the effectiveness of the methodology in Section 4. The results are also presented in this section. Finally, in Section 5, we conclude our study and provide recommendations for future research.

2. Literature review

During the analysis of internal and external of metro ridership flows, we classified studies into the following two parts: Metro travel activity patterns and the application of mobile phone data in metro transit trips.

2.1. The studies in metro activity pattern recognition

With the application of metro smart card data, the spatiotemporal features of ridership in the metro transit system can be analyzedResearchers mainly focused on the activity patterns of riderships in the metro transit system. For example, Sun et al. [30] analyzed the smart card data of the Shanghai metro transit and obtained the spatial distribution of the ridership and the uneven characteristics in and out of the metro station. Moreover, the reliability of the travel time of the metro transit system has also been evaluated in the ridership experience from a multimodal travel perspective. For instance, analyzing the features of metro ridership flow, policy recommendations can also be provided for transportation departments. Chen and Wei [31] analyzed the characteristics of morning and evening peaks and the daily variable characteristics of ridership flow. On this basis, Luo et al. [14] analyzed the hourly peak ridership flow in the Nanjing metro system at different stations and then identified the hot zones of ridership flow and the temporal characteristics of ridership flow. In summary, these studies provide support for the construction of a ridership flow prediction model by exploring the temporal characteristics of the short-term ridership flow.

Furthermore, Kim et al. [32] identified the location of residence and work riderships. Considering the socioeconomic attributes of ridership, commuting distance, commuting time, and other characteristics, they modeled the ridership' destination, and these results can be well-explained. Some researchers also combined the land use around the stations to analyze the characteristics of ridership. For example, Liu et al. [33] analyzed the smart card data of Shenzhen metro transit by surrounding POI data and obtained the precise coverage of ridership during the Spring Festival. Similarly, Shi et al. [34] built a geographical time-weighted regression model and quantitatively analyzed the correlation between the ridership flow and the characteristics of land use and network topology.

The analysis of the metro ridership flow can also be used to provide a theoretical and precise data support for a metro timetable schedule. For example, Jungang and Yang [35] built an optimization of the integer programming model for the scheduling lines with metro ridership flow and combined local optimization with the CPLEX optimization studio solver to obtain results. This showed that this approach can effectively reduce the overall ridership waiting time and reduce the potential risk caused by crowd gathering. Based on the goal of optimizing operating costs and the number of waiting passengers, Liu et al. [36] built a Lagrangian relaxation optimization model to disassemble and solve the problem. Li et al. [37] considered the departure time of the metro and the full load rate of the metro at each station, and numerical experiments showed that this method can better optimize the stability of the hour distance of the time headway.

In the study of the transfer behavior of metro ridership, based on the path-based logit model, considering several factors such as travel time, transfer times, and transfer time, Jin et al. [38] quantified the influence of different transfer times on the different path utility perceived by ridership. Considering the attitudes of different people towards risk and integrating smart card data, Zhang et al. [16] established a path selection model based on a nested logit model. Assuming that there is an emergency at the track site, Wang et al. [39] used questionnaire data to analyze the travel preference of ridership when a line is delayed. Furthermore, Cascajo et al. [40] studied the influence of in-vehicle time, transfer distance, waiting time, and crowd density on the transfer behavior of ridership.

There are studies based on machine learning models for modeling metro ridership. Based on Bayesian network optimization, Sun et al. [41] obtained the travel time of the metro transit network and then constructed the estimation of travel path in the metro transit network. Considering the ridership transfer or without transfer Zhao et al. [42] built a probability-based model to analyze the user's path routing behavior in Shenzhen. Wu et al. [43] clustered the smart card data of the same OD pair by the travel time and considered the uncertainty of the transfer time and distance. Some researchers focused on the mobility of metro ridership by traffic simulation models. For instance, Zhang and Han [44] studied the transfer behavior in the metro transit system and obtained important factors in transfer distance and time. Cheng and Tseng [45] first divided ridership based on the information theory and analyzed the willingness of different groups of passengers to transfer. The analysis results showed that the transfer cost, the perceived value of ridership, and the transfer bus significantly influence the transfer willingness of ridership. Considering the inter-station travel time in the temporal perspective, Yang et al. [46] built a bi-level integer optimization model, and the optimization goal is to minimize the expected travel time and the extra time for transfer.

2.2. The studies of mobile phone data on metro mobility

Several studies have proved that the mobility of ridership can be effectively restored by mobile phone data. For example, Schlaich et al. [47] studied the method of converting raw data into location sequences by mobile phone location data. Based on a framework of interpolation methods, Hoteit et al. [48] assessed the difference between the mobility obtained using mobile phone data and the real mobility of ridership. Considering the low market penetration rate of mobile phone data, Li et al. [49] built a data-driven mobility reconstruction model to effectively alleviate the low mobility reliability caused by the sparseness of signaling data.

Based on the travel mobility detection method, some researchers have analyzed the travel characteristics of residents. For instance, Steenbruggen et al. [50] identified the residence and workplace of ridership and analyzed the corresponding commuting time distribution using mobile phone data to verify the reliability of mobile phone data. Phithakkitnukoon [51] identified the daily activity patterns of residents and showed that travelers' whose work sites are close to each other have similar travel patterns. Jiang et al. [52] proposed a method that can convert mobile signaling phone data into interpretable residential space activity, which has a large advantage over the traditional survey method. Widhalm et al. [53] proposed a ridership staying behavior recognition method and identified the travel mode of ridership by integrating the stay location, time, and land type. Yuan and Martin [54] analyzed metro travel patterns of different regions and compared the similarities of different time series based on the DTW algorithm. Combining traditional resident travel survey data. Diao et al. [55] obtained fine travel characteristics of residents using mobile phone data.

Other researchers mainly combined special events to analyze the travel mobility by integrating mobile phone data. For example, Wang et al. [56] built an accurate train speed and position calibration model, which can provide a data basis for the fine operation of metro transit. With the analysis of the mobile phone data before and after an accident in Shanghai's metro transit, Duan et al. [57] extracted the spatiotemporal features in the evacuation of passengers and their choice of travel routes. Before travelers enter the metro transit, they will pass through the base stations over the metro station. If the information about the surrounding base stations can be extracted, it can be beneficial to the in-advance service for the ridership of the metro station [58–62]. Considering the full spaitotemporal coverage advantages of mobile phone data, researchers developed an integration framework between mobile phone data and smart card data.

In summary, for metro ridership flow, researchers usually focused on the interior of the metro transit system or have not analyzed the mobility of metro travelers from boarding stations to alighting stations [63–65]. However, for a ridership trip, we should not only analyze boarding and alighting stations but also analyze the whole mobility. Therefore, it is difficult to make a more comprehensive assessment of the service characteristics of the metro transit, such as the service scope of the metro transit system.

3. Methodology

3.1. Research framework

We focus on how to extract the mobility of metro ridership using mobile phone data. With the aid of dedicated communication base stations inside the metro, communicating with external base stations is limited by the influence of the thick soil layer. The ridership extraction can be performed by these underground signaling base stations. Furthermore, with the aid of the aboveground signaling base station, we can also identify the ridership on the aboveground. This framework is illustrated in Figure 1. Our research mostly includes the following steps:

1) The timestamp and base station location information when the mobile phone user communicates with the base station or change locations were collected, and missing and outlier data were preprocessed.

2) Identifying the internal metro boarding and alighting station, as well as the stations which ridership transfers to another metro line. The ridership volumes for each metro network were extracted and aggregated on different spatiotemporal coverage.

3) With the aid of the metro smart card data, the ridership volumes acquired by mobile phone data have been explored to verify its authenticity.

4) Recognition of the origin of metro transit ridership was performed in each aboveground metro station, and transfer transport modes for destination ridership were also estimated.

