

Article

# Uncovering Spatio-temporal Travel Patterns Using a Tensor-Based Model from Metro Smart Card Data in Shenzhen, China

Jinjun Tang <sup>1</sup>, Xiaolu Wang <sup>1</sup>, Fang Zong <sup>2,\*</sup> and Zheng Hu <sup>3,\*</sup>

<sup>1</sup> Smart Transport Key Laboratory of Hunan Province, School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China; Jinjuntang@csu.edu.cn (J.T.); xiaoluwang@csu.edu.cn (X.W.)

<sup>2</sup> College of Transportation, Jilin University, Changchun 130012, China

<sup>3</sup> School of Intelligent transportation, Hunan Communication Engineering Polytechnic, Changsha 410132, China

\* Correspondence: zongfang@jlu.edu.cn (F.Z.); jameshu9941@163.com (Z.H.)

Received: 8 December 2019; Accepted: 14 February 2020; Published: 17 February 2020



**Abstract:** Individual mobility patterns are an important factor in urban traffic planning and traffic flow forecasting. How to understand the spatio-temporal distribution of passengers deeply and accurately, so as to provide theoretical support for the planning and operation of the metro network, is an urgent issue of wide concern. In this paper, we applied NCP decomposition to uncover the characteristics of travel patterns from temporal and spatial dimensions in the metro network of Shenzhen City. Utilizing matrix factorization and correlation analysis, we extracted several stable components from the collective mobility and find that the departure and arrival mobility patterns have different characteristics in both the temporal and spatial dimension. According to the point of interest (POI) data in the Shenzhen City, the function attributes of the station are identified and then we found that the spatial distribution characteristics of different patterns are different. We explored the distribution of travel time classified according to the spatio-temporal characteristics of stable patterns. The proposed method can decompose stable travel patterns from the collective mobility and the results in this study can help us to better understand different mobility patterns in both spatial and temporal dimensions.

**Keywords:** tensor decomposition; urban mobility; spatial and temporal characteristics; metro network; point of interest

## 1. Introduction

Urban transportation is an important support for the economic development of the cities, which profoundly affects the living quality of citizens, the efficiency of logistics, as well as other factors related with transport. With the increase of population and the continuous expansion of the urban city, urban traffic is increasing rapidly and this massive traffic puts enormous pressure on road networks. Traffic congestion has become more serious due to the imbalance between infrastructure capacity and travel demands. Due to their high-capacity, speed and reliability, metro systems can effectively alleviate the problems of urban traffic congestion. Therefore, how to understand the spatial and temporal distribution of passengers deeply and accurately, so as to provide theoretical support for the planning and operation of the metro network, is an urgent issue of wide concern.

As an essential part of the public transportation systems in big cities, metro systems can address the massive travel demand of citizens thanks to their low price and arrival accuracy. Recently, a large amount of works have focused on passenger forecasting [1–3], network planning [4] and human travel

analysis [5–7] in metro systems. The fast development of information technology has enabled researchers to obtain data reflecting travels through various means, such as GPS [8,9], mobile phones [10], and smart card systems [11–14]. The emergence of large-scale data brings us new opportunities to better understand the characteristic of individual movement. Specifically, some scholars have explored human activity patterns by analyzing the spatial-temporal characteristic of travel [15–19]. Zhang et al. [15] applied a density-based method for identifying so-called temporal areas of interest (TAI), and found that there are four major types of TAIs on weekdays, namely work-like, morning, afternoon and nightlife TAIs, and three on weekends, namely work-like, day activity and nightlife TAIs. Using visualization techniques, Sun et al. [16] demonstrated the spatial and temporal distributions of passenger flows in a holistic manner, as well as the flow directional imbalances. Zhao et al. [17] used statistical-based and unsupervised clustering-based methods to understand the hidden regularities and anomalies of travel patterns and classify passengers in terms of the similarity of their travel patterns.

Furthermore, the prediction of passengers' future trajectories [20–24] is also an important issue in transit systems. Yang et al. [20] argued that future movements of different types of groups can be predicted with high confidence based on previous records. Truong et al. [21] effectively predicted the in- and outflow of passengers at a station over time based on the patterns of passenger flow with respect to time and stations. In the study [22], Shanghai was taken as a case study to discuss the location choice of after-work activities, passengers are more likely to choose a station even closer to home in terms of network distance. Based on previous research, we can draw the conclusion that passengers' behavior has a high potential to be recognized, categorized as well as predicted [24]. Besides, the study of individual mobility patterns has been paid much attention from researchers [25–30]. Sun et al. [25] utilized a data-driven approach to characterize the collective mobility patterns from high dimensional structured datasets. Taking Shanghai as a case study, Du et al. [26] adopted a dual-perspective on passengers' activity patterns to investigate the spatial and temporal dimensions of individuals travelling decision making. However, the appearance of new metro lines may significantly affect individual mobility in the metro network. Kim et al. [27] studied the changes of metro passenger flow and travel time due to the operation of new metro line. Liu et al. [29] evaluated the impact of new lines on passenger flow of existing stations in urban metro system and found new transfer stations attracted more passengers than other existing stations.

Although many scholars have studied the passenger behavior of metro systems, there still exist some challenges in current works. The data collected from smart cards in metro systems is huge and complicated, and traditional methods find it difficult to deal with the issues concerning the multi-dimensional features of original data. Principal component analysis (PCA) and non-negative matrix factorization (NMF) [31] usually expand the original data into a two-dimensional matrix in the data process, which loses the structure information of the original data and the solution often appears unreasonable. Given its strength in retrieving and storing information from large datasets, tensor decomposition [32] has also attracted more and more attention in the field of transportation data analysis. By using tensor storage strategy, the smart card data from metro systems can retain the original structure information, so we can analyze the information of different dimensions at the same time. Some scholars have proposed the use of non-negative CANDECOMP/PARAFAC (NCP) factorization [33,34], which adds non-negative constraints to the tensor decomposition process to reduce the ambiguity of the decomposition results.

Negative values are difficult to explain in the matrix elements of tensor decomposition, so we apply NCP decomposition [35] with alternating Poisson regression on a two-dimensional tensor constructed from smart card data to analyze the spatial-temporal characteristics of passengers' travel. The urban mobility discussed in this study refers to a specific travel mode, and this paper is dedicated to explore passengers' travel patterns based on smart card data from an urban metro network. We also want to introduce a framework using a tensor-based model to understand travel behaviors. The travel behavior discussed in this paper refers to the trips of passengers with different travel purposes using the urban metro network. We applied NCP decomposition to uncover the characteristic of travel patterns from temporal and spatial dimension in the metro network of Shenzhen City. Then, utilizing correlation

analysis, we establish a correlation among different days and extract some stable components from the individual mobility. Excluding the occasional factors such as weather and activity, the characteristics of the stable mobility patterns can be explored both in the temporal and spatial dimension. Next, based on the POI data collected in Shenzhen City, the function of the station is set and then we further analyze the spatial distribution characteristics of different travel patterns. Furthermore, we explored the distribution of travel time classified according to the distribution characteristics based on the three basic stable patterns in the temporal and spatial dimension. The individual records are classified into different patterns and the most likely travel time is analyzed by category.

The remainder of this study is organized as follows: In Section 2, we describe the data sources for the study. In Section 3, we introduce the framework of probabilistic factorization based on NCP decomposition and construct the tensor model. In Section 4, we analyze the spatial and temporal characteristics of the daily collective mobility patterns based on the tensor decomposition results. In Section 5, we use correlation analysis to identify stable basic patterns for the spatio-temporal distribution under the stable basic patterns. Finally, Section 6 summarizes the key findings.

## 2. Data Source Collection

This section briefly introduces the metro network, the smart card data collected during three weeks and the point of interest (POI) information of Shenzhen City. Smart card data reflect urban passengers' travel behavior. Meanwhile, POI information reflects land use in a city.

