Physica A 576 (2021) 126058

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Assessing temporal–spatial characteristics of urban travel behaviors from multiday smart-card data

Yue Deng [a], Jiaxin Wang [b], Chao Gao [a,c,*], Xianghua Li [c], Zhen Wang [c], Xuelong Li [c]

[a] *College of Computer and Information Science, Southwest University, Chongqing 400715, China*
[b] *State Grid Information & Telecommunication Group Co. Ltd., Beijing 100031, China*
[c] *School of Mechanical Engineering and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xi'an 710072, China*

ARTICLE INFO

Article history:
Received 14 December 2020
Received in revised form 28 February 2021
Available online 1 May 2021

Keywords:
Rail transit
Travel behaviors
Community detection

ABSTRACT

The rail transit has difficulties in meeting daily travel needs of passengers owing to a large population and accelerating urbanization. Analyzing urban travel behaviors with big data helps the design in infrastructures and the optimized personnel allocation. Furthermore, travel behaviors are characterized by dynamic at different time and locations, displaying the rule of urban traffic operation. This paper utilizes smart card data in two cities with different geographical features to analyze the temporal–spatial characteristics of urban travel behaviors. More specifically, by creating travel networks based on the pick-up and drop-off stations and the passenger population among these stations, an interesting observation is that the community structure of travel networks owns a metabolic trend and a stable feature simultaneously. The finding shows that the traffic system can be managed in several parts. Moreover, similar mobility patterns exist in some stations, which can be organized and controlled in the same way. Finally, travel behaviors are related to the urban layout and structure, so the distribution of urban areas can be understood better. Experiments provide enlightening insights for policy makers to comprehend the urban travel behaviors, thus improving the rail transit service plans and scheduling strategies.

1. Introduction

The acceleration of the urbanization and the increase of population lead to the station overcrowding [1,2]. It significantly affects the travel efficiency of passengers and slows down urban development [3]. Revealing and understanding urban travel behaviors can help transportation departments ensure travel needs while improving efficiency and saving management costs [4–8]. More specifically, managers are able to control the number of trains and adjust the allocation of personnel according to the different travel behaviors [9–13]. What is more, it is able to identify urban functional areas (e.g., commercial district, residential area, industrial park), thus finding travel rules and providing scientific evidence in formulating operation and service strategies for the public transportation sector [14–16]. Nonetheless, no matter in time or in space, urban travel behaviors are changing [17–19]. Therefore, it is critical to identify the urban travel behaviors, especially to describe such dynamic behaviors [20].

* Corresponding author at: College of Computer and Information Science, Southwest University, Chongqing 400715, China.
E-mail address: cgao@swu.edu.cn (C. Gao).

https://doi.org/10.1016/j.physa.2021.126058
0378-4371/

Table 1
Examples of Dataset 1. *ID* denotes the passenger ID; *Time* denotes the pick-up or drop-off time; *Station* represents the pick-up or drop-off station; *Spending* stands for the trip expense.

Example	ID	Time	Station	Spending
1	2201252167	2015-04-24 18:46:19	11	0.00
2	2201252167	2015-04-24 19:13:01	21	4.00
3	2801581273	2015-04-24 09:00:06	18	0.00
4	2801581273	2015-04-24 09:32:05	74	3.60

The existing researches of travel behaviors based on the rail transit card data are mainly divided into two categories. One is from the perspective of the user classification to reveal characteristics of travel behaviors for different types of users [21–24]. The other focuses on the dynamic travel patterns of passengers [25–29]. Concerning the first category about the user classification, passengers can be classified as transit commuter, regular OD passenger, habitual time passenger and irregular passengers [30]. For example, Kieu et al. adopted a density-based noise application spatial clustering (DBSCAN) algorithm to mine travel patterns, assessing and identifying the types of smart card users [30]. Ma et al. first used the spatial clustering and multi-standard decision analysis to identify commuters and non-commuters in Beijing, and then employed the spatial data visualization method to research the living and working places of commuters on the map [31]. There is an unbalanced distribution of working and living places in Beijing. The research provides some ideas for management departments to balance job-housing relationship.

In terms of the second category concerning dynamic travel patterns, Zhong et al. proposed a basic variability measurement to evaluate the stability of travel rules in the time dimension based on the smart card data in Beijing, Singapore and London, in order to understand the variability of passenger travel patterns [32]. However, the spatial variability of passenger travel rules is not addressed in their researches. Ali et al. analyzed the characteristics of the waiting time and transfer passenger flows to provide new ideas for rail transit operators to improve the infrastructures [33]. Although the temporal variability has been analyzed in some researches, the spatial dynamic characteristics are not contained simultaneously. Therefore, in this paper, temporal and spatial characteristics are both considered to display the urban travel behaviors. The main contributions of this paper can be listed as follows.

- This paper reveals the temporal–spatial characteristics of urban travel behaviors, which can be applied in the field of the public transportation.
- This paper compares two cities with different geographical features in China (i.e., Shanghai and Chongqing) so as to make targeted strategies based on the temporal–spatial characteristics of the city.

The remaining paper is organized as follows. Section 2 depicts the preliminary including datasets and the travel networks. Section 3 involves temporal–spatial analysis of urban travel behaviors in terms of statistics and travel patterns in Shanghai and Chongqing. Following that, Section 4 discusses the differences of travel behaviors in the two cities. Section 5 summarizes this paper.

2. Preliminary

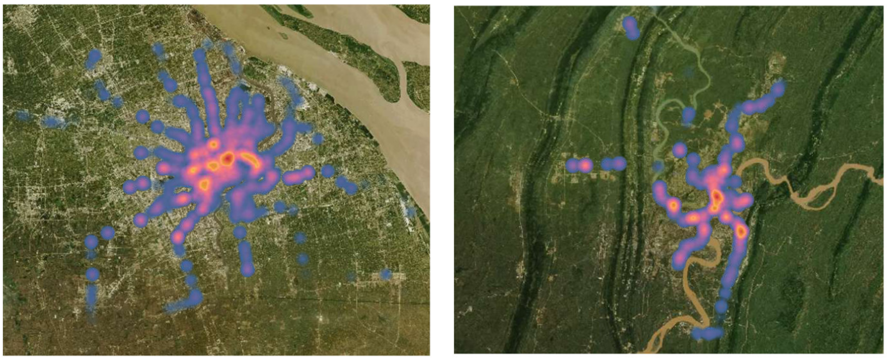
In this section, we first introduce the datasets in Section 2.1. Following that, the process of network construction is displayed in Section 2.2.

2.1. Dataset

Considering the influence of topographies on travel behaviors, we collect data from two cities with different geographical features shown in Fig. 1.

Dataset 1 (Shanghai Subway): The land topography in Shanghai is slightly inclined from east to west. Except for a few hills in the southwest, the terrain is flat. This paper collects the dataset of Shanghai metro from 6:00 to 24:00 between April 20, 2015 and April 26, 2015. There are 289 stations and 16 lines in total. The data generally have more than 12 million records every day, including the IDs of users, the card swiping time at different stations, the spending of trips that can be used with the card swiping time to identify the pick-up and drop-off stations. The specific examples of records are shown in the Dataset 1 of Table 1.

Dataset 2 (Chongqing Subway): The city of Chongqing spans mountains and two rivers (namely the Yangtze River and the Jialing River) which divide Chongqing into several parts connected by the expressway and light rail [34]. This section employs the data of the subway in Chongqing for one week which own almost 1.5 million records per day. There are 89 stations and 4 lines in the subway in total. The data records the IDs of users, the pick-up and drop-off time as well as the corresponding stations. The specific examples are shown in the Dataset 2 of Table 2.