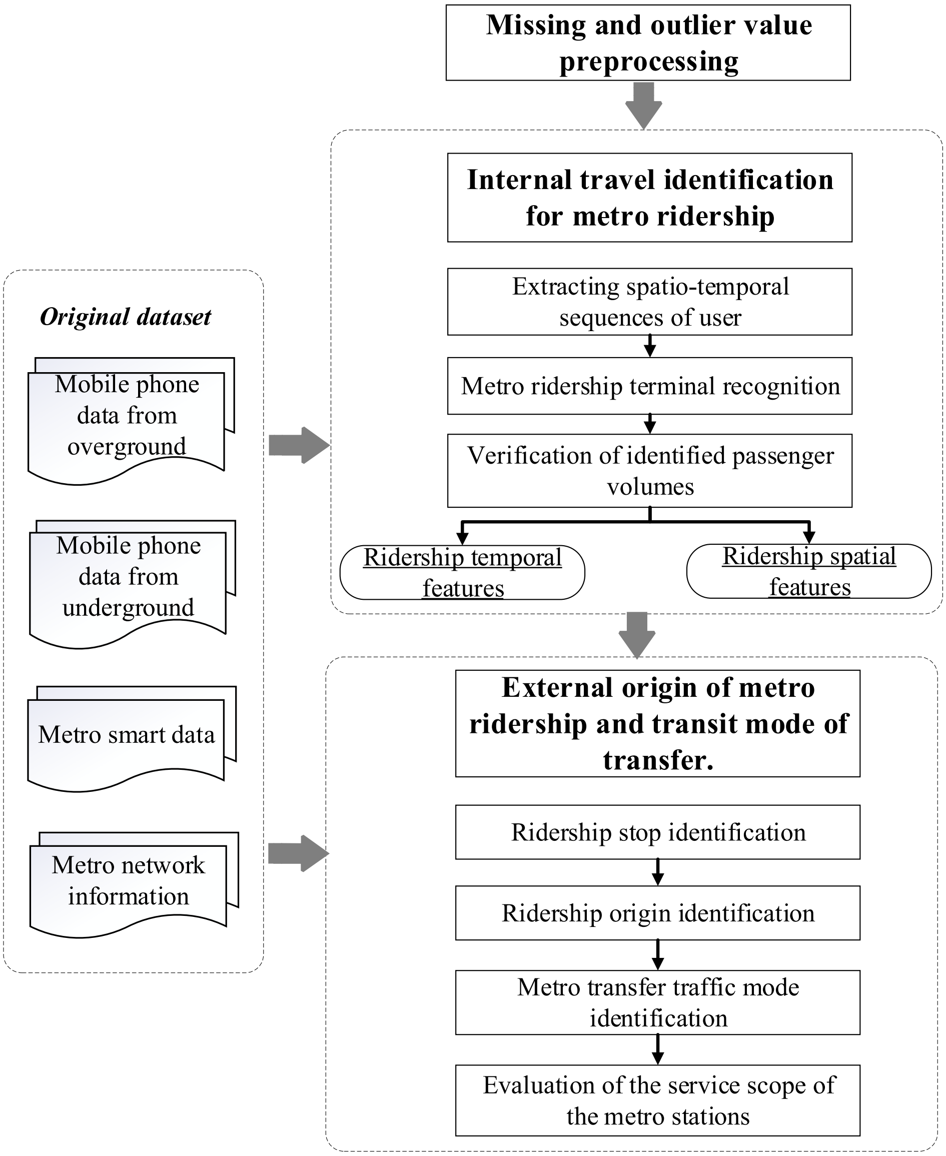


Figure 1. Flowchart of our study framework.

3.2. Identification of internal ridership boarding and alighting station

3.2.1. Problem definition

Mobile phone data belongs to discrete spatiotemproal data. When the mobile phone user communicates with the base station, the signaling position data will be generated. Note that the communication device will communicate with the base station when the mobile phone users make calls or send information to the base station regularly. Because of this, the position change information can be obtained. This sequence positioning base station can be used to represent the complete travel trajectory of metro ridership, as it is shown in Figure 2. The label of mobile phone data is listed in Table 1.

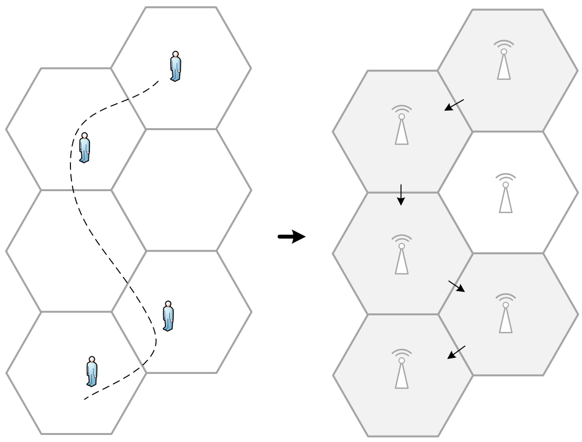


Figure 2. Metro ridership trajectory by mobile phone data.

Table 1. Label of mobile phone data.

Label	Description	Example
MDN	Phone number	156****6109
LAC	Base station location code	DFGDEC
CELLID	Base station label	50

As internal base stations are installed inside the metro line network, the positioning information of ridership can be identified by the signaling data in each base stations, which include metro line stations and signaling base stations. The former (railway base station) is installed in the intermediate area between each metro stations, and the latter (underground base station) is installed in each internal metro station. These are shown in Figure 3.

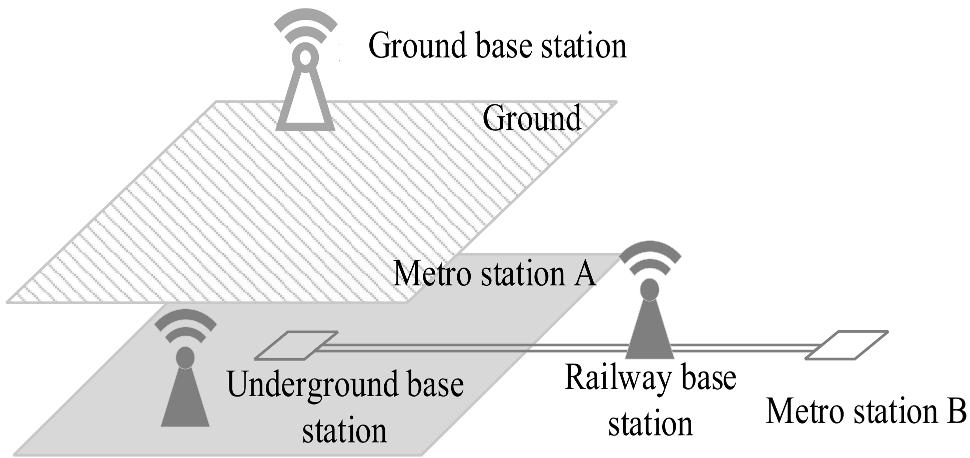


Figure 3. Types of base stations from aboveground and underground stations.

3.2.2. Reliability analysis of metro ridership by mobile phone data

The mobile phone data we applied was provided by China Telecom, which has almost one-third of the market penetration rate for mobile phone users. The identical distribution of ridership obtained by mobile phone data and that smart card data should be validated, which is necessary to be distributionally tested in advance. The ridership volume from two types of datasets is standardized in Eq (1):

$$x_{norm} = \frac{x - \bar{x}}{s} \tag{1}$$

where x is true value of sample dataset, $\bar{x}$ is the mean value, and s is the standard deviation.

SPSS was used to test the two-way variance analysis hypothesis of Friedman by rank to test the distribution of the two types of readerships. The results showed that the distribution of the two types was significantly the same, as shown in Table 2. Hence, it can be determined that ridership volume is obtained using mobile phone data recognition and that metro smart data are identically distributed. This proved the accuracy of the metro ridership identification by mobile phone data.