### 2.1. Smart Card Dataset

Shenzhen City, located in the south of China, is one of the fast-growing and densely populated metropolitan cities in the world. The range of the general study area in Shenzhen City is limited in longitude [113.684206, 114.658294] and altitude [22.243608, 22.862324]. The metro smart card data were collected from the automatic fare collection (AFC) system. The dataset covers a population of millions of passengers for three weeks from 3 April 2017 to 23 April 2017, which include the codes of entry and exit metro stations and the corresponding time when getting in and out of the stations. However, there are some problems such as missing information, invalid or repeated card. As for the structure information of metro system, it has a total of 167 stations and eight metro lines (including Shenzhen Metro Lines 1–5, Line 7, Line 9, and Line 11), shown in Figure 1. Furthermore, it is essential to match the latitude and longitude of the metro stations with the data analysis results, the latitude and longitude information of the metro station were also collected from the open source map.

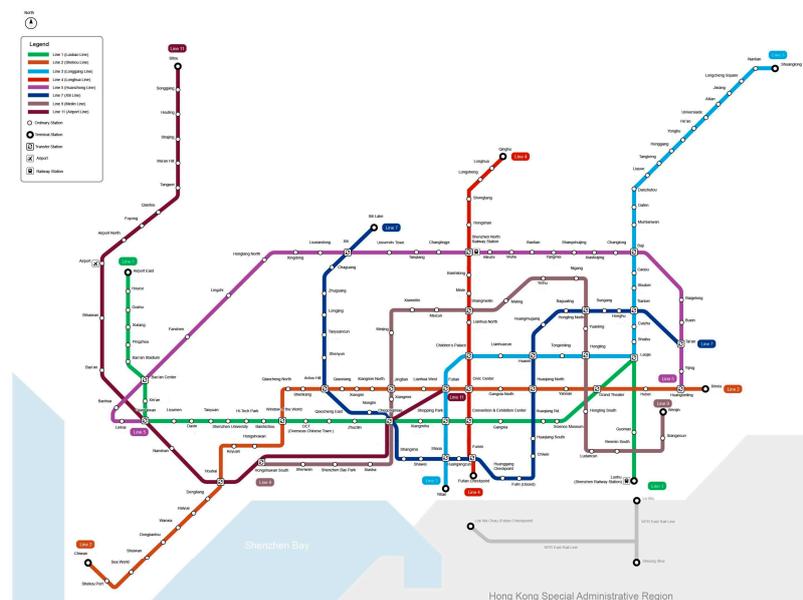


Figure 1. The metro network in Shenzhen City.

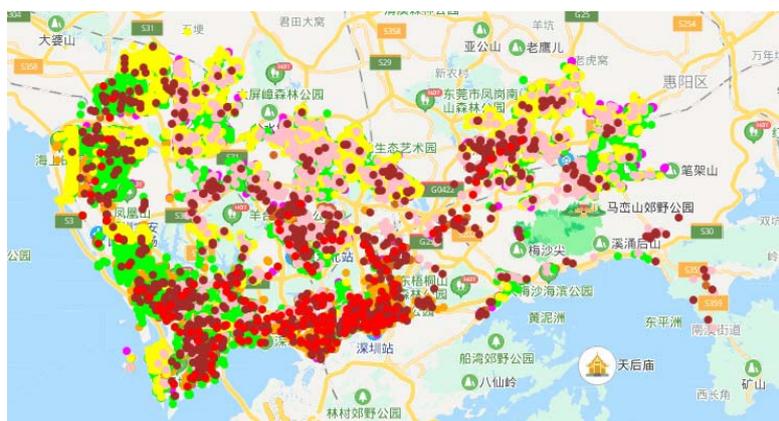
## 2.2. POI Dataset

The travel behaviors of passengers are closely linked with the type of land use in the area around the metro stations. Urban POI data refers to all geographical entity objects that can be abstracted into points without area or volume which are closely related to people's daily life and can reflect the functional attribution of urban areas. The POIs of cities are distributed with different density around arrival stations. The places with more attractive points of interest are generally the locations with higher possibility to travel. In this paper, we extract the POI information of Shenzhen City to identify the characteristics of the urban functional area, and then use this information to analyze the spatial patterns of the travels. At present, application programming interfaces (APIs) such as Baidu Map, Google Maps and Gaode Map can be used to obtain the POI information. We utilize the requests module in Python to obtain the POI information of Shenzhen City provided by the Baidu map development platform [36]. The platform divides the POI information of cities into 19 categories and 140 subcategories according to default function attributes. As we cannot collect the travel purpose from the data in smart card, the travel behavior we introduced in this study is the combination of daily travel trips with different purpose under metro network. We divide different types of POI into seven major categories: residential area (represented by R), corporate companies (indicated by C), transportation hub (indicated by T), leisure and entertainment (indicated by L), medical institution (indicated by M), education institution (indicated by E) and gourmet restaurant (indicated by F). The specific contents of the seven categories of POIs are shown in Table 1.

**Table 1.** POI category classification.

Category	Contents
residential area (R)	Residential area, dormitory.
corporate companies (C)	Company, factory.
leisure and entertainment (L)	Cinema, KTV, theater, shopping center, department store, etc.
transportation hub (T)	Airports, railway stations, bus stations, ports.
medical institution (M)	General hospitals, specialist hospitals, clinics, etc.
education institution (E)	Colleges, elementary schools, kindergartens, adult education, etc.
gourmet restaurant (F)	Chinese restaurants, foreign restaurants, cafes, etc.

Records within each category include category attributes, POI names and the longitude and latitude information. Figure 2 shows the seven major categories of POI information on the map.



**Figure 2.** POI data points. (R: orange, C: yellow, T: green, L: purple, M: brown, E: blue, F: chocolate color, metro stations: red.).

## 3. Methodology

As the amount of data collected from smart cards in metro systems is huge, and the complexity of the data increases with the dimensions of its contents, it is difficult for traditional and basic statistical

methods to deal with this data source. The tensor-based model we applied in this study can extract multi-dimensional features from an original dataset. The passengers' spatio-temporal characteristics of different travel patterns can be quantified based on tensor decomposition. Tensor decomposition is a high-dimensional generalization of singular value decomposition (SVD) and principal component analysis (PCA), which includes two types of tensor decomposition techniques: Tucker decomposition and CP decomposition. Tucker decomposition is a high-order principal component analysis method that decomposes a tensor into a core tensor multiplied by a corresponding factor matrix along each dimension. The CP decomposition decomposes a high-dimensional tensor into a sum of component rank-one tensors. The rank one tensor is a special tensor type. If an N-way tensor can be expressed by the outer product of N vectors, then this is a rank one tensor. In this study, we apply the CP decomposition to analyze the spatial-temporal characteristics of urban travels.

The CP decomposition process can be explained by a third-order tensor  $X \in R^{I \times J \times K}$  show in Figure 3, where  $I, J, K$  are the dimension along the  $p$ th way ( $p \in \{1, 2, 3\}$ ). The approximate decomposition of a three-dimensional tensor  $X$  can be defined [32] as follows:

$$X \approx \sum_{r=1}^R a_r \circ b_r \circ c_r, \quad (1)$$

where  $R$  is a positive integer and  $a_r \in R^I, b_r \in R^J, c_r \in R^K, r = 1, \dots, R$ . The symbol  $\circ$  means vector outer product and  $a_r \circ b_r \circ c_r$  is a rank-one tensor. The CP decomposition of a tensor is a sum of  $R$  rank-one component tensors (we call  $R$  the rank of tensor  $X$ ).

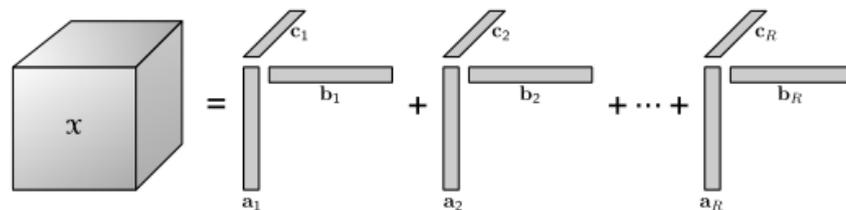


Figure 3. Approximate decomposition of a three-dimensional tensor  $X$ .