(a) Satellite heat map in Shanghai (b) Satellite heat map in Chongqing

Fig. 1. The satellite heat maps during the morning peak in Shanghai and Chongqing, respectively. The brighter the color is, the more passenger flows are. It shows that there is a flat topography and concentrated population in Shanghai, while a steep terrain and relatively dispersive population in Chongqing.

Table 2
Examples of Dataset 2. *ID* denotes the passenger ID; *Time* denotes the pick-up or drop-off time; *Station* represents the pick-up or drop-off station; *In_* and *Out_* represent pick-up and drop-off, respectively.

Example	ID	In_Station	Out_Station	In_Time	Out_Time
1	7927395	0212	0323	06:36:25	07:13:50
2	7861020	0303	0315	06:34:03	07:13:42
3	7981456	0113	0102	06:35:47	07:10:54
4	4457964	0111	0107	06:58:02	07:13:20

2.2. Network construction

To begin with, we clean the redundant information for accurately and rigorously analyzing the characteristics of travel behaviors. Then, we extract the origin and destination data (OD data) in an interval of 1 h and 30 min, respectively, so as to explore the temporal–spatial travel behaviors and their characteristics. For *Dataset 1 (Shanghai Subway)*, according to the two records with the same *ID*, the pick-up and drop-off stations can be judged based on the time sequence and the spending. More specifically, a user must swipe a card in the pick-up station with 0.00 yuan early and in the drop-off station with more than 0.00 yuan after to show the trip expense. Taking the first two records as an example in Table 1, for the user with $ID = 2201252167$, the *Time* is 18:46:19 and 19:13:01, and the *Spending* is 0.00 yuan and 4.00 yuan, respectively. We can infer that the user gets on the train at v_{11} (Shaanxi South Road) in *Line*1 and gets off at v_{21} (Pengpu New Village) in *Line*1. For *Dataset 2 (Chongqing Subway)* in Table 2, each record represents a complete trip which can obtain the pick-up and drop-off stations directly. According to above information, we count the passenger flows from pick-up stations to drop-off stations in a period in order to obtain the OD data in different cities.

Based on the obtained OD data, travel networks of passengers are constructed by the SSPN method [35]. The network of traveling in a metro system, abbreviated as a travel network, is formulated as a directed weighted graph $G = (V, E)$. V represents a set of stations ($V \neq \emptyset$) and E refers to a set of edges ($\|E\| \geq 1$). According to the passenger flows, there is a directed edge $e = \left(v_i \to v_j\right) \in E$ with an origin–destination trip from an origin v_i to a destination v_j, where $v_i, v_j \in V$. Moreover, the weight of e is equal to the volume of passenger flows from v_i to v_j.

In order to overcome the effect of the noise caused by the small volume of passenger flows, an appropriate threshold on the analysis of human mobility is first identified. The edges less than such a threshold are deleted from the dynamic travel networks. Eventually, we extract networks of 18 periods from 06:00 to 24:00 every day in an interval of 1 h and those of 36 periods in an interval of 30 min.

3. Results

The temporal–spatial analysis of urban travel behaviors is an important part for the research of traffic data. Section 3.1 first describes the temporal characteristics of urban travel behaviors in Shanghai and Chongqing. The spatial characteristics of urban travel behaviors are displayed in Section 3.2.

3.1. Urban travel behaviors from the time dimension

The travel behaviors of passengers are dynamic in different periods. In order to clarify the travel characteristics of passengers with respect to time, this paper summarizes the temporal characteristics and displays the specific analysis process in terms of statistics and travel patterns.

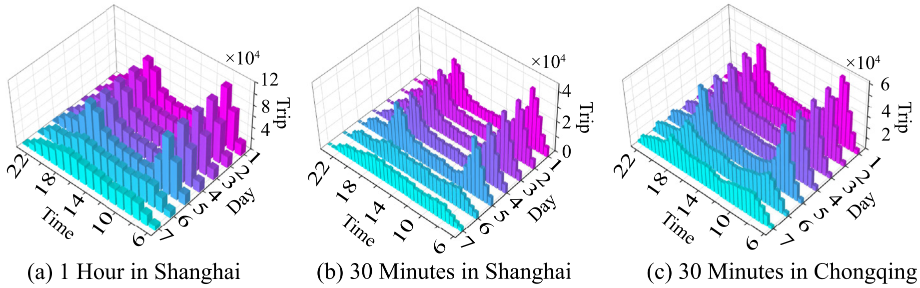


Fig. 2. The statistical charts of passenger flows for one week. The horizontal axis represents the time period, and the vertical axis stands for the passenger flows corresponding to the time quantum. (a) and (b) are in an interval of 1 h and 30 min in Shanghai, respectively. (c) is in an interval of 30 min in Chongqing. The results show that there are obvious morning and evening peaks on weekdays. Moreover, the passenger flows in the morning period (08:00–09:00) are larger than those in the evening period (18:00–19:00).

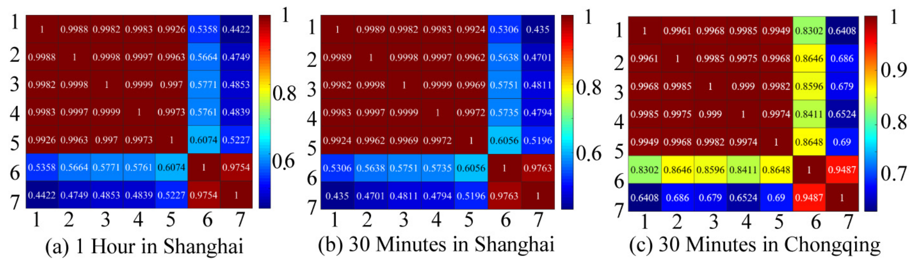


Fig. 3. The correlation coefficient matrices in a week in an interval of 1 h and 30 min in Shanghai, as well as 30 min in Chongqing. The axes are on behalf of the days of a week, and the color blocks for every two days represent the similarity degree of travel behaviors. There are highly similar travel behaviors at working days and rest days, respectively. And Friday is the transition from working mode to rest mode.

Table 3
The specific morning and evening peaks.

Peak	Period	Time granularity
Morning peak	08:00–09:00	1 h
Evening peak	18:00–19:00	1 h
Morning peak	08:00–08:30	30 min
Evening peak	18:00–18:30	30 min

3.1.1. Statistical analysis

Fig. 2 depicts passenger flows in different periods of one week, which shows the change of passenger flows from 06:00 to 24:00 in Shanghai and Chongqing, respectively. The results show that the passenger flows of weekdays have typical peaks which are concluded in Table 3. Moreover, the morning period (08:00–09:00 or 08:00–08:30) owns higher passenger flows than the evening period (18:00–19:00 or 18:00–18:30) on weekdays. Possibly because after work residents in various companies set off at different time. Compared with weekdays, the overall trend of passenger flows is flatter on weekends. As a result, public transportation managers can increase the number of vehicles, raise the number of security personnel, or add iron railings in subway stations to lengthen the walking time, so as to reduce the pressure on the platforms during peaks.