Table 2. Results of distribution examination.

Zero hypothesis	Significance	Decision
The ridership volume distribution of stations obtained by two types of data sources is the same.	0.254	Zero hypotheses were retained.
The ridership volume volume of the station obtained by two types of data sources is the same.	0.917	Zero hypotheses were retained.
The ridership volume distribution of stations obtained by two types of data sources is the same.	0.178	Zero hypotheses were retained.
The ridership volume of station obtained by two types of data sources is the same.	0.756	Zero hypotheses were retained.

3.2.3. Identification of metro passengers boarding and alighting

With the applicaiton of mobile phone data, the sequence passengers passing metro stations can be obtained. The first and last metro stations of this sequence can be regarded as the boarding and alighting stations, respectively. Due to the possible offset of the mobile terminal in the process of communication with the base station, there may be deviations in the first and last records of the station sequence. Furthermore, limited by mobile drift and other reasons, when the first and last record of a traveler is the communication with the metro-inside based base station, the communication is affected by noise, causing problems to accurately judge whether the base station is upstream or downstream of base station. Hence, this modified method can be generated by the following assumption: Before entering the metro station, the mobile phone terminal held by the metro ridership will communicate with the base station within the radius of 150 meters in each metro-inside based base station. Based on the time label recorded at the origin and destination of each trip, the metro-inside based base station, in which the ridership is previously located, can be determined, the ID information of the station is located, and the correction of the boarding and alighting information can be obtained. The details of the algorithm can be listed as follows:

Algorithm 1. Recognition of boarding and alighting stations of ridership.

Input: metro-inside based base station ID No. list, original signaling data;
Output: the boarding and alighting metro stations of the passengers' trip.

1. Filter out the traveler mobility by the base station matching ID No. table:
Connect the distribution of metro base stations with mobile signaling data, and filter out the mobility of ridership appearing in the metro base station.
2. Elimination of erroneous data:
Remove mobile phone users with less than 2 metro base stations (including station halls and station tracks) all day.
3. Extract mobility information of ridership:
Extracting passengers' mobility period, station, and next metro station recording information; calculating the time slot between each positioning trajectory, to obtain the mobility which composed of adjacent stations in the signaling data sorted by user time.
4. Elimination of erroneous data:
Remove mobile phone users with less than 2 metro base stations (including station halls and station tracks) all day, or with travel time less than 5 min.
5. Filter out mobile phone user data of aboveground base station and obtain ridership volume.
Filter out mobile phone data by travel speeds from car users; select speed ≤ 15 km/h as the metro threshold value of speed, where speed can be calculated by the number of spacing station × 1000 /time, this research regard the distance between each station as 1000 m.
6. Add the ground base station around the stations to re-recognize the ridership volumes.
Associate the ridership volumes obtained from the metro station hall inside based base station with the base stations in the range of 500 m around the station.

3.2.4. Metro transfer identification

For the recognition of passenger transfer, the following constraints should be satisfied: Transfer behavior of ridership must have occurred at the transfer station, and the ridership appearing at the transfer station can be divided as follows:

Type 1: The ridership volume that transfer metro station.

Type 2: The passenger do not get off at that metro station.

Type 3: The passenger departure or destination station is the transfer station.

The three types of ridership have certain characteristics, and we designed recognition methods based on these characteristics. For Type 1 of ridership volume, its distinctive feature is that with the transfer station as the demarcation point, the stations passing by ridership before and after the transfer station are on different metro lines. For Type 2, the characteristic of ridership is on the same line with the transfer station as the demarcation point. For Type 3, the characteristics are relatively obvious. It is only necessary to judge whether the first and last stations of the station are transfer stations.

The signaling records that appear at the transfer station are filtered out, and then the boarding and alighting stations of each ridership are obtained and matched by their phone number. Finally, a sequence of stations is obtained, including the information on boarding and alighting stations and transfer information. This order can be listed as follows: Boarding station, transfer station 1, transfer station 2, transfer station 3, and alighting station. The details of the algorithm is listed in Algorithm 2.

Algorithm 2. Transfer metro station identification framework.
Input: the metro inside-based base station list table, original signaling data; **Output:** the boarding and alighting stations of the passenger's metro travel.
1. Extract the station inside-based hall base station and base station label information of the transfer station. The base station collection can be denoted as $\{T\}$. 2. Filter the signaling record of the base station in the collection $\{T\}$ from the original signaling table and sort it by the user's unique recognition number and time to obtain the sequence of the transfer station each user passes (the user may transfer or may not transfer in the station). 3. Match the information table obtained in step 2 with the user's boarding and alighting station obtained to obtain the sequence of each user's boarding and alighting station and the transfer stations passing by. 4. Infer the real line passed by the user and the transfer station sequence of the actual transfer occurred based on the line conversion sequence that may be realized at each transfer station.

3.3. Identification of the origin and destination of metro ridership volume

Traditional metro smart card data contains only the terminal information of ridership, which cannot record the mobility of ridership outside the metro system. With the application of mobile phone data, the origin and destination of passengers can be accessed and then the spatiotemporal characteristics of the passengers served by the metro transit system can be described. The stay sites can be recognized based on the time of the mobile phone user's stay at each base station and the time it takes to change their position between the base stations. The boarding and alighting stations of the ridership, as well as the corresponding time information, can be obtained. For the recognition of the user's origin station, the last stay site one hour before the time the user enters the metro can be considered as the origin station. For the recognition of the user's destination station, the first stay site within one hour, after the user leaves the metro, can be considered as the user's destination station.

This algorithm obtains the passenger's origin and destination information by considering the spatiotemporal correlation between the entry and departure time gap, as well as the site of residence. Hence, the accuracy of this algorithm is influenced by the accuracy of the recognition performance of the passenger's residence location. Since the ridership may have a short stopover at some sites, the time threshold in this process of stay site recognition should be adjusted by the traveling characteristics of different residents.

Algorithm 3. Recognition algorithm for the origin location of ridership.
Input: information on passenger's boarding and alighting station and stay sites. **Output:** passenger's origin and destination.
1. Obtain the passenger's departure and arrival time according to his boarding and alighting station information. 2. Take the passenger's nearest stop time within 2 hours before the departure time of the boarding station greater than or equal to 15 minutes as his origin location. 3. Take the passenger's first stop time greater than or equal to 15 minutes from the first stop within 2 hours after reaching the alighting station as his destination.

In terms of space, all stations have a certain scope of service. Therefore, intuitively, the origin and destination recognition results should theoretically be concentrated in a certain area around the station. According to this basic assumption, we put forward two indicators to judge the reliability of the recognition results. The symbols used in the indicators are shown in Table 3.

Table 3. Description of notations.

Symbol	Description
S_i	ith station
$B_{i,j,t}$	The jth origin base station of the ith station at the time of t
$S_{i,j}^{nearest}$	The ith station, the nearest station to the jth origin base station
$n_{i,j,t}$	The number of ridership in the jth source base station of the ith station at the time of t
$\{S_i\}$	Origin collection of the ith station

(i) Possible incorrect base stations

In theory, ridership will choose to enter the nearest metro station to take the train. Therefore, if the base station is the origin base station of the station i, the station i should theoretically be the nearest station to the base station j. However, due to the large coverage of the base station, its location is not accurate. Assuming that the nearest station to the base station j is actually k, the distance difference between the base station j and station i and k should be as small as possible. Therefore, the weighted average of the number of base stations with the distance difference is considered as an indicator.

$$\mathrm{criterion}_{1,i,t} = \frac{\sum_{j=0} \Delta s \times n_{i,j,t}}{\sum_{j=0} n_{i,j,t}} \tag{2}$$

where Δs is the absolute value of the distance between $B_{i,j,t}$ and the distance difference between $S_{i,j}^{nearest}$ and S_i.