The element of the tensor  $X$  can be expressed as:

$$x_{ijk} \approx \sum_{r=1}^R a_{ir} b_{jr} c_{kr}, \quad (2)$$

where  $i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, K$ , we use  $i, j, k$  to denote the index of each dimension and we denote by  $x_{ijk}$  the value of the element of the tensor  $X$ .

Fibers [32] are the higher order analogue of matrix rows and columns. A fiber is defined by fixing every index but one. Third-order tensors have column, row, and tube fibers, denoted by  $X_{:jk}$ ,  $X_{ik}$  and  $X_{ij}$ , respectively. Slices are two-dimensional sections of a tensor, defined by fixing all but two indices. Third-order tensors have the horizontal, lateral, and frontal slides, denoted by  $X_{i::}$ ,  $X_{:j}$  and  $X_{::k}$ , respectively. The  $k$ th frontal slice  $X_{::k}$  of a third-order tensor may also be denoted as  $X_{(k)}$ .

The factor matrix refers to the combination of the vectors from the rank-one component,  $A = [a_1 a_2 \dots a_R]$ ,  $B = [b_1 b_2 \dots b_R]$ ,  $C = [c_1 c_2 \dots c_R]$ . Matricization is the process of recording the elements of a multi-dimension tensor into a matrix. Using the factor matrix, the three matricized versions on one per dimension are:

$$X_{(1)} \approx A(C \circ B)^T, \quad (3)$$

$$X_{(2)} \approx B(C \circ A)^T, \quad (4)$$

$$X_{(3)} \approx C(B \circ A)^T, \quad (5)$$

where  $\odot$  denotes the Khatri-Rao product. Given matrices  $A \in R_{I \times K}$  and  $B \in R_{J \times K}$ , their Khatri-Rao product is denoted by  $A \odot B$ . The result is a matrix of size  $(IJ) \times K$ .

The columns of  $A$ ,  $B$  and  $C$  are normalized to length one with the weights of lambda. Where  $\lambda \in R^R$ ,  $X$  can be approximately decomposed into Equation (6):

$$X \approx \sum_{r=1}^R \lambda_r a_r \circ b_r \circ c_r = [\Lambda; A, B, C], \quad (6)$$

The three-dimensional tensor can be extended to N-dimensional tensor [32] and written as:

$$X \approx \sum_{r=1}^R \lambda_r a_r^{(1)} \circ a_r^{(2)} \circ \dots \circ a_r^{(N)} = [\Lambda; A^{(1)}, A^{(2)}, \dots, A^{(N)}], \quad (7)$$

However, the matrix elements of tensor decomposition can be positive or negative. The existence of negative values in the decomposition results is correct from the point of view of mathematical operation, but negative values are difficult to explain in practical problems. We apply NCP decomposition [35] with alternating Poisson regression to uncover the characteristic of travel patterns in the Shenzhen metro network. The method we applied is appropriate for nonnegative integer values used in smart card data, which adds non-negative constraints to the tensor decomposition process.

We have extracted a total of 167 metro stations from eight metro lines, and the original data collected from 5:00 to 23:00 is divided into 18 time periods of one hour. Meanwhile, the passenger volumes were collected at entry and exit stations in the metro system at different time periods. Therefore, the departure station tensor  $Mo = (\text{departure station, time slot})$  and arrival station tensor  $Md = (\text{arrival station, time slot})$  are constructed, and its collective mode is  $M \in R_+^{167 \times 18}$ . The collective mobility is decomposed into a linear combination of several basic modes of movement by NCP decomposition:

$$M \approx \sum_{r=1}^R \lambda_r a_r \circ b_r, \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_R, \quad (8)$$

Equation (8) decomposes the matrix  $M \in R_+^{m \times n}$  into  $R$  components, each component  $i$  includes one eigenvalue and two vectors,  $a_r \in R_+^{m \times 1}$  and  $b_r \in R_+^{n \times 1}$ . In this study, the collective mobility patterns  $M \in R_+^{167 \times 18}$  are decomposed into several basic patterns.  $b_r \in R_+^{18 \times 1}$  represents the temporal distribution.  $a_r \in R_+^{167 \times 1}$  represents the corresponding spatial information, which indicates the probability of arriving or leaving the metro stations. Therefore, we obtain the components representing the basic mobility patterns from the collective mobility patterns through NCP decomposition.

Taking the departure station tensor  $Mo = (\text{departure station, time slot})$  in a specific day as example, its mobility patterns can be expressed as  $Mo \in R_+^{167 \times 18}$ . We define the NCP decomposition of the departure station tensor  $Mo$  as Equation (9):

$$M_O \approx \sum_{r=1}^R \lambda_r a_r \circ b_r, \quad (9)$$

This can be further expressed as Equation (10):

$$M_o(z, t) \approx \sum_{r=1}^R \lambda_r a_r(z, 1) \circ b_r(1, t), \quad (10)$$

#### 4. The Spatial-Temporal Characteristics Analysis of Travel Patterns

Using the tensor toolbox [35] in MATLAB, we performed NCP decomposition to explore the number of basic mobility patterns. Through repeated experiment and analysis of tensor composition

results, we found that when the number of basic mobility patterns  $R = 3$ , the extracted travel modes expressed obvious differences in spatio-temporal dimension. If we set  $R = 4$  or more, the travel modes extracted from tensor decomposition are unreasonable and cannot be in accordance with daily travel regularity. The three basic mobility patterns have been confirmed by scholars [31,37] to be appropriate for understanding the spatial-temporal characteristics of travel patterns.

The collective model is decomposed into three components, and the characteristics of travel behavior of the three components in the temporal dimension show in Figure 4a,c,e. Departure-1 indicates the temporal distribution in the morning peak hours of passengers traveling in Shenzhen metro system during the weekday. Departure-2 indicates the temporal distribution in the evening peak hours during the weekday. Departure-3 indicates the temporal distribution of the daily hours in weekday. All the results shown in Figure 4 are the average calculating results from three weeks.