To further find the temporal mobility patterns in one week, this paper counts the correlation of travel patterns in different days, displayed in Fig. 3. More specifically, it calculates the real-time passenger flows at different time as passenger flow vectors and quantifies the correlation of the passenger flow vectors among one week by the Pearson correlation analysis. The results show the degree of similarity in travel patterns between different days. Fig. 3(a) and (b) exhibit the correlation coefficient matrices in an interval of 1 h and 30 min in Shanghai, respectively. The matrix in an interval of 30 min in Chongqing is shown in Fig. 3(c). The closer the value on the color blocks is to 1, the higher the similarity is. In contrast, the closer it is to 0, the less similarity the two days share. It is obvious that the correlation coefficients of weekdays are around 0.99 and those of weekends are about 0.97. Thus, the conclusion is that there are highly similar travel behaviors at working days and rest days, respectively. Additionally, the change of correlation coefficients from Friday to Sunday indicates the transformation of lifestyles from working mode to leisure mode. Therefore, Friday is the transition from working days to rest days.

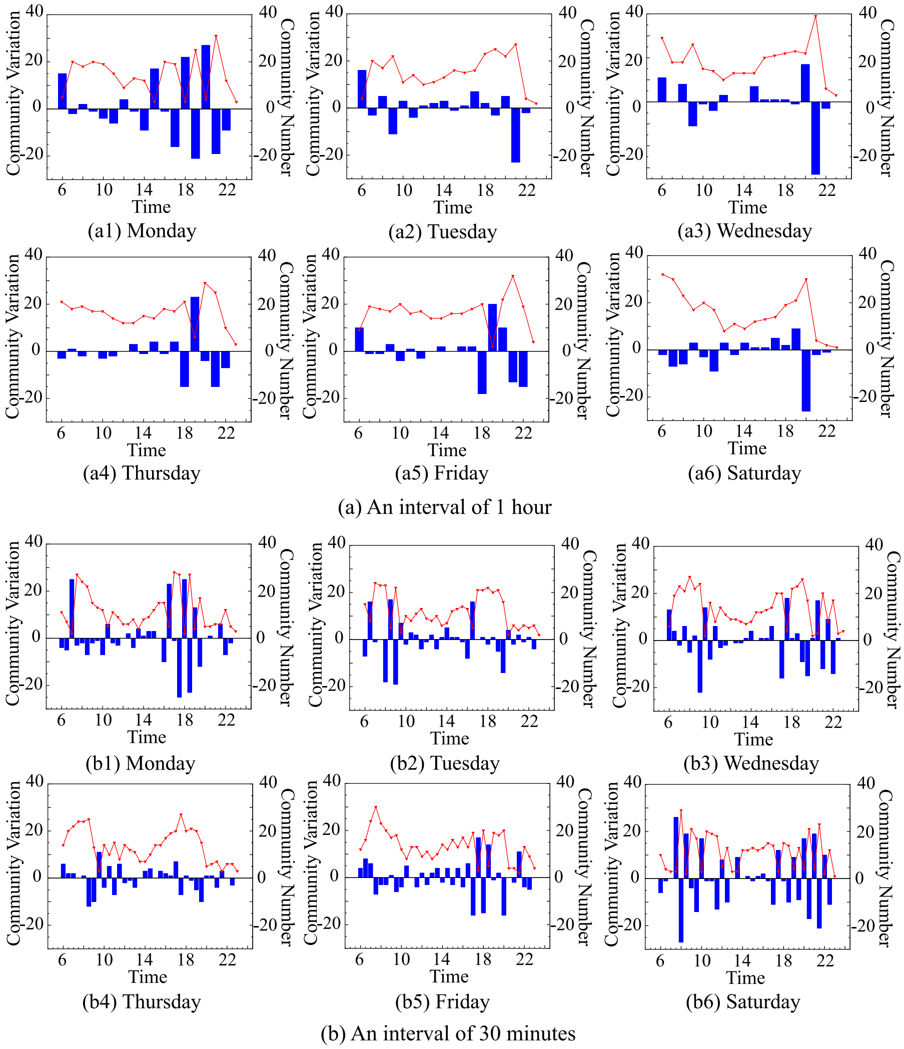


Fig. 4. (a) and (b) display the variation of the community number for one week, in an interval of 1 h and 30 min in Shanghai, respectively. Each number on the horizontal axis represents the time period. The left vertical axis is the change of the community number in adjacent time periods (the blue bar chart), and the right vertical axis refers to the total number of communities in each time period (the red solid line). The results show that the mobility patterns of passengers are changeable but regular simultaneously.

3.1.2. Travel pattern analysis

In the context of daily flows, passenger flows connect inner areas of a city and display certain mobility patterns. The mobility patterns are able to reveal the latent relationship of stations. The analyzing of travel networks by community detection is an effective way to identify the mobility patterns. Based on the travel networks in Section 2.2, the infomap community detection algorithm is applied for observing the mobility patterns of passengers [36]. Such an algorithm has been proved to be an effective method for directed weighted networks.

Fig. 4 shows the number of communities and the change of communities in an interval of 1 h and 30 min in Shanghai, respectively. We find that although the travel of passengers is dynamic, there are regular mobility patterns. More specifically, Fig. 4(a) demonstrates that there are around 20 communities both in the morning and evening periods (08:00–09:00 and 18:00–19:00) on weekdays except the evening period (18:00–19:00) on Monday, while the community number of many periods is about 10. Moreover, the number of communities slightly decreases in the morning period (08:00–09:00) on weekdays except Thursday. Despite the large passenger flows in the morning period (08:00–09:00), passengers mostly travel to working areas, so the number of communities becomes large but slightly decreases. Further, Fig. 4(b) demonstrates that the variation trend of the total community number is similar to that of passenger flows. The community number is larger in the morning and afternoon, but less in the noon and evening. Because the time granularity is reduced in Fig. 4(b) compared with Fig. 4(a), the community number is more greatly affected by the passenger flows.