(ii) Number of repeats of origin base stations of neighboring k stations

Theoretically, the collection of origin base stations of ridership at different stations should not intersect with each other (regardless of the base stations in the middle area of the two stations). Therefore, the reasonableness of the recognition results can be measured by the distance between the two stations apart from k stations.

$$criterion_{1,i,t}^{k} = \left|\{S_i\} \cap \{S_{i+k}\}\right| \tag{3}$$

where k can be positive or negative, representing the k th station downstream of the current station or the k th station upstream, respectively.

3.4. Identification transfer transit mode of metro ridership

Through the mobility information of the ridership' path base station obtained using mobile phone data, the non-linear coefficient of the user's path can be analyzed. Combined with time information, the ridership's movement speed, acceleration, and other characteristics in each mobility segment can be further calculated. According to these characteristics, based on the triangular positioning data derived from mobile phone data, we constructed a preliminary method for identifying the transfer mode of metro ridership. Considering the effect of the positioning error of the mobile phone signalling data, the recognition method of ridership transfer transit mode has been proposed as shown in Algorithm 4.

Algorithm 4. Recognition method of ridership transfer transit mode.

Input: Mobile phone positioning data, latitude, and longitude information of ridership origin and destination.

Output: Metro ridership transfer mode.

1. Filtering metro passengers from the mobile phone positioning data;
2. Extracting mobile phone positioning data before and after the passenger enters the metro station based on the ridership terminal information;
3. Obtaining the time point of departure and arrival of the ridership from the origin according to the information of the origin and destination;
4. Recording the timestamp extracted from the previous step as the boarding station of travel before entering the station or the alighting station;
5. Extracting passengers from 3 minutes to 30 minutes of transfer time of mobile phone data;
6. Calculating the driving speed and acceleration of each section of the travel mobility;
7. Determining the travel mode;

If the speed is less than 4 m/s, it is considered to be slow; If the speed is greater than 4 m/s and the acceleration is greater than 0.15 m/s^2, it is considered a car; If the acceleration is less than 0.15 m/s^2, it is considered to be a bus.

4. Case analysis

4.1. Research data and metro network

The study area of this research is the city of Suzhou, China, which is a prefecture-level city located in the southeast region of China. It is situated in the core area of the Yangtze River Delta region, connecting Shanghai to the Eastern provinces. Suzhou is a unique blend of industrialization and historical significance, thus offering its residents a diverse and vibrant lifestyle encompassing both work and leisure activities. This socioeconomic and demographic landscape makes Suzhou an intriguing area to examine the impact of human mobility patterns. Furthermore, as a developed city in China, the design of metro lines and metro stations was justified based on the mobility of existing travel modes such as public transport over ground. Thus, the metro lines are designed to cover most of the downtown area of Suzhou, and the metro stations have been designed to be located in critical areas. Our research data is from the Suzhou metro transit system. The Suzhou metro transit system has a total of 4 operating lines and 93 stations, including 5 transfer stations, which are shown in Figure 4.

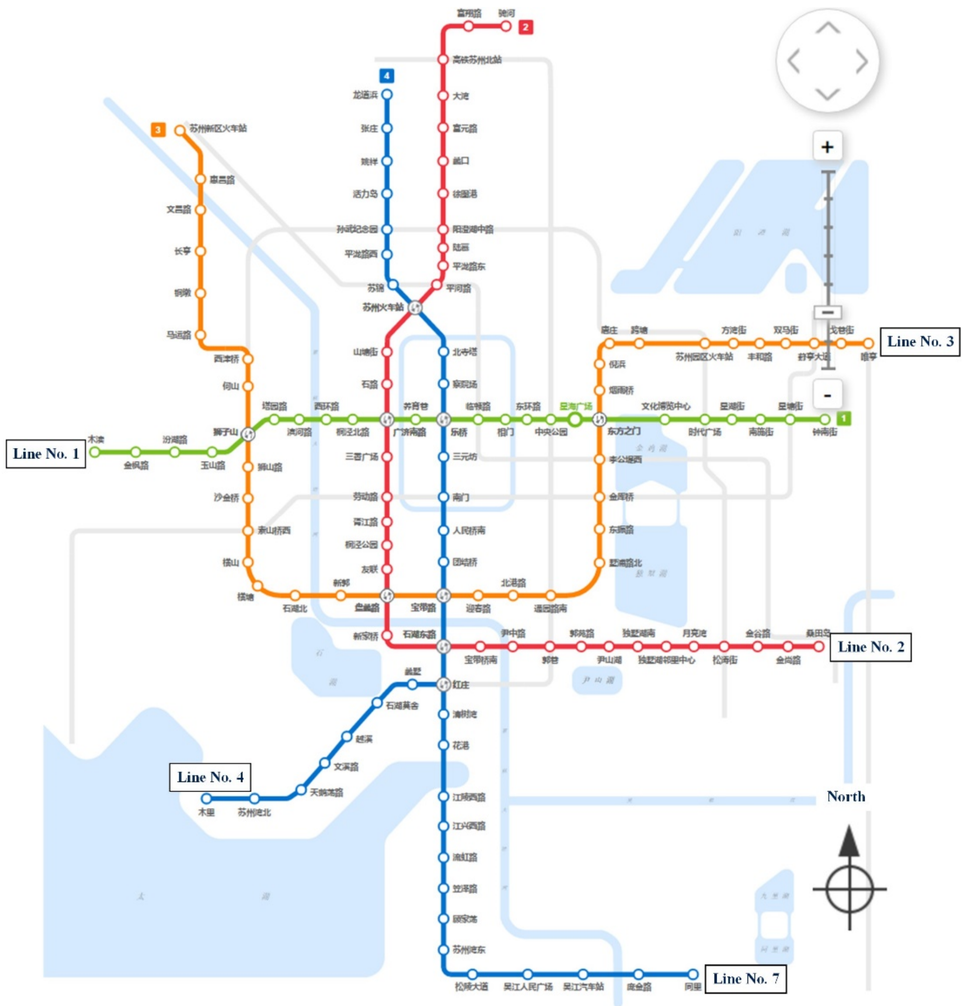


Figure 4. Metro line network in Suzhou, China.

By taking urban mobile phone data and metro smart card data as the major data sources, we used the mobile phone data and metro smart card data of Suzhou from August 15, 2018 to October 15, 2018, and site information and base station information. The market penetration of our selected mobile phone data was above 30%. In order to use these data for traffic volumes estimation, we applied the Time Difference of Arrival (TDOA) location technique to identify the location of each mobile phone signal transmission. The positioning accuracy of this technique was within 150 m and is not affected by buildings or other surfaces. Due to the influence of physical conditions and equipment mechanisms in the process of data collection and storage, there were some errors and missing data. Therefore, to ensure the effectiveness of the performance, we needed to preprocess these datasets. This process included the processing of drift data generated by the frequent switching of base stations in mobile phone data, the processing of numerical abnormal data, and the integration and aggregation of data.

Furthermore, the metro smart card data has recorded the timestamps of passengers' entrance and exit stations, and the corresponding metro station location information. Herein, the smart card data can be used to validate the traffic mode identification results based on mobile phone data.

4.2. Ridership volume analysis

As method in Section 3, the complete trip of metro ridership that includes the boarding to the alighting can be effectively identified using mobile phone data. Then, the station-based ridership volume in each metro line network, the transfer ridership volume, and the spatiotemporal characteristics of the origin of ridership can be extracted.

4.2.1. Ridership volume analysis in each metro line

As shown in Figure 5, hourly ridership volumes extracted by metro smart card data and mobile phone data are described for metro lines No.1, No.2, No.4, and No.7 in the solid line and the dotted line, respectively. Both are indicated by different coordinate axes. We found that the evolving trends in time series by these two types of datasets have similar fluctuations. We can conclude that the ridership extracted by applied mobile phone data can effectively reflect the real metro ridership trips.