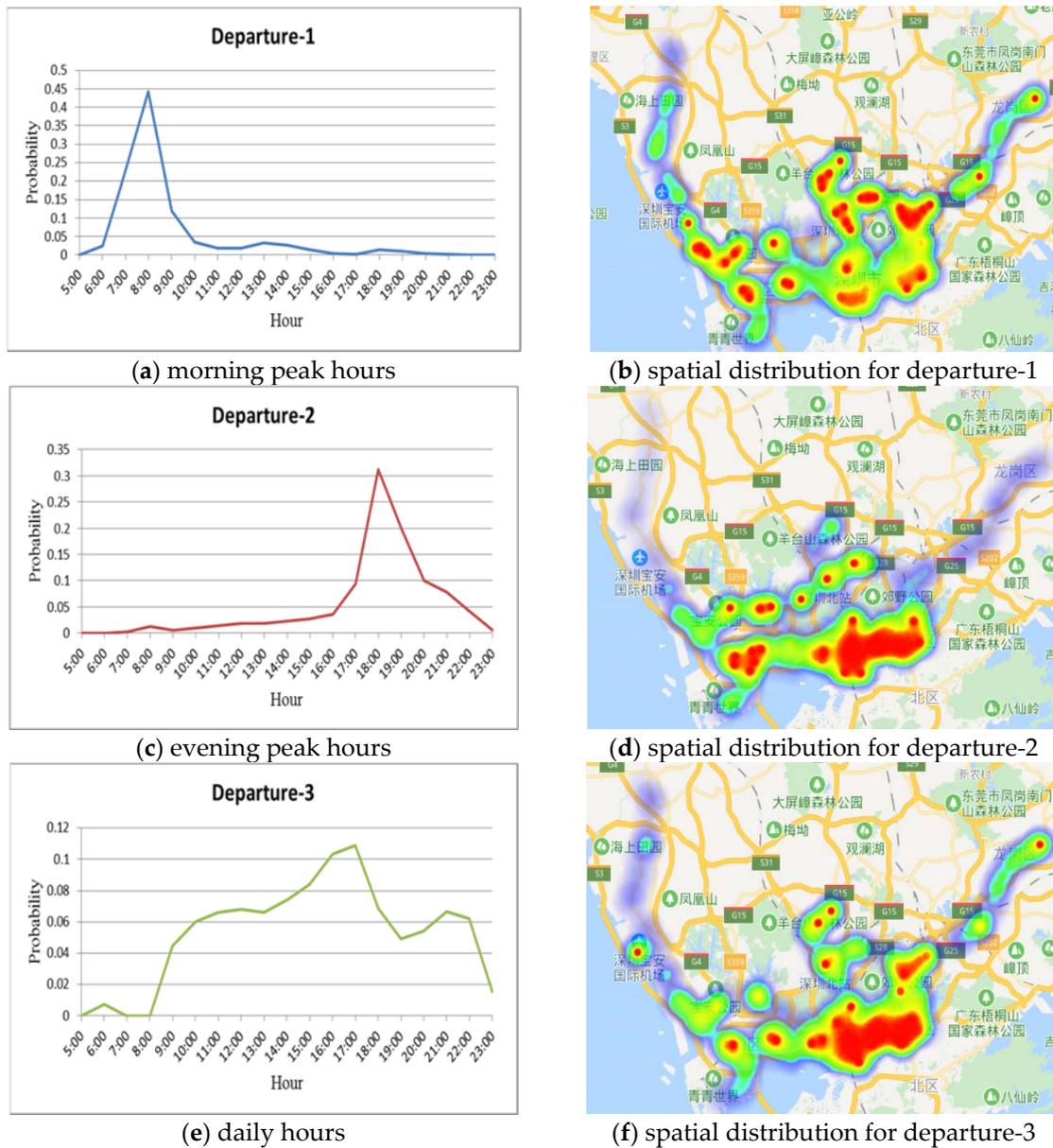


Figure 4. Temporal and spatial distribution of three components.

According to the decomposition result, the spatial distribution characteristics of travel behavior corresponding to the temporal distribution of the three components can be analyzed simultaneously.

In the spatial dimension, we use a thermodynamic chart to show the spatial distribution of each metro station. The right part of Figure 4b,d,f shows the spatial distribution of travel behavior.

The Shenzhen metro network covers most areas from the city center to the suburbs. The network structure of the route makes travel in Shenzhen City become more convenient. The results in Figure 4 show that Departure-1 has obvious peak traffic during the morning period, the peak hours range between 7:30 and 9:00, and the value reaches peak at 8:00. At this time, passenger trip is mainly distributed in the surrounding areas and scattered in the downtown areas of the city, specifically Pengzhou, Wuhe, Buji, Qinghu and Minzhi are the main stations. According to the extracted POI data of Shenzhen City, Figure 5 shows that most of the residential areas in Shenzhen are located in the surrounding areas and scattered downtown areas of the city. We can find that traffic travel is mainly distributed in residential areas during the morning peak hours on weekdays.

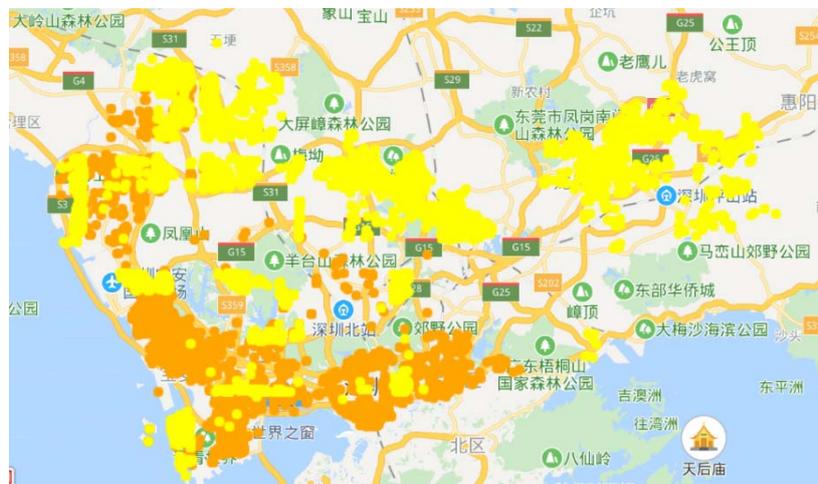


Figure 5. POI distribution of residential area (orange) and corporate (yellow).

Departure-2 expresses peak passenger volume during the evening hours, and the peak hours range between 17:00 and 20:00, and the value reaches maximum at 18:00. At this time, traffic travel is mainly distributed in the downtown area of the city, with Chegongmiao, Shenda, Gaoxinyuan, Convention and Exhibition Center and Huaqiang North being the main stations. According to the extracted POI data of Shenzhen city, Figure 5 shows that companies are mostly distributed around the center of the city. We can find that traffic travel is mainly distributed in companies' areas during the evening peak hours on weekdays. So, on weekdays, passengers leave the residential area around the city during the morning peak hours, and return to the residential areas from the work area of the city center during the evening peak hours.

Departure-3 shows a significant decline in the proportion of passenger volume in each period, and there is no obvious peak volume. The urban travel is mainly distributed in the central area and scattered around city, which belongs to the daily transition mode between Departure-1 and Departure-2.

## 5. Analysis of Travel Characteristics Combined with POI Data

We have decomposed the departure station tensor  $M_o = (\text{departure station, time slot})$  into several basic patterns in Section 4. However, for the components of the basic patterns, some are stable, and others are occasional due to specific reasons (such as weather and activities). Utilizing correlation analysis, we establish correlation among different days and extract some stable components from the collective mobility to analyze the temporal distribution in stable basic mobility patterns in Section 5.1. Next, the functions of the station are set based on the POI data collected in Shenzhen City, and then we further analyze the spatial distribution characteristics in stable basic mobility patterns in Section 5.2. Furthermore, we classify individual records into different patterns and the travel time is analyzed for each category in Section 5.3.

### 5.1. Analysis of the Correlation Among Daily Travels

After constructing the departure station tensor  $Mo = (\text{departure station, time slot})$  and arrival station tensor  $Md = (\text{arrival station, time slot})$ , through the process of nonnegative CP decomposition, different mobility patterns are explored. Next, in order to compare the basic patterns of daily travels, it is necessary to establish correlations among different days so that we can examine the changes in passenger flow at the stations on different days. For all the stations, we used the average of passenger volume in one specific day (from Monday to Sunday) of several weeks, and use correlation coefficients to describe the similarity of each basic pattern. In order to establish correlation between day  $m$  and day  $n$ , the tensor decomposition results are as shown in Equations (11) and (12):

$$M^m \approx \sum_{i=1}^R \lambda_i^m a_i^m \circ b_i^m, \quad (11)$$

$$M^n \approx \sum_{j=1}^R \lambda_j^n a_j^n \circ b_j^n, \quad (12)$$

Then, we calculate the correlation coefficient of the spatial distribution of the  $i^{\text{th}}$  component of day  $m$  and the  $j^{\text{th}}$  component of day  $n$  by:

$$r_{ij}^{mn} = \frac{\sum_{p=1}^{167} (a_{ip}^m - \bar{a}_i^m)(a_{jp}^n - \bar{a}_j^n)}{\sqrt{\sum_{p=1}^{167} (a_{ip}^m - \bar{a}_i^m)^2 \cdot \sum_{p=1}^{167} (a_{jp}^n - \bar{a}_j^n)^2}}, \quad (13)$$

In this formula,  $a_{ip}^m$  denotes the  $p^{\text{th}}$  element of the  $i^{\text{th}}$  component of day  $m$ , and  $a_{jp}^n$  denotes the  $p^{\text{th}}$  element of the  $j^{\text{th}}$  component of day  $n$ . We set that if the correlation coefficient  $r_{ij}^{mn}$  is greater than the threshold, then the  $i^{\text{th}}$  component of day  $m$  and the  $j^{\text{th}}$  component of day  $n$  are considered to be the same basic pattern. If for all  $n$ , there is  $j$ , the correlation coefficient  $r_{ij}^{mn}$  is greater than the threshold, then the  $i^{\text{th}}$  component of day  $m$  and the  $j^{\text{th}}$  component of day  $n$  are considered to be a stable basic pattern. If, for all  $n$ , all  $j$  makes the correlation coefficient  $r_{ij}^{mn}$  less than the threshold, then the  $i^{\text{th}}$  component of day  $m$  and the  $j^{\text{th}}$  component of day  $n$  are considered to be an occasional basic pattern.