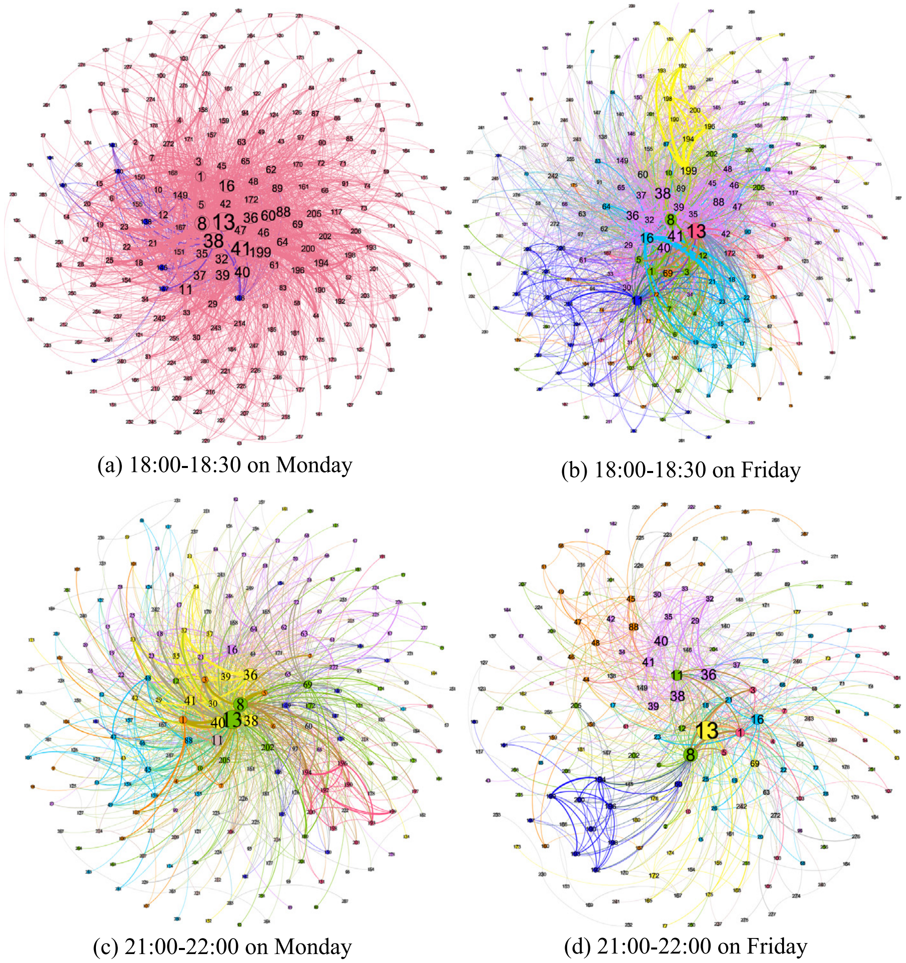


Fig. 5. The networks during (a) 18:00–18:30 on Monday, (b) 18:00–18:30 on Friday, (c) 21:00–22:00 on Monday and (d) 21:00–22:00 on Friday in Shanghai. Nodes of the same color belong to the same community. The size of the points stands for the degree, and the number of the point represents the label. The thickness of the edges means the weights. (a) has more stations with high degree than (b), and these stations are located in the center of Shanghai. Thus, there are more congregated passenger flows in (a) than (b). In (c) and (d), the stations with a high degree, such as v_{13} (People's Square), v_8 (Xujiahui), v_{38} (Jing'an Temple) and v_{36} (Zhongshan Park) are all in the business district.

Therefore, the passenger flows are simultaneously large and congregated in the morning period (08:00–09:00), which provide suggestions for traffic managers to flexibly schedule vehicles.

However, the community number significantly decreases in the evening period (18:00–18:30) on Monday in Fig. 4(b). In order to deeply analyze the plunge, Fig. 5(a) and (b) depict the network between 18:00 and 18:30 on Monday and Friday, respectively. It shows that there are more stations with larger degree on Monday than Friday (e.g., v_{41} (Lujiazui), v_{13} (People's Square), v_8 (Xujiahui), v_{38} (Jing'an Temple)), which stand in the center of Shanghai. In addition, the number of nodes and edges on Monday is almost equal to that on other working days, shown in Table 4, which means the decrease has nothing to do with the number of nodes and edges. Thus, it is the stronger congregated passenger flows in the evening period (18:00–19:00) on Monday in the downtown that contributes to the decrease of community number. Therefore, the traffic department ought to add vehicles and increase security personnel in v_{41}, v_{13}, v_8, v_{38}, especially in the evening period (18:00–18:30) on Monday.

Except the morning and evening periods (08:00–09:00 and 18:00–19:00), the mobility patterns of other periods also deserve attention. Fig. 4(a) shows the community number is large during 21:00–22:00 or 20:00–21:00 on weekdays. To investigate the reasons, the networks during 21:00–22:00 on Monday and Friday are indicated in Fig. 5(c) and (d), respectively. The results show the large and dispersive passenger flows going shopping or entertaining causes a large number of communities during 21:00–22:00. More specifically, we find that the degrees of v_{13}, v_8, v_{38} and v_{36} in the two subgraphs are large shown in Fig. 5(c) and (d). Combined with the geographic information of Shanghai, these nodes all locate in the commercial areas. Therefore, traffic managers should reasonably allocate vehicles from 20:00 to 22:00. Moreover, they can select some stations (e.g., v_{13} (People's Square), v_8 (Xujiahui), v_{38} (Jing'an Temple) and v_{36} (Zhongshan

Table 4
The basic information of networks in the period of 18:00 to 18:30 for a week in Shanghai.

Date	Edges	Nodes
Monday	2823	251
Tuesday	2803	249
Wednesday	2761	251
Thursday	2763	248
Friday	2809	248

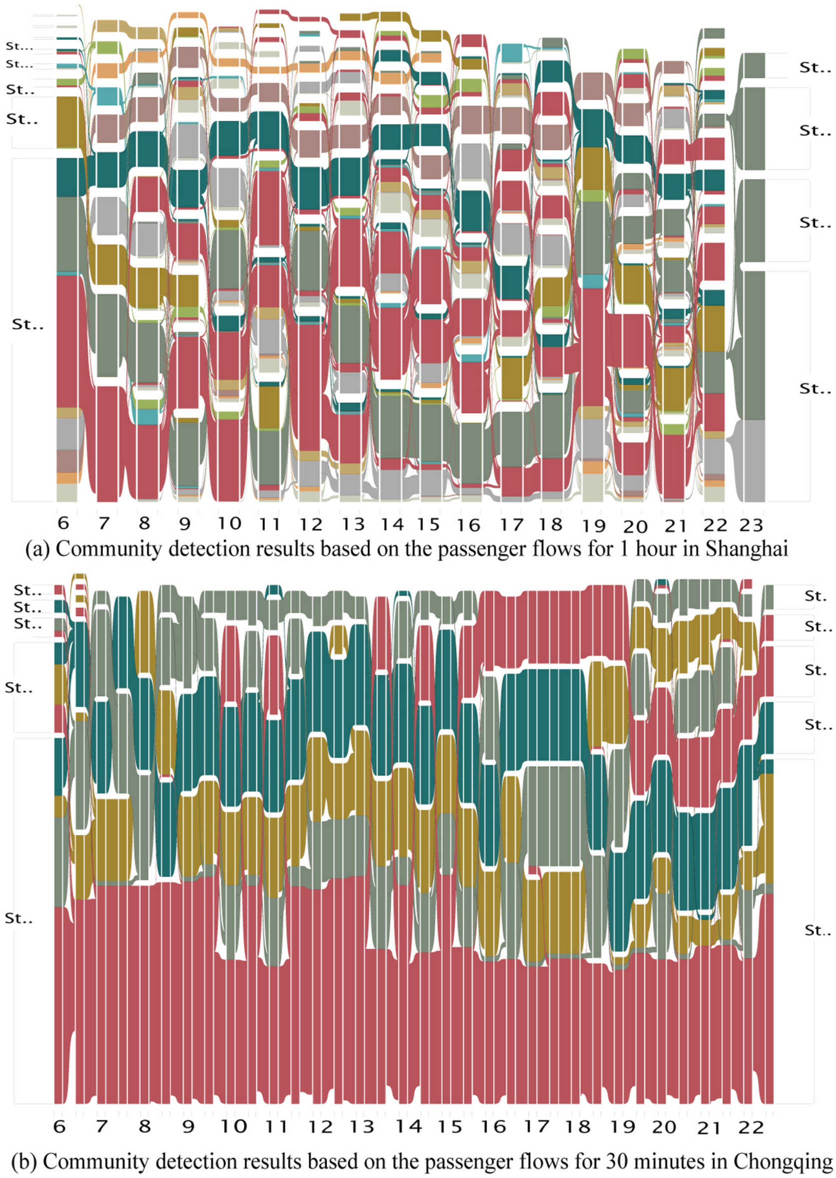


Fig. 6. Results of community detection in Shanghai and Chongqing. Horizontally, Friday is divided by 1 h in Shanghai and 30 min in Chongqing from 6:00 to 24:00. Vertically, the diagram has many rectangles with white lines and colored parts, and each represents a community. The colors represent the merging and splitting of the community. The clustering changes with time, but some stations always exist in the same community. It indicates that the community structure of travel networks is varied and stable at the same time.