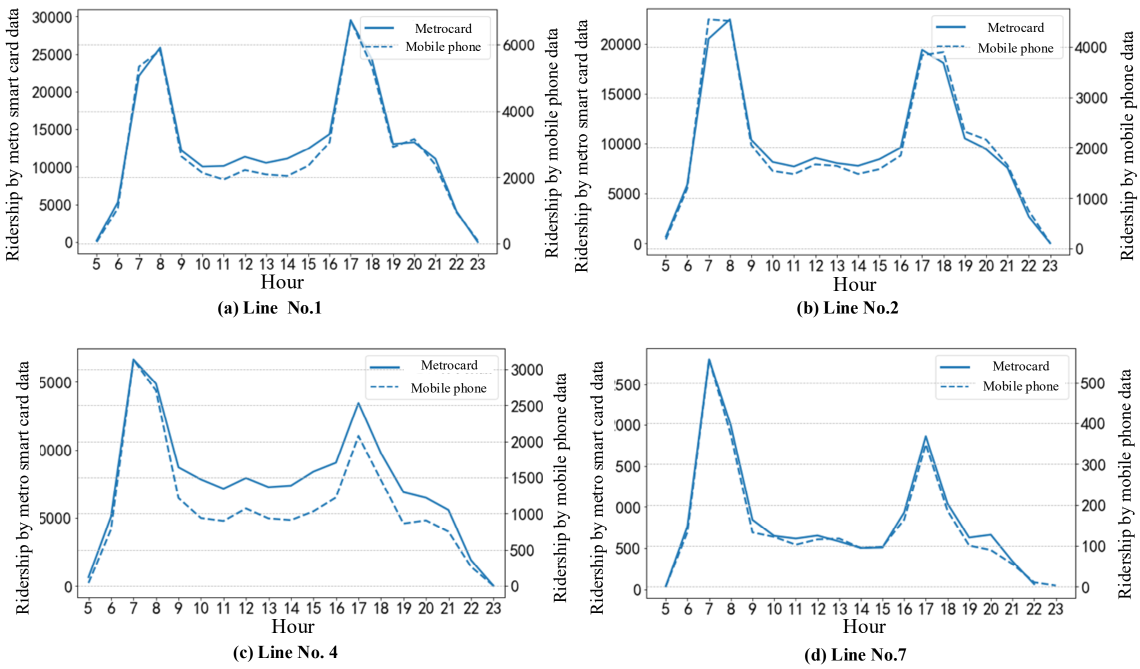
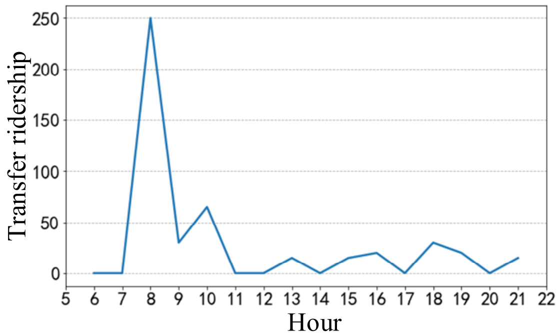


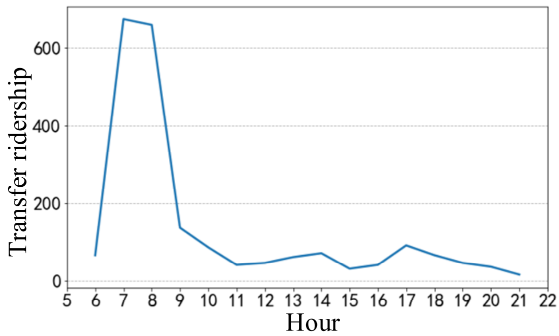
Figure 5. Hourly ridership volume statistics in each metro line in Suzhou.

Observing fluctuations of ridership volume, the ridership volume on Line No.1 is the highest of the four, and the following sequence is Line No.2, Line No.4, and Line No.7. Furthermore, each metro line has obvious fluctuations in the morning and evening peaks. The ridership volume of Line No.1 in the evening peak is slightly higher than that in the morning peak. Furthermore, the morning peak is slightly higher than the evening peak in other metro lines.

For transfer ridership volume, we take the transfer line from Line No.7 to Line No.1 and Line No.2 as our examples. As shown in Figure 6, there are two transfer metro lines, which include Line No.2, Line No.1. We found that most of the ridership volume from Line No.2 to Line No.1 occurs in the morning peak hour.



(a) Ridership volume on Line No. 1



(b) Ridership volume on Line No. 2

Figure 6. Transfer ridership from Line No.7 to Line No.1 and Line No.2.

4.2.2. Traffic mode preference from originating ridership at different distances

(i) The characteristics of origins and destinations

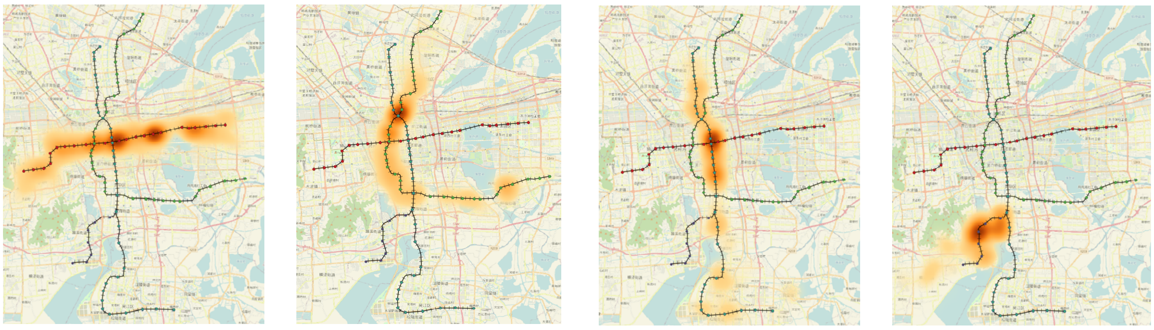
We recorded the location information of aboveground ridership via the geographic coordinates of mobile signaling base stations. Considering the effect of the density of mobile base station and the positioning error of mobile phone signaling data, we found that there is an error of almost one hundred meters between the actual location of the ridership and the recorded location information. However, as the metro generally serves long-distance trips, this error is in the tolerable range for metro ridership. As shown in Table 4, we aggregated all the origins and destinations of ridership at different distances: The average distance from the ridership's origin is 2142 m, and the average distance from the ridership's destination is 3625 m; the origin of 50% of ridership is less than 900 m, and the destination of 50% of ridership is less than 1200 m; and the origin of 75% of ridership is less than 1700 m, and the destination of 75% of ridership is less than 3800 m.

Table 4. Distance from different origins and destinations of ridership.

Statistical description	Origin distance (m)	Destination distance (m)
mean	2142	3625
25%	350	500
50%	900	1200
75%	1700	3800

(ii) Visualization of each metro line ridership

Figure 7 shows the spatial distribution of the origin of ridership volume in each metro line. The color from light to dark indicates that the ridership volumes in the metro station are from less to more. We can find that most of the origins are from fringe areas of the metro line network. Moreover, we found that for the line network, there are some hot spots of ridership volume, such as the transfer station in some metro lines.