We decompose the departure station tensor  $Mo = (\text{departure station, time slot})$  and arrival station tensor  $Md = (\text{arrival station, time slot})$  constructed from the Shenzhen Metro smart card data for three weeks. The specific days (from Monday to Sunday) in the three weeks have similar spatial and temporal distribution based on tensor decomposition results. According to the correlation coefficient formula, the correlation coefficients between the components of average  $Mo$  and  $Md$  in specific days of three weeks were calculated respectively. Take Monday as an example,  $M_O^1$  and  $M_d^1$  are constructed from the average of passenger volume of Monday in three weeks. The calculation results are shown in the Tables A1 and A2 of Appendix A. The threshold of correlation coefficient is set at 0.8. The positions with correlation coefficient greater than 0.8 are marked in red in the tables.

According to the calculation results of the correlation coefficient, the stable patterns appear almost every day. We extract the data of the stable basic patterns of the departure station tensor  $Mo$  and the arrival station tensor  $Md$ , and take the average value of each component as the departure and arrival stable mobility pattern. Figure 6 shows the temporal distribution of stable departure and arrival patterns for Metro passengers on weekdays and weekends. We find the stable basic patterns of weekdays and weekends have different characteristics for travel departure and arrival.

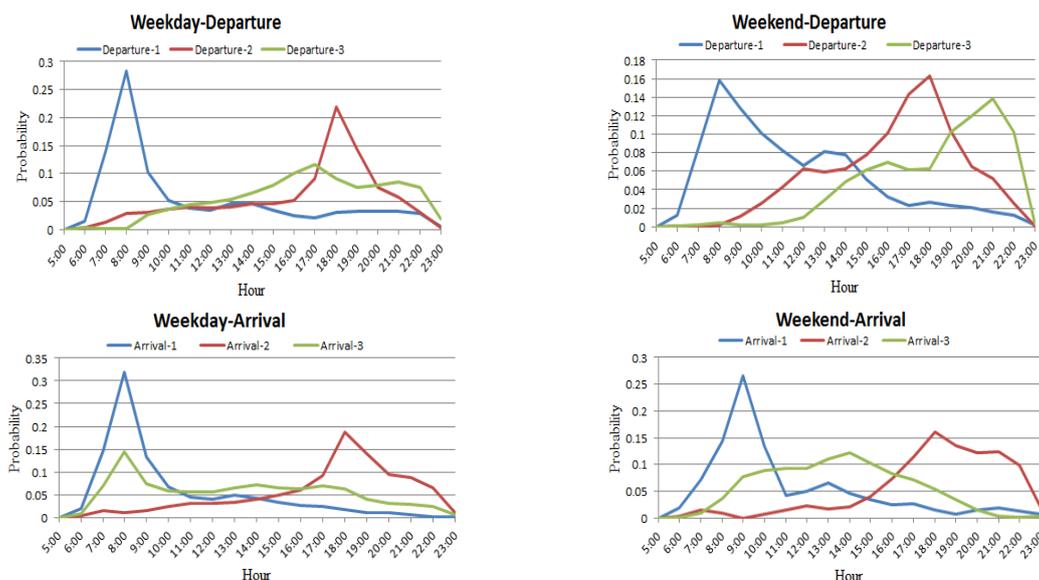


Figure 6. Temporal distribution in stable basic mobility patterns.

As can be seen from Figure 6, during the weekday, in the departure station, passenger volume reaches the peak at 8:00, the probability accounts for nearly 0.3. On the weekend, the peak volume is expressed at 8:00, but the proportion of the trip numbers greatly reduced. The trips are mainly concentrated during the 7:00–9:00 period. During the weekday, passenger volume in arrival station has obvious evening peak at 18:00. While trip in arrival station during weekends is mainly concentrated between 17:00–22:00, and the time is relatively scattered.

There are still some differences in the temporal distribution. During peak hours on weekdays, whether it is departing or arriving at the metro station, urban trips mainly occur at 8:00 am and 6:00 pm. The peak value is relatively high and the duration of the relatively high capacity is short. This means that passenger volume is generated in the morning and evening peak hours on weekdays, and a large number of people take the metro network as the first choice for commute in a short time. However, during the weekends, the peak hours appear slightly later than on weekdays. The peak value is lower and duration of the relatively high capacity is longer. Since passengers are not limited by rigid working timetables on weekends, the change of passenger flow over time is slower than that of weekdays.

## 5.2. Analysis of Travel Spatial Characteristics

The analysis of the temporal distribution of travel is studied in Section 5.1. In order to further understand the stable movement pattern of passengers' travel in Shenzhen Metro, the analysis of spatial distribution of travel will be studied in this section. We extracted more than 21,000 point-of-interest information from APIs in the Baidu map and classified them into seven major categories: R, C, L, T, F, M and E. Based on the extracted POI information, one primary function is assigned to each metro station using following method: we firstly search all the POIs within a radius of 1000 m for each metro station, then the specific category that covers the largest proportion of POIs is set to be the primary function for the station. For example, the station "ShenZhenBeiZhan" is described as transportation hub (T) and the station "ShaoNianGong" is assigned as the function of leisure and entertainment (L). Furthermore, the 30 metro stations with highest visiting frequency are identified for each mode based on the result of tensor decomposition, and then set their primary functions as the feature of the mode.

Table 2 shows the spatial distribution of the stable departure and arrival patterns of urban travel on weekdays and weekends. From the Table 2, we can see Weekday-Departure is decomposed into three basic modes. For Mode-1, we extracted the 30 stations with high visiting frequency in the spatial dimension and found that the passenger source comes from residential areas and leisure and entertainment. The spatial distribution characteristic of Mode-1 is composed of 40% [R] +16.7% [L]

+ 13.3% [C] +13.3% [T] +10% [F] + 3.3% [M] +3.3% [E], then the feature of Mode-1 can be defined as [R] for its highest proportion. For the Mode-2, in the stations with high visiting frequency, ten of them belong to the corporate companies, six stations belong to the residential area, five stations belong to leisure and entertainment, four stations belong to transportation hub, three stations belong to gourmet restaurant, one station belongs to medical institution and one station belongs to education institution. The spatial distribution characteristic of Mode-2 is composed of 33.3% [C] +20% [R] + 16.7% [L] +13.3% [T] +10% [F] + 3.3% [M] + 3.3% [E], then the feature of Mode-2 can be identified as [C]. Similarly, the spatial distribution characteristic of Mode-3 is composed of 30% [L] +20% [R] + 16.7% [C] +13.3% [F] +10% [T] + 6.7% [E] + 3.3% [M], then the feature of Mode-3 can be identified as [L]. For the weekday, passengers travel from the residential area to the corporate area at the morning peak, and go to the residential area from the corporate area at the evening peak. For the weekend, passengers travel from the residential area to the entertainment area and corporate area at the morning peak, and from the entertainment area, corporate area to the residential area at the evening peak. Most of the metro trips do not provide door-to-door service, so it may be necessary to transfer by other means in the transportation hub area.