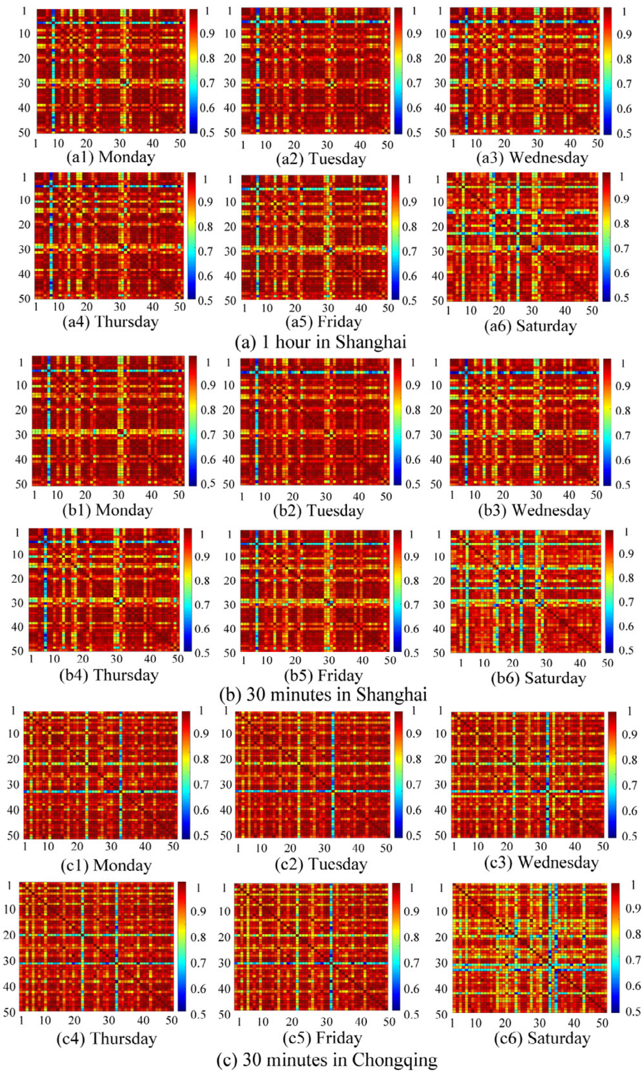
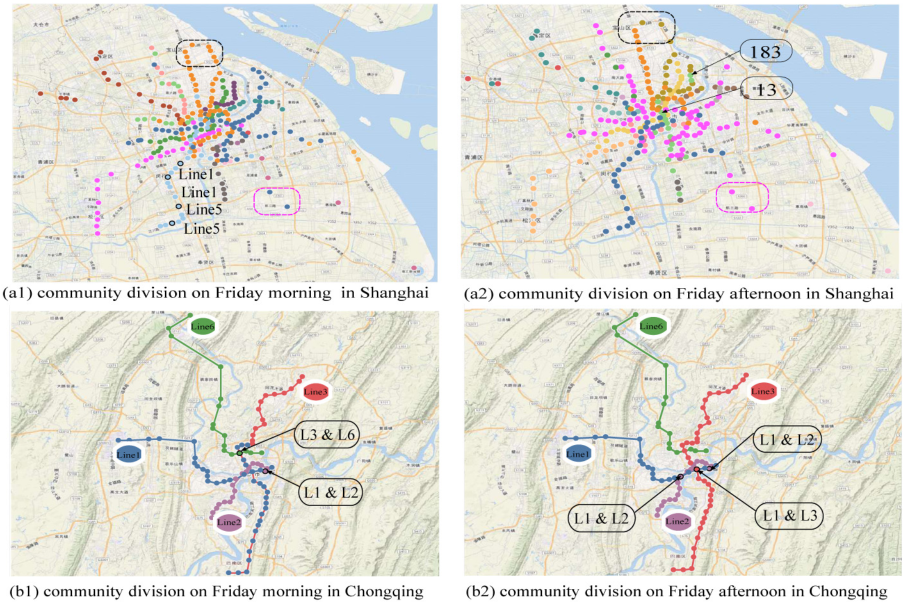


Fig. 7. The correlation coefficient matrices between stations. The coordinate values stand for the different stations, and the color blocks corresponding to the two sites represent the similarity between them. (a) matrices in an interval of 1 h in Shanghai. (b) matrices in an interval of 30 min in Shanghai. (c) matrices in an interval of 30 min in Chongqing. It shows significant differences of mobility patterns between working days and rest days. However, the mobility patterns in some stations bear resemblance to each other.

Park)) as the central stations of the line to provide the unified technical and business support for other stations. The adjacent stations (e.g., v_{173} (Qufu Road), v_{172} (Big World), v_{151} (Bang Road), v_{60} (Yishan Road)) can be managed as satellite stations. Hence, these central stations become the management and cost center to improve the efficiency when dealing with emergencies.

Apart from analyzing the single passenger travel network, the overall community structure is also a significant part for revealing the characteristics of travel behaviors. Therefore, Fig. 6 visualizes the results of community detection on Friday. It shows that although the variation of the community structure exists on the whole, some stations always remain in the same cluster, in spite of the different number of trips, changing travel time or travel distance. More specifically, these color blocks, which denote different communities, are continuously merging and splitting. Moreover, according to the data of Shanghai metro in Fig. 6(a), v_{21}, v_{23}, v_{26}, v_{25}, v_{22} are always in the same community. In Fig. 6(b), v_1, v_{26}, v_{32}, v_{27}, v_{23}, v_{24} are in the same community all the time in Chongqing. Therefore, these stations remain their roles in terms of urban functions on different days. We can partition the urban system into several local systems, a structure of city appears consequently, thus the urban traffic system can be organized and managed in several parts instead of by many stations.



(a1) community division on Friday morning in Shanghai
(a2) community division on Friday afternoon in Shanghai
(b1) community division on Friday morning in Chongqing
(b2) community division on Friday afternoon in Chongqing

Fig. 8. Community division results on maps in Shanghai and Chongqing. The dots stand for stations and the lines represent subway lines. As for dots, different colored dots belong to different communities. As to lines, lines with different colors mean different subway lines. (a) Community division on Friday morning and evening in Shanghai. (b) Community division on Friday morning and evening in Chongqing. Stations in the same community have the same color. The results show that stations in the same community own similar passenger clustering characteristics.

3.2. The urban travel behaviors from the spatial dimension

In addition to the temporal travel behaviors, this paper further reveals the spatial characteristics. More specifically, it projects the travel records into geographic space to analyze the spatial characteristics dynamically. We demonstrate results in terms of statistical analysis and travel pattern analysis.

3.2.1. Statistical analysis

The spatial mobility patterns can be identified by the passenger flows of each station. This article randomly selects 50 stations and then renumbers the 50 stations as v_1-v_{50}. Fig. 7 displays the correlation coefficient matrices of passenger flows in stations in an interval of 1 h and 30 min in Shanghai and Chongqing, respectively. The results show the great differences in the spatial mobility patterns between working days and rest days. More specifically, the matrices on weekdays are similar, but the matrix of Saturday is very different from them.