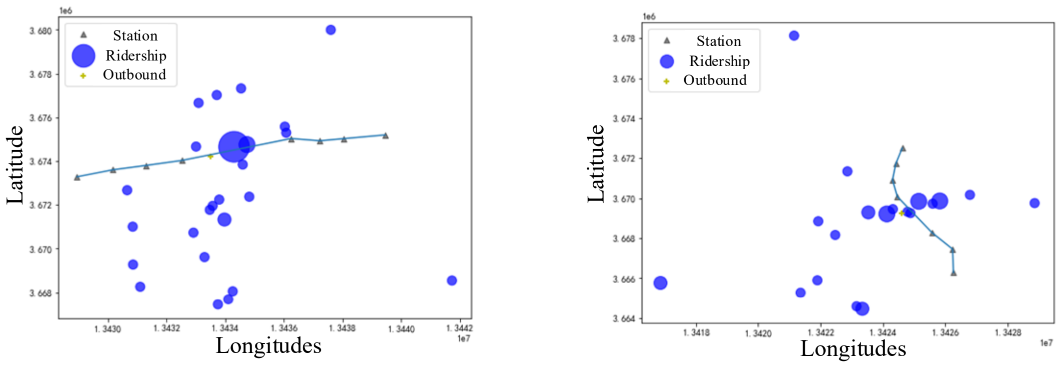


(a) Line No.1 (b) Line No.2 (c) Line No.4 (d) Line No.7

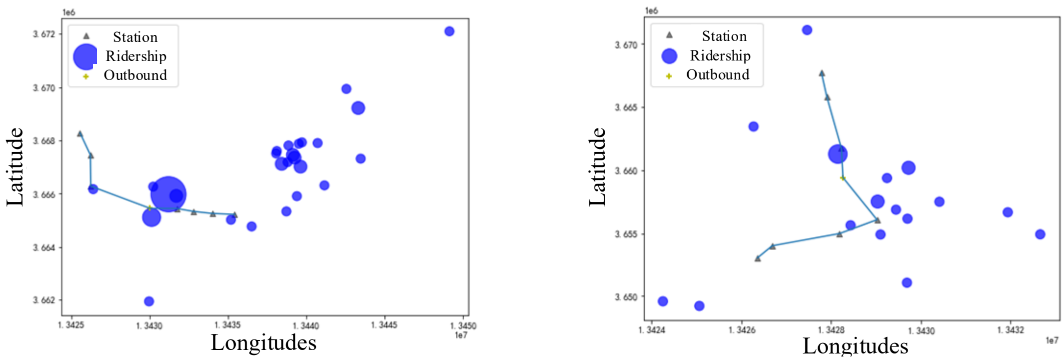
Figure 7. Visualization of the origins of different metro lines.

4.3. Identification and verification of ridership volume

4.3.1. Effectiveness analysis of aboveground ridership location



（a）Visualization of stations with less than 1000



（b）Visualization of stations with larger than 3000

Figure 8. Visualization of metro stations considering indicator I.

Affected by the density of base station deployment aboveground, we need to consider the impact when base stations are sparsely deployed and when the distance between neighboring metro stations is short. This can lead to the corresponding identification of ridership being identified as entering and exiting the metro station incorrectly during the identification of the origin and destination of ridership. As shown in Figure 8, the visualization results demonstrate the origin distribution of the distance between different metro stations. Compared with stations with an indicator value of more than 3000, those with an indicator value of less than 1000 are significantly dense and reasonable to be close to the current metro station. Thus, the performance of the recognition of the origin of ridership for indicator I is acceptable.

4.3.2. Repetition number of origin base station at neighboring k metro stations

As shown in Figure 9, the asterisk represents the station, and the circles of different colors represent the recognized origin base station. The number of overlapped base stations of the two neighboring stations in the figure on the left is small, while that in the figure on the right is great. Hence, the recognition results of the two stations in the figure on the left are highly credible. Both indicators can be used to evaluate the origin and destination results of different stations. Overall, the results of this study are highly accurate for most stations.

(a) Good discrimination (b) Confused discrimination

Figure 9. Illustration of different stations with different numbers of overlapped stations.

5. Conclusions

With the aid of spatiotemporal coverage of mobile phone data, we proposed a framework for the identification of ridership from the metro station of boarding and alighting location, and then estimated the effect of distance from different origins and destinations of metro ridership aboveground. Furthermore, the same distribution test of mobile phone data and metro smart data proved the effectiveness of mobile phone data. The results of our aggregated ridership volume also proved the reliability of our proposed method. The boarding, alighting, and transfer metro stations from ridership mobility can be identified. We also recognized the origin location of metro ridership in the aboveground station, which shows that the origin of most ridership comes from the range of 2 to 3 kilometers, and the destination distance of ridership is larger than the origin distance. Moreover, it has been verified that the origin and destination are consistent with the actual metro scenario, which can serve in the planning of the macro-level metro line network.

However, there are some discussions that are worthy of being explored in the future. Even our proposed transfer recognition method can accurately recognize the occurrence of the transfer by combining the station sequence and topology of the metro network. However, this origin and destination verification method is mainly based on an intuitive spatial constraint hypothesis, and the verification indicator system of origin and destination recognition can be further developed, by the number of people at each origin-destination. Furthermore, among the recognition method for the transfer mode of metro riderships, the waiting time and individual traveling pattern of metro passengers need to be considered. Moreover, additional types of data need to be fused to calibrate the correctness and diversity of the estimation results.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This study is supported by the National Natural Science Foundationof China (No. 42261144745, No. 52131203) and Natural Science Foundation of Jiangsu Province (No. BK20232019) and Jiangsu Provincial Scientific Research Center of Applied Mathematics (No. BK20233002).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. Q. Cheng, Y. Lin, X. Zhou, Z. Liu, Analytical formulation for explaining the variations in traffic states: A fundamental diagram modeling perspective with stochastic parameters, *Eur. J. Oper. Res.*, **12** (2023), 452–473. https://doi.org/10.1016/j.ejor.2023.07.005
2. Y. Lin, Y. Xu, Z. Zhao, S. Park, S. Su, M. Ren, Understanding changing public transit travel patterns of urban visitors during COVID-19: A multi-stage study, *Travel Behav. Soc.*, **32** (2023), 100587. https://doi.org/10.1016/j.tbs.2023.100587
3. M. Amirgholy, H. Gao, Optimal traffic operation for maximum energy efficiency in signal-free urban networks: A macroscopic analytical approach, *Appl. Energy*, **329** (2023), 120128. https://doi.org/10.1016/j.apenergy.2022.120128
4. M. Amirgholy, M. Shahabi, Traffic automation and lane management for communicant, autonomous, and human-driven vehicles, *Transp. Res. Part C Emerging Technol.*, **111** (2020), 112–124. https://doi.org/10.1016/j.trc.2019.12.009
5. T. Li, M. Xu, H. Sun, J. Xiong, X. Dou, Stochastic ridesharing equilibrium problem with compensation optimization, *Transp. Res. Part E Logist. Transp. Rev.*, **170** (2023), 102999. https://doi.org/10.1016/j.tre.2022.102999