**Table 2.** Spatial distribution table in stable patterns.

State	Mode	R	C	L	T	F	M	E	Feature
Weekday-Departure	Mode-1	12	4	5	4	3	1	1	[R]
	Mode-2	6	10	5	4	3	1	1	[C]
	Mode-3	6	5	9	3	4	1	2	[L]
Weekday-Arrival	Mode-1	6	9	6	4	3	1	1	[C]
	Mode-2	12	3	5	4	4	1	1	[R]
	Mode-3	7	8	6	3	3	2	1	[C]
Weekend-Departure	Mode-1	12	4	5	3	4	1	1	[R]
	Mode-2	7	4	10	4	3	1	1	[L]
	Mode-3	10	7	5	3	3	1	1	[R]
Weekend-Arrival	Mode-1	4	7	10	4	2	1	2	[L]
	Mode-2	12	4	4	4	4	1	1	[R]
	Mode-3	8	7	5	3	3	2	2	[R]

### 5.3. Analysis of Travel Time Distribution

We extract the records of different modes based on the tensor decomposition results and then analyze their statistical characteristics. The travel time distribution of passengers is a key factor to evaluate the operating efficiency of a metro network. Taking the travel time as an example, the daily record data is divided into three categories according to their departure/arrival time and the metro station. Millions of passengers' records of the specific weekday (6 April 2017) and the weekend (8 April 2017) were extracted by category, and the travel time of each passenger was calculated. The distribution characterization of travel time on weekday and weekend was analyzed according to different classes.

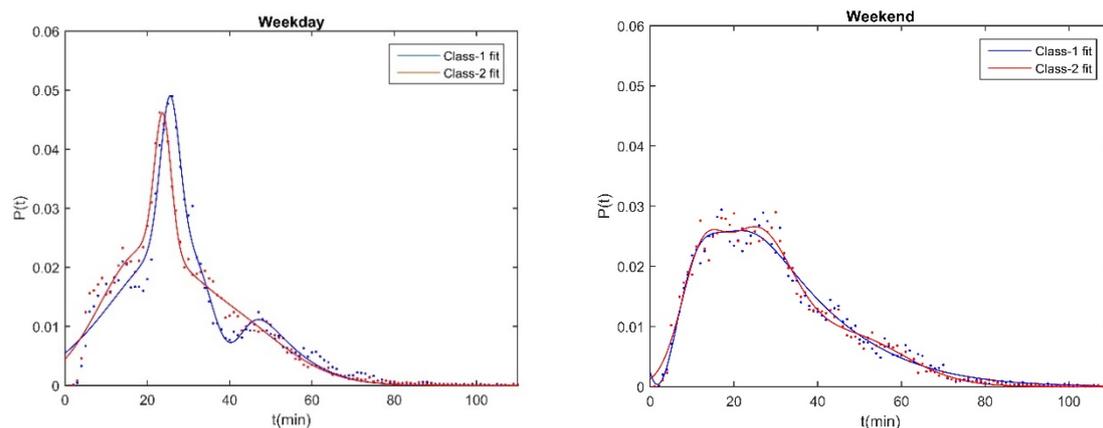
Class-1 corresponds to the early peak period represented by mode-1. On weekdays, we extracted the records from 7:00 to 8:00, the departure station function is R, and the arrival station function is C. On weekends, we extracted the records from 7:00 to 8:00, the departure station function is R, and the arrival station function is L. Class-2 corresponds to the early peak period represented by mode-2. On weekdays, we extracted the records from 17:00 to 18:00, the departure station function is C, and the arrival station function is R. On weekends, we extracted the records from 17:00 to 18:00, the departure station function is L, and the arrival station function is R. Class-3 corresponds to the daily period represented by mode-3, its travel time distribution becomes more stable, there is no obvious peak time, and its spatial distribution characteristics have no obvious statistical rules, so we did not perform further analysis for Class-3. Based on the classification, the proportion of travel time at one minute's intervals is calculated. The fitting degree results of four function distribution are shown in Table 3. We

can see that the Gaussian mixture model has the best fitting results, whose probability density obeys the distribution  $p(t) = a_1 \times \exp\left(-\left(\frac{t-b_1}{c_1}\right)^2\right) + a_2 \times \exp\left(-\left(\frac{t-b_2}{c_2}\right)^2\right) + a_3 \times \exp\left(-\left(\frac{t-b_3}{c_3}\right)^2\right)$ .

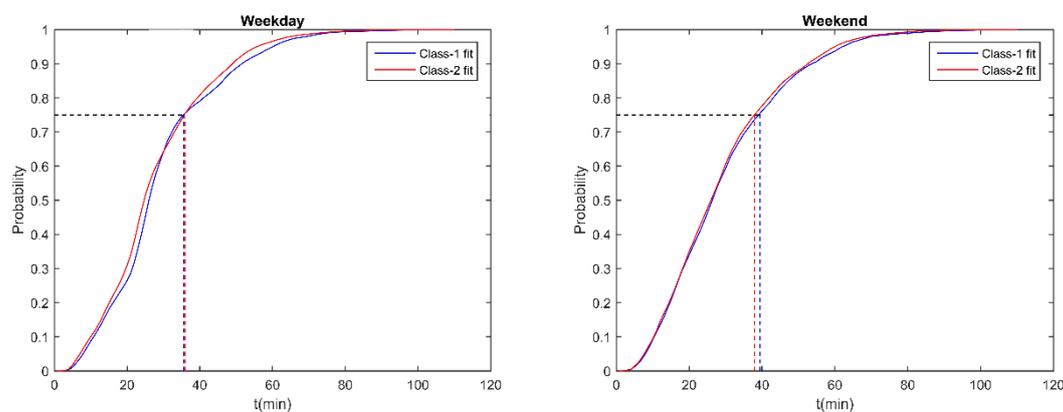
**Table 3.** Fitting degree results of four function distribution.

	Class	Gaussian Distribution	Weibull Distribution	Gaussian Mixed Distribution	Polynomial Distribution
Weekdays	Class-1	0.8122	0.7965	0.9648	0.7592
	Class-2	0.8715	0.8846	0.9763	0.8614
Weekends	Class-1	0.9045	0.9544	0.9767	0.947
	Class-2	0.9147	0.9611	0.9797	0.951

Figure 7 expresses the proportion of travel volume in five minutes' interval in travel records of the class. We found that the probability distribution of travel time of different classes had slight differences, which may be caused by the travel of the same station function (R & C in weekday, R & L in weekend). The cumulative probabilistic distribution of travel time is similar between weekdays and weekends in the Figure 8. We obtained the position of the third-quartile  $t_0$ , which represent high travel probability within the travel time  $t_0$ . On weekdays, for Class-1,  $t_0$  is 35.5 min; for Class-2,  $t_0$  is 35.6 min. On weekends, for Class-1 and Class-2,  $t_0$  is 37.9 min and 39.5 min.



**Figure 7.** Probabilistic distribution of travel time.



**Figure 8.** Cumulative probabilistic distribution of travel time.

## 6. Conclusions

The smart card data collected from a metro system were used to analyze the characteristics of urban metro passenger flow, urban spatial functional structure and travel behaviors. The whole

mobility is considered as a combination of several basic patterns. Applying tensor decomposition and correlation analysis, stable components are extracted from collective movement data, and each pattern is analyzed according to the primary function of the metro station by using POI information. The main conclusion of this study can be summarized as follows:

- (1) The departure and arrival mobility of weekdays and weekends can be decomposed into several stable basic modes by tensor decomposition. For example, it can be decomposed into morning peak hours, evening peak hours and daily pattern.
- (2) Based on the tensor decomposition results and correlation analysis, stable and occasional patterns are distinguished.
- (3) On weekdays, passengers' travel broke out in the morning and evening peak hours. However, on weekends, the peak hours are slightly later than on weekdays. The peak period lasted for a long time, and the passengers' travel in the peak period decreased significantly. Since passengers are not limited by rigid work requirements on weekends, liquidity is more random and uniform and the change of passengers' flow is slower than that of weekdays.
- (4) The typical stable patterns on weekdays mainly involve travel between residential and workplace. The typical stable patterns on weekends mainly involve travel between residential and entertainment.
- (5) The travel time of passengers is calculated based on the classification. We find the travel time distribution is similar between different classes, which mainly caused by the travel of stations with same function, and the metro travel time with highest travel probability is within 35–40 min. Above all, we try to apply a tensor decomposition perspective to depict the spatial-temporal characteristic of passenger flow on the metro network.