The variations of correlation coefficients can tell different components of mixed land use and urban functional areas. More specially, Fig. 7 shows there are some stations with a high similarity all the time, which tend to form a neighborhood with very similar land uses and jointly demonstrate the spatial structure of the city. On weekdays, the stations near residential areas (e.g., v_1, v_5, v_8, v_{45} and v_{34}) or stations near work areas (e.g., v_3, v_{46}, v_{44}, v_{42}, v_{38}) own a high similarity in passenger flows, respectively. While on weekends, stations close to the commercial and recreation areas (e.g., v_{40}, v_2, v_{19}, v_{26} and v_{43}) are similar to each other. It indicates that the stations close to residential areas (e.g., houses) are similar in travel patterns and those near working areas (e.g., offices) bear resemblance to each other on working days. Further, the stations near commercial areas (e.g., shopping malls) and recreation areas (e.g., parks) have similar travel patterns on rest days. Therefore, the public transportation managers can adopt similar management modes for the stations with a high similarity. For example, the same number of trains, channels and escalators should be configured in the uniform period to ensure travel needs of passengers, improve management efficiency and save management costs.

3.2.2. Travel pattern analysis

Based on the theory of community detection in the field of network science, the connectivity of nodes inside their own cluster is stronger than that between other clusters [37–41]. Community structures are of significance for exploring social network behaviors [42,43]. An algorithm of community detection aims to divide a network into various structures according to the sparsity of connections [44]. In the context of urban mobility, stations inside the same cluster have more interaction between each other and together they form a neighborhood system.

In order to clearly explore the connections between stations, this article visualizes the results of station partition on a map. Fig. 8 shows the community structure of travel networks during the morning period (08:00–08:30) and evening period (18:00–18:30) on Friday in Shanghai and Chongqing, respectively. The results show that the stations of the same line almost belong to the same communities, and the stations in the same community are mostly adjacent in the geographical locations. Therefore, stations in the same community, such as adjacent stations and stations of the same line, have similar passenger mobility patterns. However, some stations in the same community can also belong to different lines. As shown in Fig. 8(a1), the light blue stations marked with subway lines belong to *Line*1 and *Line*5 respectively. Since *Line*1 and *Line*5 are located close to each other, and passengers travel closely between these stations. In reality, the traffic department can merge *Line*1 and *Line*5 to reduce the disadvantages brought by the frequent transfer.

Moreover, the stations in the same community can be distant to each other. In Fig. 8(a), the stations circled in the pink dotted box belong to the same community as further stations both in the morning period (08:00–09:00) and evening period (18:00–19:00). Because the two stations are surrounded by the wildlife park and transportation hub stations, which play an important role in gathering the crowd. So they are more closely related to other subway stations. For such stations gathering passenger flows, billboards can be added inside, and rest areas or shops can be built to promote the economic benefits of subway departments. Another example is shown in Fig. 8(a2), v_{183} (Xiangyin Road) is light green instead of the same color as the adjacent stations (i.e., yellow). However, v_{183} belongs to the same community as distant v_{13} (People's Square). The reason is that passengers often commute between v_{183} and v_{13}. More specifically, v_{13} is mainly surrounded by office and commercial areas, while v_{183} is located in residential areas. Fig. 9 indicates the inbound and outbound traffic of v_{183} and v_{13} on Friday. There are large inbound traffic in the morning and remarkable outbound traffic in the afternoon in v_{183}. On the contrary, in v_{13}, the outbound traffic is impressive in the morning and the inbound traffic is large in the afternoon. What is more, v_{13} is located in the downtown in Shanghai, where residents shop and entertain, so the passenger flows into the station at night are also large. For these stations which are not adjacent but in the same community (e.g., v_{183} and v_{13}), the transport operations may formulate preferential policies. If passengers only get on and off in these stations, the ticket price can be relatively lower. At the same time, the policy can increase passenger flows to improve the economic benefits of the stores around the stations.

There are changeable community structures for a fraction of stations among morning and evening peaks. More specifically, in the box dotted with black lines of Fig. 8(a), these stations belong to the same community as the surrounding stations in the evening period (18:00–19:00). However, these stations are in different communities in the morning period (08:00–09:00). Another example is that a part of stations in *Line*3 are in a community alone in Fig. 8(b1). Moreover, the rest of stations in *Line*3 and stations in *Line*1 belong to the same community in the morning period (08:00–09:00). While in Fig. 8(b2), the stations in the same subway line are almost in the same community in the evening period (18:00–19:00). Based on the geography in Chongqing, there are bus stations, railway stations, airports and other transportation hubs near the dots colored with red. Passengers who will travel in the morning generally live nearby in advance, so there is no close contact between these and other stations. So these stations belong to a separate community during the morning rush hour. Therefore, traffic managers can adopt different management strategies for these stations during the morning and evening peaks.

The intersection of two subway lines plays a significant role in transportation and economy. Therefore, we analyze the intersections in Fig. 8(b1). Taking the green station with a black stroke as an example, it can belong to *Line*3 or *Line*6 in theory. However, it is green but not blue, showing that the transit station has more link with *Line*6 at that time. Therefore, around the transit station, *Line*6 is more important than *Line*1. The results show that the intersection of two subway lines can identify the importance of the two lines around the transit stations. Additionally, it proves the paramount significance of transit stations for the two subway lines, so traffic workers can add security personnel to the transit stations to strengthen the management.

On the basis of the above analysis, in the time dimension, community structures change constantly on the whole, but some stations always exist in the same cluster. It demonstrates that the urban system can be divided into certain local systems to form a urban structure, so that operators can organize and manage the traffic system in different parts rather than by a lot of stations. In spatial dimension, this paper finds different components of mixed land use and the functional types of surrounding areas, so as to understand the distribution of urban functional areas. And it also identifies the inner potential relationship of stations and lines. Therefore, traffic managers can employ a fixed management mode for stations that are stable in the same community. Similarly, dynamic management mode should be adopted for stations with frequently changing communities. What is more, according to the urban regional functions and the relationship of stations and lines, public decision-makers are able to manage stations, build infrastructures and improve service strategies.

4. Discussion

This section presents the differences of travel behaviors in the two cities. To begin with, Fig. 2(a) and (b) in Shanghai show the flat passenger flows on weekends. However, Fig. 2(c) in Chongqing has relatively obvious peaks.

On the whole, the traffic networks in Shanghai are more varied and complex than those in Chongqing. For example, the community number of Shanghai is 23 while that of Chongqing is 4, shown in Fig. 8. What is more, the variability of communities in Shanghai is much more obvious in Fig. 6. That is because there are more passenger flows and stations in Shanghai.