6. T. Li, Y. Cao, M. Xu, H. Sun, Optimal intersection design and signal setting in a transportation network with mixed HVs and CAVs, *Transp. Res. Part E Logist. Transp. Rev.*, **175** (2023), 103173. https://doi.org/10.1016/j.tre.2023.103173
7. J. Zhang, W. Wu, Q. Cheng, W. Tong, A. Khadka, X. Fu, et al., Extracting the complete travel trajectory of subway passengers based on mobile phone data, *J. Adv. Transp.*, **2022** (2022), 8151520. https://doi.org/10.1155/2022/8151520
8. Q. Du, Y. Zhou, Y. Huang, Y. Wang, L. Bai, Spatiotemporal exploration of the non-linear impacts of accessibility on metro ridership, *J. Transp. Geogr.*, **102** (2022), 103380. https://doi.org/10.1016/j.jtrangeo.2022.103380
9. X. Fu, Y. Zuo, J. Wu, Y. Yuan, S. Wang, Short-term prediction of metro passenger flow with multi-source data: A neural network model fusing spatial and temporal features, *Tunnelling Underground Space Technol.*, **124** (2022), 104486. https://doi.org/10.1016/j.tust.2022.104486
10. Y. Liu, Y. Ji, T. Feng, Z. Shi, A route analysis of metro-bikeshare users using smart card data, *Travel Behav. Soc.*, **26** (2022), 108–120. https://doi.org/10.1016/j.tbs.2021.09.006
11. C. Yang, X. Fu, Z. Liu, Estimation of joint activity–travel benefit with metro smart card data, *J. Transp. Eng. Part A. Syst.*, **148** (2022),112–124. https://doi.org/10.1061/JTEPBS.0000751
12. L. Yang, B. Yu, Y. Liang, Y. Lu, W. Li, Time-varying and non-linear associations between metro ridership and the built environment, *Tunnelling Underground Space Technol.*, **132** (2023), 104931. https://doi.org/10.1016/j.tust.2022.104931
13. P. Peng, Z. Liu, J. Guo, C. Wang, Dynamic metro stations importance evaluation based on network topology and real-time passenger flows, *KSCE J. Civ. Eng.*, **7** (2023), 321–335. https://doi.org/10.1007/s12205-023-0954-7
14. Y. F. Luo, L. Bai, G. Chen, X. Yan, Analysis of space-time variation of passenger flow and commuting characteristics of residents using smart card data of Nanjing metro, *Sustainability*, **11** (2019), 4989. https://doi.org/10.3390/su11184989
15. E. Chen, Z. Ye, C. Wang, W. Zhang, Discovering the spatio-temporal impacts of built environment on metro ridership using smart card data, *Cities*, **95** (2019), 242–253. https://doi.org/10.1016/j.cities.2019.05.028
16. Y. Zhang, E. Yao, J. Zhang, K. Zheng, Estimating metro passengers' path choices by combining self-reported revealed preference and smart card data, *Transp. Res. Part C Emerging Technol.*, **92** (2018), 76–89. https://doi.org/10.1016/j.trc.2018.04.019
17. J. Xing, R. Liu, Y. Zhang, C. F. Choudhury, X. Fu, Q. Cheng, Urban network-wide traffic volume estimation under sparse deployment of detectors, *Transportmetrica A: Transp. Sci.*, **3** (2024), 112–140. https://doi.org/10.1080/23249935.2023.2197511
18. J. Xing, R. Liu, K. Anish, Z. Liu, A customized data fusion tensor approach for interval-wise missing network volume imputation, *IEEE Trans. Intell. Transp. Syst.*, **10** (2023), 1–16. https://doi.org/10.1109/TITS.2023.3289193
19. A. M. Pitale, M. Parida, S. Sadhukhan, Factors influencing choice riders for using park-and-ride facilities: A case of Delhi, *Multimodal Transp.*, **2** (2023), 100065. https://doi.org/10.1016/j.multra.2022.100065
20. X. Ma, Y. Ji, Y. Yuan, N. Van Oort, Y. Jin, S. Hoogendoorn, A comparison in travel patterns and determinants of user demand between docked and dockless bike-sharing systems using multi-sourced data, *Transp. Res. Part A Policy Pract.*, **139** (2020), 148–173. https://doi.org/10.1016/j.tra.2020.06.022

21. Z. Shi, W. Pan, M. He, Y. Liu, Understanding passenger route choice behavior under the influence of detailed route information based on smart card data, *Transportation*, 2023. https://doi.org/10.1007/s11116-023-10432-x
22. Z. Shi, J. Wang, K. Liu, Y. Liu, M. He, Exploring the usage efficiency of electric bike-sharing from a spatial-temporal perspective, *Transp. Res. Part D Transp. Environ.*, **129** (2024), 104139. https://doi.org/10.1016/j.trd.2024.104139
23. Q. Cheng, Z. Liu, J. Guo, X. Wu, R. Pendyala, B. Belezamo, et al., Estimating key traffic state parameters through parsimonious spatial queue models, *Transp. Res. Part C Emerging Technol.*, **137** (2022), 103596. https://doi.org/10.1016/j.trc.2022.103596
24. Q. Cheng, Z. Liu, Y. Lin, X. Zhou, An s-shaped three-parameter (S3) traffic stream model with consistent car following relationship, *Transp. Res. Part B Methodol.*, **153**, (2021), 246–271. https://doi.org/10.1016/j.trb.2021.09.004
25. M. Amirgholy, M. Nourinejad, H. Gao, Balancing the efficiency and robustness of traffic operations in signal-free networks, *Transp. Res. Interdiscip. Perspect.*, **19** (2023), 100821. https://doi.org/10.1016/j.trip.2023.100821
26. M. Amirgholy, M. Nourinejad, Optimal traffic control at smart intersections: Automated network fundamental diagram, *Transp. Res. Part B Methodol.*, **10** (2020), 1016. https://doi.org/10.1016/j.trb.2019.10.001
27. X. Ma, S. Zhang, M. Zhu, T. Wu, M. He, H. Cui, Non-commuting intentions during COVID-19 in Nanjing, China: A hybrid latent class modeling approach, *Cities*, **137** (2023), 104341. https://doi.org/10.1016/j.cities.2023.104341
28. M. He, X. Ma, J. Wang, M. Zhu, Geographically weighted multinomial logit models for modelling the spatial heterogeneity in the bike-sharing renting-returning imbalance: A case study on Nanjing, China, *Sustainable Cities Soc.*, **83** (2022), 103967. https://doi.org/10.1016/j.scs.2022.103967
29. H. Cui, F. Wang, X. Ma, M. Zhu, A novel fixed-node unconnected subgraph method for calculating the reliability of binary-state networks, *Reliab. Eng. Syst. Saf.*, **226** (2022), 108687. https://doi.org/10.1016/j.ress.2022.108687
30. Y. Sun, S. Jungang, P. Schonfeld, Identifying passenger flow characteristics and evaluating travel time reliability by visualizing AFC data: a case study of Shanghai Metro, *Public Transp.*, **8** (2016), 23–34. https://doi.org/10.1007/s12469-016-0137-8
31. M. C. Chen, Y. Wei, Exploring time variants for short-term passenger flow, *J. Transp. Geogr.*, **19** (2011), 488–498. https://doi.org/10.1016/j.jtrangeo.2010.04.003
32. J. Kim, S. Tak, J. Lee, H. Yeo, Integrated design framework for on-demand transit system based on spatiotemporal mobility patterns, *Transp. Res. Part C Emerging Technol.*, **150** (2023), 104087. https://doi.org/10.1016/j.trc.2023.104087
33. 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** (2020), 651–675. https://doi.org/10.3390/ijgi9110651
34. Z. Shi, N. Zhang, Y. Liu, W. Xu, Exploring spatiotemporal variation in hourly metro ridership at station level: the influence of built environment and topological structure, *Sustainability*, **10** (2018), 12–34. https://doi.org/10.3390/su10124564
35. S. Jungang, L. Yang, Service-oriented train timetabling with collaborative passenger flow control on an oversaturated metro line: An integer linear optimization approach, *Transp. Res. Part B Methodol.*, **110** (2018), 112–134. https://doi.org/10.1016/j.trb.2018.02.003