What we found in this study has important reference value for the planning and operation of metro networks, which contributes to a better understanding travel behavior. Identifying the spatial-temporal characteristic of trips can help administrations strengthen the management and operations of stations so that the service level and quality of safety can be improved in the public traffic system. Meanwhile, the departure interval of metros can be specifically adjusted in all directions and metro ridership can be improved efficiently.

This study could be extended in the following ways: (1) The research based on metro smart card data does not fully reflect all aspects of urban public transport travel, such as bicycle trips, private cars, taxis, buses and other travel behavior, so we could further study the spatial-temporal interplay between different mass transit modes in the future; (2) We only used smart card data for three consecutive weeks on normal conditions, in the next step, the spatial-temporal characteristic of travel patterns is necessary to be explored under different conditions, such as severe weather, large-scale activities, emergencies and so on; (3) The tensor can be further constructed a multi-dimensional model, which not only includes spatial-temporal information but also information about transport mode and different types of public transport passengers and so on; (4) Travel behavior is affected by many factors, such as passengers' income level, personal preference, age and so on, so travel behavior characteristics under the influence of multiple factors can be further explored.

**Author Contributions:** Conceptualization, J.T. and F.Z.; Methodology, X.W. and Z.H.; Supervision, J.T.; Validation, F.Z.; Visualization, X.W.; Writing—original draft, J.T. and X.W.; Writing—review & editing, F.Z. and Z.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded in part by the National Natural Science Foundation of China (No. 71701215), Innovation-Driven Project of Central South University (No. 2020CX041), Foundation of Central South University (No. 502045002), Postdoctoral Science Foundation of China (No. 2018M630914 and 2019T120716).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Appendix A

Table A1. Coefficient of correlation of Mo component.

(m, n)	r <sub>ij</sub>	r11	r12	r13	r21	r22	r23	r31	r32	r33
1,2		0.315	0.989	0.238	0.591	0.277	0.94	0.843	0.164	0.561
1,3		0.928	0.01	0.431	0.204	0.389	0.608	0.061	0.398	0.74
1,4		0.91	0.025	0.392	0.189	0.393	0.651	0.029	0.41	0.736
1,5		0.915	0.049	0.271	0.191	0.368	0.696	0.025	0.371	0.803
2,1		0.315	0.591	0.843	0.989	0.277	0.164	0.238	0.94	0.561
2,3		0.224	0.306	0.883	0.915	0.002	0.499	0.235	0.282	0.417
2,4		0.171	0.317	0.787	0.888	0.014	0.451	0.229	0.28	0.484
2,5		0.165	0.289	0.778	0.893	0.04	0.318	0.234	0.263	0.517
3,1		0.928	0.204	0.061	0.01	0.389	0.398	0.431	0.608	0.74
3,2		0.224	0.915	0.235	0.306	0.002	0.282	0.883	0.499	0.417
3,4		0.996	0.044	0.271	0.05	0.999	0.341	0.267	0.287	0.968
3,5		0.996	0.023	0.152	0.05	0.992	0.441	0.268	0.263	0.909
4,1		0.91	0.189	0.029	0.025	0.393	0.41	0.392	0.651	0.736
4,2		0.171	0.888	0.229	0.317	0.014	0.28	0.787	0.451	0.484
4,3		0.996	0.05	0.267	0.044	0.999	0.287	0.271	0.341	0.968
4,5		0.999	0.016	0.119	0.035	0.993	0.446	0.234	0.319	0.956
5,1		0.915	0.191	0.025	0.049	0.368	0.371	0.271	0.696	0.803
5,2		0.165	0.893	0.234	0.289	0.04	0.263	0.778	0.318	0.517
5,3		0.996	0.05	0.268	0.023	0.992	0.263	0.152	0.441	0.909
5,4		0.999	0.035	0.234	0.016	0.993	0.319	0.119	0.446	0.956
6,7		0.275	0.976	0.133	0.835	0.065	0.728	0.709	0.164	0.891
7,6		0.275	0.835	0.709	0.976	0.065	0.164	0.133	0.728	0.891

Table A2. Coefficient of correlation of Md component.

(m, n)	r <sub>ij</sub>	r11	r12	r13	r21	r22	r23	r31	r32	r33
1,2		0.332	0.963	0.387	0.981	0.436	0.239	0.232	0.116	0.92
1,3		0.267	0.225	0.756	0.938	0.118	0.34	0.088	0.253	0.605
1,4		0.272	0.691	0.146	0.936	0.254	0.073	0.101	0.558	0.188
1,5		0.308	0.713	0.156	0.942	0.268	0.073	0.135	0.597	0.194
2,1		0.332	0.981	0.232	0.963	0.436	0.116	0.387	0.239	0.92
2,3		0.914	0.138	0.365	0.344	0.199	0.676	0.166	0.349	0.617
2,4		0.911	0.278	0.09	0.346	0.618	0.129	0.178	0.588	0.284
2,5		0.918	0.29	0.092	0.375	0.628	0.138	0.225	0.63	0.285
3,1		0.267	0.938	0.088	0.225	0.118	0.253	0.756	0.34	0.605
3,2		0.914	0.344	0.166	0.138	0.199	0.349	0.365	0.676	0.617
3,4		0.999	0.115	0.032	0.003	0.583	0.993	0.22	0.934	0.173
3,5		0.979	0.133	0.034	0.051	0.522	0.994	0.258	0.958	0.189
4,1		0.272	0.936	0.101	0.691	0.254	0.558	0.146	0.073	0.188
4,2		0.911	0.346	0.178	0.278	0.618	0.588	0.09	0.129	0.284
4,3		0.999	0.003	0.22	0.115	0.583	0.934	0.032	0.993	0.173
4,5		0.982	0.143	0.034	0.167	0.993	0.505	0.014	0.425	0.999
5,1		0.308	0.713	0.156	0.942	0.268	0.073	0.135	0.597	0.194
5,2		0.918	0.375	0.225	0.29	0.628	0.63	0.092	0.138	0.285
5,3		0.979	0.051	0.258	0.133	0.522	0.958	0.034	0.994	0.189
5,4		0.982	0.167	0.014	0.143	0.993	0.505	0.034	0.505	0.999
6,7		0.971	0.491	0.057	0.222	0.876	0.626	0.349	0.125	0.784
7,6		0.971	0.222	0.349	0.491	0.876	0.125	0.057	0.626	0.784