Moreover, we find the structure and layout of cities can contribute to the differences of urban travel behaviors, which helps understand the distribution of urban areas. For example, Fig. 3 displays that the similarity of weekends and weekdays in Chongqing is around 0.70 while that in Shanghai is around 0.45. Especially, the correlation coefficient of weekdays and

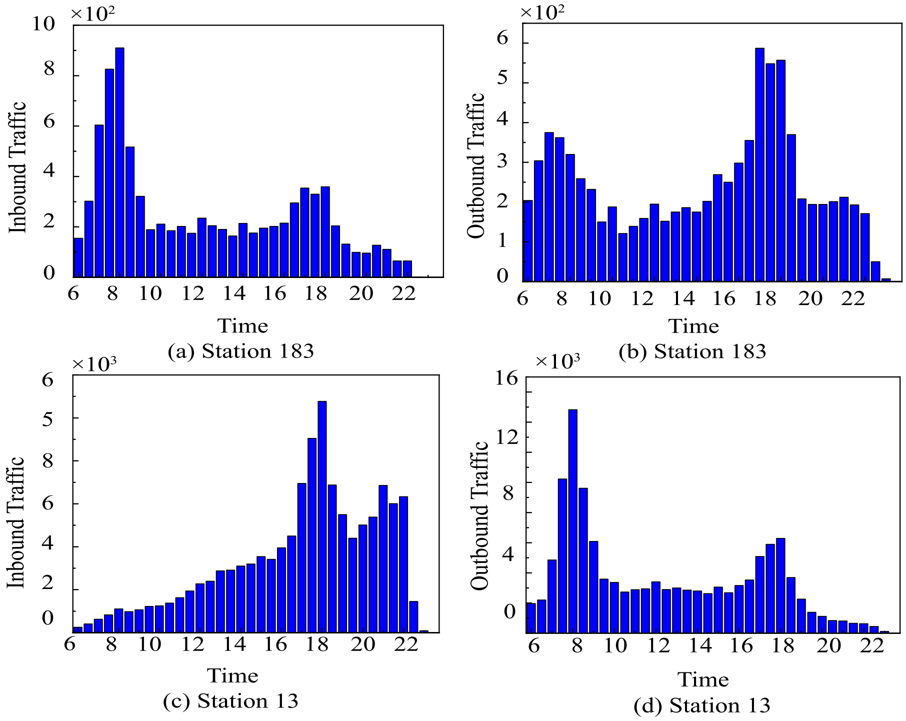


Fig. 9. The statistics of passenger flows of v_{183} (Xiangyin Road) and v_{13} (People's Square) in Shanghai. The horizontal axis represents the time period, and the vertical axis stands for the inbound or outbound traffic. (a)–(b) the inbound and outbound traffic of v_{183}, respectively. (c)–(d) the inbound and outbound traffic of v_{13}. The results show that passengers commute between the two stations.

Saturday in Chongqing reaches about 0.84. The topographies of the two cities is very different, shown in Fig. 1. In Shanghai, the terrain of Shanghai is flat with various means of transportation such as buses, taxis and bikes. However, Chongqing is divided into separate regions by mountains as well as the Yangtze and Jialing Rivers. Only long-distance buses and subway exist among these regions. Thus, the subway is the first choice to daily travel when it comes to service and the efficiency. Therefore, passengers in Chongqing are more willing to take the subway whether it is weekends or weekdays which causes the higher similarity.

5. Conclusion

Smart card data help researchers have a solid grasp of travel behaviors. Through the data in Shanghai and Chongqing, this paper analyzes the temporal–spatial characteristics to identify the urban travel behaviors. The main findings of this paper are as follows.

- The overall travel behaviors are varied, but the connection of certain stations is so strong that city planners can partition the urban system into several local systems, so as to organize the traffic in parts instead of by lots of stations.
- There are quite different travel patterns in working days and rest days, but similar mobility patterns in some adjacent stations.
- The structure and layout of cities play a significant role in the operation rules of cities, so transport agencies can control and manage the traffic system based on the specific city features.

To sum up, this paper reveals the temporal–spatial characteristics for transportation departments to develop urban facilities, improve the management efficiency and adjust service strategies. Moreover, the urban phenomena can be further explored in combination with datasets like population, economy and taxis. In the future, our goal is to connect multi-source data to further research travel characteristics, explore potential rules and make contributions to the development of smart cities.

CRediT authorship contribution statement

Yue Deng: Dataset curation, Investigation, Writing - original draft. **Jiaxin Wang:** Software, Validation. **Chao Gao:** Supervision, Conceptualization. **Xianghua Li:** Visualization, Writing - review & editing. **Zhen Wang:** Formal analysis, Project administration. **Xuelong Li:** Methodology, Conceptualization.

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.

Acknowledgments

This work was supported by the National Key R&D Program of China (No. 2019YFB2102300), National Natural Science Foundation of China (Nos. 61976181, U1803263, 11931015), Fok Ying-Tong Education Foundation, China (No. 171105), Key Technology Research and Development Program of Science and Technology Scientific and Technological Innovation Team of Shaanxi Province, China (No. 2020TD-013) and the Science and Technology Support Program of Guizhou, China (No. QKHZC2021YB531).