36. R. Liu, S. Li, L. Yang, Collaborative optimization for metro train scheduling and train connections combined with passenger flow control strategy, *Omega*, **90** (2018), 12–27. https://doi.org/10.1016/j.omega.2018.10.020
37. S. Li, M. Dessouky, L. Yang, Z. Gao, Joint optimal train regulation and passenger flow control strategy for high-frequency metro lines, *Transp. Res. Part B Methodol.*, **99** (2017), 113–137. https://doi.org/10.1016/j.trb.2017.01.010
38. F. Jin, E. Yao, Y. Zhang, S. Liu, Metro passengers' route choice model and its application considering perceived transfer threshold, *PloS One*, **12** (2017), e0185349. https://doi.org/10.1371/journal.pone.0185349
39. X. Wang, E. Yao, S. Liu, Travel choice analysis under metro emergency context: utility? regret? or both?, *Sustainability*, **10** (2018), 3852. https://doi.org/10.3390/su10113852
40. R. Cascajo, A. Garcia-Martinez, A. Monzón, Stated preference survey for estimating passenger transfer penalties: design and application to Madrid, *Eur. Transp. Res. Rev.*, **9** (2017), 1–11. https://doi.org/10.1007/s12544-017-0260-x
41. L. Sun, Y. Lu, J. G. Jin, D. H. Lee, K. Axhausen, An integrated Bayesian approach for passenger flow assignment in metro networks, *Transp. Res. Part C Emerging Technol.*, **52** (2015), 123–147. https://doi.org/10.1016/j.trc.2015.01.001
42. J. Zhao, F. Zhang, L. Tu, C. Z. Xu, D. Shen, L. Ruyue, et al., Estimation of Passenger route choice pattern using smart card data for complex metro systems, *IEEE Trans. Intell. Transp. Syst.*, **18** (2016), 231–246. https://doi.org/10.1109/TITS.2016.2587864
43. J. Wu, Y. Qu, H. Sun, H. Yin, X. Yan, J. Zhao, Data-driven model for passenger route choice in urban metro network, *Physica A*, **524** (2019), 351–366. https://doi.org/10.1016/j.physa.2019.04.231
44. Q. Zhang, B. Han, Simulation model of passenger transfer behavior in metro station, *Appl. Mech. Mater.*, **253** (2012), 1791–1796. https://doi.org/10.4028/AMM.253-255.1791
45. Y. H. Cheng, W. C. Tseng, Exploring the effects of perceived values, free bus transfer, and penalties on intermodal metro-bus transfer users' intention, *Transp. Policy*, **47** (2016), 127–138. https://doi.org/10.1016/j.tranpol.2016.01.001
46. L. Yang, Y. Zhang, S. Li, Y. Gao, A two-stage stochastic optimization model for the transfer activity choice in metro networks, *Transp. Res. Part B Methodol.*, **83** (2016), 271–297. https://doi.org/10.1016/j.trb.2015.11.010
47. J. Schlaich, T. Otterstätter, M. Friedrich, Generating trajectories from mobile phone data, in *Proceedings of the 89th Annual Meeting Compendium of Papers, Transportation Research Board of the National Academies*, **1** (2010).
48. S. Hoteit, S. Secci, S. Sobolevsky, C. Ratti, G. Pujolle, Estimating human trajectories and hotspots through mobile phone data, *Comput. Networks*, **64** (2014), 121–134. https://doi.org/10.1016/j.comnet.2014.02.011
49. M. Li, S. Gao, F. Lu, H. Zhang, Reconstruction of human movement trajectories from large-scale low- frequency mobile phone data, *Comput. Environ. Urban Syst.*, **77** (2019), 112–134. https://doi.org/10.1016/j.compenvurbsys.2019.101346
50. J. Steenbruggen, M. T. Borzacchiello, P. Nijkamp, H. Scholten, Mobile phone data from GSM networks for traffic parameter and urban spatial pattern assessment: a review of applications and opportunities, *GeoJournal*, **78** (2013), 223–243. https://doi.org/10.1007/s10708-011-9413-y

51. S. Phithakkitnukoon, Analysis of weather effects on people's daily activity patterns using mobile phone GPS data, in *Urban Informatics Using Mobile Network Data: Travel Behavior Research Perspectives*, Singapore, Springer Nature Singapore, (2022), 161–181.
52. S. Jiang, J. Ferreira, M. C. Gonzalez, Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore, *IEEE Trans. Big Data*, **3** (2017), 208–219. https://doi.org/10.1109/TBDATA.2016.2631141
53. P. Widhalm, Y. Yang, M. Ulm, S. Athavale, M. C. Gonzalez, Discovering urban activity patterns in cell phone data, *Transportation*, **42** (2015), 597–623. https://doi.org/10.1007/s11116-015-9598-x
54. Y. Yuan, R. Martin, Extracting dynamic urban mobility patterns from mobile phone data, in *Geographic Information Science: 7th International Conference*, Springer Berlin Heidelberg, (2012), 354–367.
55. M. Diao, Y. Zhu, J. Ferreira, C. Ratti, Inferring individual daily activities from mobile phone traces: A Boston example, *Environ. Plann. B: Urban Anal. City Sci.*, **43** (2015), 122–134. https://doi.org/10.1177/0265813515600896
56. Y. Wang, J. Cong, P. Wang, X. Liu, H. Tang, A data-fusion approach for speed estimation and location calibration of a metro train based on low-cost sensors in smartphones, *IEEE Sens. J.*, **19** (2019), 10744–10752. https://doi.org/10.1109/JSEN.2019.2933638
57. Z. Duan, Z. Lei, H. Zhang, W. Li, J. Fang, J. Li, Understanding evacuation and impact of a metro collision on ridership using large-scale mobile phone data, *IET Intel. Transp. Syst.*, **11** (2017), 156–173. https://doi.org/10.1049/iet-its.2016.0112
58. S. Liu, F. Zhang, Y. Ji, X. Ma, Y. Liu, S. Li, et al., Understanding spatial-temporal travel demand of private and shared e-bikes as a feeder mode of metro stations, *J. Cleaner Prod.*, **398** (2023), 136602. https://doi.org/10.1016/j.jclepro.2023.136602
59. S. Liu, F. Zhang, Y. Ji, X. Ma, Y. Liu, S. Li, et al., Exploring the usage efficiency of electric bike-sharing from a spatial-temporal perspective, *Transp. Res. Part D Transp. Environ.*, **129** (2024), 104139. https://doi.org/10.1016/j.trd.2024.104139
60. C. Ding, T. Liu, X. Cao, L. Tian, Illustrating nonlinear effects of built environment attributes on housing renters' transit commuting, *Transp. Res. Part D Transp. Environ.*, **112** (2022), 103503. https://doi.org/10.1016/j.trd.2022.103503
61. X. Ma, S. Zhang, M. Zhu, T. Wu, M. He, H. Cui, Non-commuting intentions during COVID-19 in Nanjing, China: A hybrid latent class modeling approach, *Cities*, **137** (2023), 104341. https://doi.org/10.1016/j.cities.2023.104341
62. X. Ma, Y. Ji, Y. Yuan, N. Van Oort, Y. Jin, S. Hoogendoor, A comparison in travel patterns and determinants of user demand between docked and dockless bike-sharing systems using multi-sourced data, *Transp. Res. Part A Policy Pract.*, **139** (2020), 148–173. https://doi.org/10.1016/j.tra.2020.06.022
63. Q. Zhang, Y. Shi, R. Yin, H. Tao, Z. Xu, Z. Wang, et al., An integrated framework for real-time intelligent traffic management of smart highways, *J. Transp. Eng. Part A. Syst.*, **149** (2023), 04023055. https://doi.org/10.1061/JTEPBS.TEENG-7729
64. X. Yang, Y. Shi, J. Xing, Z. Liu, Autonomous driving under V2X environment: state-of-the-art survey and challenges, *Intell. Transp. Infrastruct.*, **1** (2022), 12–23. https://doi.org/10.1093/iti/liac020

65. Q. Cheng, Z. Liu, J. Lu, G. List, P. Liu, X. S. Zhou, Using frequency domain analysis to elucidate travel time reliability along congested freeway corridors, *Transp. Res. Part B Methodol.*, **184** (2024), 102961. https://doi.org/https://doi.org/10.1016/j.trb.2024.102961