## References

1. Ma, X.; Liu, C.; Wen, H.; Wang, Y.; Wu, Y.J. Understanding commuting patterns using transit smart card data. *J. Transp. Geogr.* **2017**, *58*, 135–145. [[CrossRef](#)]
2. Dou, M.; He, T.; Yin, H.; Zhou, X.; Chen, Z.; Luo, B. Predicting passengers in public transportation using smart card data. *Australas. Database Conf.* **2015**, *9093*, 28–40.
3. Gordillo, F. *The Value of Automated Fare Collection Data for Transit Planning: An Example of Rail Transit OD Matrix Estimation*; Massachusetts Institute of Technology: Cambridge, MA, USA, 15 September 2006.
4. Ma, Y.; Xu, W.; Zhao, X.; Li, Y. Modeling the hourly distribution of population at a high spatiotemporal resolution using subway smart card data: A case study in the central area of Beijing. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 128. [[CrossRef](#)]
5. Zhao, J.; Zhang, F.; Zhang, J.; Sun, L. Exploring Human Mobility Patterns Related to Shenzhen-Hong Kong Ports: A Preliminary Study for Futian Port. In Proceedings of the 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA), Beijing, China, 10–12 March 2017; pp. 950–954.
6. Yang, C.; Yan, F.F.; Xu, X.D. Clustering Daily Metro Origin—Destination Matrix in Shenzhen China. *Appl. Mech. Mater.* **2015**, *743*, 422–432. [[CrossRef](#)]
7. Ma, X.L.; Wang, Y.H.; Chen, F.; Liu, J.F. Transit Smart Card Data Mining for Passenger Origin Information Extraction. *J. Zhejiang Univ. Sci. C* **2012**, *13*, 750–760. [[CrossRef](#)]
8. Yue, M.; Kang, C.; Andris, C.; Qin, K.; Liu, Y.; Meng, Q. Understanding the interplay between bus, metro, and cab ridership dynamics in Shenzhen, China. *Trans. GIS* **2018**, *22*, 855–871. [[CrossRef](#)]
9. Rhee, I.; Shin, M.; Hong, S.; Lee, K.; Kim, S.J.; Chong, S. On the Levy-walk nature of human mobility. *IEEE/ACM Trans. Netw.* **2011**, *19*, 630–643. [[CrossRef](#)]
10. González, M.C.; Hidalgo, C.A.; Barabási, A.L. Understanding individual human mobility patterns. *Nature* **2009**, *458*, 238. [[CrossRef](#)]
11. Zhao, J.; Zhang, F.; Tu, L.; Xu, C.; Shen, D.; Tian, C. Estimation of passenger route choice pattern using smart card data for complex metro systems. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 790–801. [[CrossRef](#)]
12. Chen, E.; Ye, Z.; Wang, C.; Zhang, W. Discovering the spatio-temporal impacts of built environment on metro ridership using smart card data. *Cities* **2019**, *95*, 102359. [[CrossRef](#)]
13. Zhang, Y.; Martens, K.; Long, Y. Revealing group travel behavior patterns with public transit smart card data. *Travel Behav. Soc.* **2018**, *10*, 45–52. [[CrossRef](#)]
14. Du, Z.; Tang, J.; Qi, Y.; Wang, Y.; Han, C.; Yang, Y. Identifying Critical Nodes in Metro Network Considering Topological Potential: A case study in Shenzhen city—China. *Phys. A* **2020**, *539*, 122926. [[CrossRef](#)]
15. Zhang, Y.; Liu, L. Understanding temporal pattern of human activities using Temporal Areas of Interest. *Appl. Geogr.* **2018**, *94*, 95–106. [[CrossRef](#)]
16. Sun, Y.; Shi, J.; Schonfeld, P.M. Identifying passenger flow characteristics and evaluating travel time reliability by visualizing AFC data: A case study of Shanghai Metro. *Public Transp.* **2016**, *8*, 341–363. [[CrossRef](#)]
17. Zhao, J.; Qu, Q.; Zhang, F.; Xu, C.; Liu, S. Spatio-Temporal Analysis of Passenger Travel Patterns in Massive Smart Card Data. *IEEE Trans. Intell. Transp. Syst.* **2017**, *11*, 3135–3146. [[CrossRef](#)]
18. Sun, L.; Jin, J.G. Modeling Temporal Flow Assignment in Metro Networks Using Smart Card Data. In Proceedings of the International Conference on Intelligent Transportation Systems IEEE, Las Palmas, Spain, 15–18 September 2015; pp. 836–841.
19. Zhong, C.; Batty, M.; Manley, E.; Wang, J.; Wang, Z.; Chen, F.; Schmitt, G. Variability in Regularity: Mining Temporal Mobility Patterns in London, Singapore and Beijing Using Smart-Card Data. *PLoS ONE* **2016**, *11*, e0149222. [[CrossRef](#)]
20. Yang, C.; Yan, F.; Ukkusuri, S.V. Unraveling traveler mobility patterns and predicting user behavior in the shenzhen metro system. *Transp. A Transp. Sci.* **2018**, *14*, 576–597. [[CrossRef](#)]
21. Truong, R.; Gkountouna, O.; Pfoser, D.; Züfle, A. Towards a better understanding of public transportation traffic: A case study of the Washington, DC metro. *Urban Sci.* **2018**, *2*, 65. [[CrossRef](#)]
22. Wang, Y.; de Almeida Correia, G.H.; de Romph, E.; Timmermans, H.J.P. Using metro smart card data to model location choice of after-work activities: An application to shanghai. *J. Transp. Geogr.* **2017**, *63*, 40–47. [[CrossRef](#)]
23. Ma, X.W.; Ji, Y.J.; Fan, Y. Exploring the evolution of passenger flow and travel time reliability with the expanding process of metro system using smartcard data. *J. Harbin Inst. Technol.* **2019**, *26*, 21–33.

24. Van Oort, N.; Brands, T.; de Romph, E. Short Term Ridership Prediction in Public Transport by Processing Smart Card Data. In Proceedings of the Transportation Research Board 94th Annual Meeting, Washington, DC, USA, 11–15 January 2015.
25. Sun, L.J.; Axhausen, K.W. Understanding urban mobility patterns with a probabilistic tensor factorization framework. *Transp. Res. Part B Methodol.* **2016**, *91*, 511–524. [[CrossRef](#)]
26. Du, Z.; Yang, B.; Liu, J. Understanding the Spatial and Temporal Activity Patterns of Subway Mobility Flows. *arXiv* **2017**, arXiv:1702.02456.
27. Kyoungok, K. Exploring the difference between ridership patterns of subway and taxi: Case study in seoul. *J. Transp. Geogr.* **2018**, *66*, 213–223.
28. Yong, N.; Ni, S.; Shen, S. A Preliminary Study of Mobility Patterns in Urban Subway. In *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*; Springer: Cham, Switzerland, 2016; pp. 61–70.
29. Liu, S.; Yao, E.; Cheng, X.; Zhang, Y. Evaluating the impact of new lines on entrance/exit passenger flow of adjacent existing stations in urban rail transit system. *Transp. Res. Procedia* **2017**, *25*, 2625–2638. [[CrossRef](#)]
30. Fu, X.; Gu, Y. Impact of a new metro line: Analysis of metro passenger flow and travel time based on smart card data. *J. Adv. Transp.* **2018**, *2018*, 9247102. [[CrossRef](#)]
31. Yong, N.; Ni, S.; Shen, S.; Chen, P.; Ji, X. Uncovering stable and occasional human mobility patterns: A case study of the Beijing subway. *Phys. A* **2018**, *492*, 28–38. [[CrossRef](#)]
32. Tamara, G.K.; Brett, W.B. *Tensor Decompositions and Applications*; Sandia Report: SAND2007-6702; SIAM Review; Sandia National Laboratories: Albuquerque, NM, USA, 2009; Volume 51, pp. 455–500.
33. Cichocki, A.; Zdunek, R.; Phan, A.H.; Amari, S.I. *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*; John Wiley & Sons: West Sussex, UK, 2009.
34. Fan, Z.P.; Shibasaki, X.R. CitySpectrum: A non-negative tensor factorization approach. In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing, New York, NY, USA, 17 May 2014.
35. Bader, B.W.; Kolda, T.G. Matlab Tensor Toolbox Version 2.5. Available online: <http://www.sandia.gov/~tgkolda/TensorToolbox/> (accessed on 12 October 2018).
36. Baidu Map. Available online: <http://lbsyun.baidu.com/index.php?title=webapi/guide/webservice-placeapi> (accessed on 15 March 2019).
37. Peng, C.; Jin, X.; Wong, K.C.; Shi, M.; Lio, P. Collective Human Mobility Pattern from Taxi Trips in Urban Area. *PLoS ONE* **2012**, *7*, e34487.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).