References

[1] Enhui Chen, Zhirui Ye, Chao Wang, Mingtao Xu, Subway passenger flow prediction for special events using smart card data, IEEE Trans. Intell. Transp. Syst. 21 (3) (2019) 1109–1120.
[2] Baoyu Hu, Shumin Feng, Jinyang Li, Hu Zhao, Statistical analysis of passenger-crowding in bus transport network of Harbin, Physica A 490 (2018) 426–438.
[3] Zulong Diao, Dafang Zhang, Xin Wang, Kun Xie, Shaoyao He, Xin Lu, Yanbiao Li, A hybrid model for short-term traffic volume prediction in massive transportation systems, IEEE Trans. Intell. Transp. Syst. 20 (3) (2018) 935–946.
[4] Jinjun Tang, Fang Liu, Yinhai Wang, Hua Wang, Uncovering urban human mobility from large scale taxi GPS data, Physica A 438 (2015) 140–153.
[5] Ilias Kalamaras, Alexandros Zamichos, Athanasios Salamanis, Anastasios Drosou, Dionysios D. Kehagias, Georgios Margaritis, Stavros Papadopoulos, Dimitrios Tzovaras, An interactive visual analytics platform for smart intelligent transportation systems management, IEEE Trans. Intell. Transp. Syst. 19 (2) (2017) 487–496.
[6] Zhaoyu Kou, Hua Cai, Understanding bike sharing travel patterns: An analysis of trip data from eight cities, Physica A 515 (2019) 785–797.
[7] Jianjun Chen, Yue Deng, Zhen Su, Songxin Wang, Chao Gao, Xianghua Li, Identifying multiple influential users based on the overlapping influence in multiplex networks, IEEE Access 7 (1) (2019) 156150–156159.
[8] Xiao Luo, Liang Dong, Yi Dou, Ning Zhang, Jingzheng Ren, Ye Li, Lu Sun, Shengyong Yao, Analysis on spatial-temporal features of taxis' emissions from big data informed travel patterns: a case of Shanghai, China, J. Cleaner Prod. 142 (2017) 926–935.
[9] Li Zhu, Fei Richard Yu, Yige Wang, Bin Ning, Tao Tang, Big data analytics in intelligent transportation systems: A survey, IEEE Trans. Intell. Transp. Syst. 20 (1) (2018) 383–398.
[10] Yi Fan, Fan Zhang, Shihong Jiang, Chao Gao, Zhanwei Du, Zhen Wang, Xianghua Li, Dynamic robustness analysis for subway network with spatiotemporal characteristic of passenger flow, IEEE Access 8 (2020) 45544–45555.
[11] Zhouhua Peng, Jun Wang, Dan Wang, Qing-Long Han, An overview of recent advances in coordinated control of multiple autonomous surface vehicles, IEEE Trans. Ind. Inf. 17 (2) (2020) 732–745.
[12] Kang Liu, Ling Yin, Zhanwu Ma, Fan Zhang, Juanjuan Zhao, Investigating physical encounters of individuals in urban metro systems with large-scale smart card data, Physica A 545 (2020) 123398.
[13] Shaoying Li, Dijiang Lyu, Xiaoping Liu, Zhangzhi Tan, Feng Gao, Guanping Huang, Zhifeng Wu, The varying patterns of rail transit ridership and their relationships with fine-scale built environment factors: Big data analytics from Guangzhou, Cities 99 (2020) 102580.
[14] Chen Zhong, Ed Manley, Stefan Müller Arisona, Michael Batty, Gerhard Schmitt, Measuring variability of mobility patterns from multiday smart-card data, J. Comput. Sci. 9 (2015) 125–130.
[15] K Mohamed, Etienne Côme, Latifa Oukhellou, Michel Verleysen, Clustering smart card data for urban mobility analysis, IEEE Trans. Intell. Transp. Syst. 18 (3) (2016) 712–728.
[16] Boris Chidlovskii, Mining smart card data for travellers mini activities, IEEE Trans. Intell. Transp. Syst. 19 (11) (2018) 3676–3685.
[17] Mohammad Sajjad Ghaemi, Bruno Agard, Martin Trépanier, Vahid Partovi Nia, A visual segmentation method for temporal smart card data, Transp. A 13 (5) (2017) 381–404.
[18] Juanjuan Zhao, Qiang Qu, Fan Zhang, Chengzhong Xu, Siyuan Liu, Spatio-temporal analysis of passenger travel patterns in massive smart card data, IEEE Trans. Intell. Transp. Syst. 18 (11) (2017) 3135–3146.
[19] Fang Zong, Yongda Tian, Yanan He, Jinjun Tang, Jianyu Lv, Trip destination prediction based on multi-day GPS data, Physica A 515 (2019) 258–269.
[20] Gabriel Goulet-Langlois, Haris N Koutsopoulos, Zhan Zhao, Jinhua Zhao, Measuring regularity of individual travel patterns, IEEE Trans. Intell. Transp. Syst. 19 (5) (2017) 1583–1592.
[21] Yu Cui, Qing He, Alireza Khani, Travel behavior classification: an approach with social network and deep learning, Transp. Res. Rec. 2672 (47) (2018) 68–80.
[22] Chunyu Wang, Fan Zhang, Yue Deng, Chao Gao, Xianghua Li, Zhen Wang, An adaptive population control framework for ACO-based community detection, Chaos Solitons Fractals 138 (2020) 109886.
[23] Pengfei Lin, Jiancheng Weng, Yu Fu, Dimitrios Alivanistos, Baocai Yin, Study on the topology and dynamics of the rail transit network based on automatic fare collection data, Physica A 545 (2020) 123538.
[24] Ekki D. Kreutzberger, Distance and time in intermodal goods transport networks in Europe: A generic approach, Transp. Res. A 42 (7) (2008) 973–993.
[25] Javier Del Ser, Javier J. Sanchez-Medina, Eleni I. Vlahogianni, Introduction to the special issue on online learning for big-data driven transportation and mobility, IEEE Trans. Intell. Transp. Syst. 20 (12) (2019) 4621–4623.
[26] Youcef Djenouri, Asma Belhadi, Jerry Chun-Wei Lin, Djamel Djenouri, Alberto Cano, A survey on urban traffic anomalies detection algorithms, IEEE Access 7 (2019) 12192–12205.
[27] Wei Wang, Kaiquan Cai, Wenbo Du, Xin Wu, Lu Carol Tong, Xi Zhu, Xianbin Cao, Analysis of the chinese railway system as a complex network, Chaos Solitons Fractals 130 (2020) 109408.
[28] Jiaoe Wang, Fengjun Jin, Huihui Mo, Fahui Wang, Spatiotemporal evolution of Chinas railway network in the 20th century: An accessibility approach, Transp. Res. A 43 (8) (2009) 765–778.
[29] Chunyu Wang, Yue Deng, Xianghua Li, Jianjun Chen, Chao Gao, Dynamic community detection based on a label-based swarm intelligence, IEEE Access 7 (1) (2019) 161641–161653.

[30] Le Minh Kieu, Ashish Bhaskar, Edward Chung, Passenger segmentation using smart card data, IEEE Trans. Intell. Transp. Syst. 16 (3) (2014) 1537–1548.
[31] Xiaolei Ma, Congcong Liu, Huimin Wen, Yunpeng Wang, Yao-Jan Wu, Understanding commuting patterns using transit smart card data, J. Transp. Geogr. 58 (2017) 135–145.
[32] Chen Zhong, Michael Batty, Ed Manley, Jiaqiu Wang, Zijia Wang, Feng Chen, Gerhard Schmitt, Variability in regularity: Mining temporal mobility patterns in London, Singapore and Beijing using smart-card data, PLOS ONE 11 (2) (2016) e0149222.
[33] Atizaz Ali, Jooyoung Kim, Seungjae Lee, Travel behavior analysis using smart card data, KSCE J. Civil Eng. 20 (4) (2016) 1532–1539.
[34] Li Zhang, Qianyun Nie, Yalei Dai, Shisheng Zhu, Jinbao Wang, Wei Wang, Xiaobo Tan, Peng Zhang, Jianbo Li, Diatomological mapping of water bodies in Chongqing section of the Yangtze River and Jialing River, Int. J. Legal Med. 134 (2020) 1375–1385.
[35] Feng Xia, Jinzhong Wang, Xiangjie Kong, Da Zhang, Zhibo Wang, Ranking station importance with human mobility patterns using subway network datasets, IEEE Trans. Intell. Transp. Syst. 21 (7) (2019) 2840–2852.
[36] Martin Rosvall, Carl T. Bergstrom, Maps of random walks on complex networks reveal community structure, Proc. Natl. Acad. Sci. 105 (4) (2008) 1118–1123.
[37] Chao Gao, Yi Fan, Shihong Jiang, Yue Deng, Jiming Liu, Xianghua Li, Dynamic robustness analysis of a two-layer rail transit network model, IEEE Trans. Intell. Transp. Syst. (2021) http://dx.doi.org/10.1109/TITS.2021.3058185.
[38] Min Ouyang, Lijing Zhao, Liu Hong, Zhezhe Pan, Comparisons of complex network based models and real train flow model to analyze chinese railway vulnerability, Reliab. Eng. Syst. Saf. 123 (2014) 38–46.
[39] Michele Guida, Funaro Maria, Topology of the Italian airport network: A scale-free small-world network with a fractal structure? Chaos Solitons Fractals 31 (3) (2007) 527–536.
[40] Xiangxiang Zeng, Wen Wang, Cong Chen, Gary G. Yen, A consensus community-based particle swarm optimization for dynamic community detection, IEEE Trans. Cybern. 50 (6) (2019) 2502–2513.
[41] Ernesto Estrada, Jesús Gómez-Gardeñes, Communicability reveals a transition to coordinated behavior in multiplex networks, Phys. Rev. E 89 (4) (2014) 042819.
[42] Zhen Wang, Chunyu Wang, Xianghua Li, Chao Gao, Xuelong Li, Junyou Zhu, Evolutionary markov dynamics for network community detection, IEEE Trans. Knowl. Data Eng. (2020) http://dx.doi.org/10.1109/TKDE.2020.2997043.
[43] Etienne Gael Tajeuna, Mohamed Bouguessa, Shengrui Wang, Modeling and predicting community structure changes in time-evolving social networks, IEEE Trans. Knowl. Data Eng. 31 (6) (2018) 1166–1180.
[44] Maoguo Gong, Qing Cai, Xiaowei Chen, Lijia Ma, Complex network clustering by multiobjective discrete particle swarm optimization based on decomposition, IEEE Trans. Evol. Comput. 18 (1) (2013) 82–97.