

Research Article

Data-Driven Approach for Passenger Mobility Pattern Recognition Using Spatiotemporal Embedding

Chao Yu ^{1,2}, Haiying Li,¹ Xinyue Xu ¹, Jun Liu ¹, Jianrui Miao ¹, Yitang Wang,³
and Qi Sun⁴

¹State Key Laboratory of Rail Traffic Control & Safety, Beijing Jiaotong University, Beijing 100044, China

²School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

³Track and Transportation Department, China Railway First Survey and Design Institute, Xi'an, Shaanxi 710043, China

⁴ACC Technical Room, Beijing Metro Network Control Center, Beijing 100101, China

Correspondence should be addressed to Xinyue Xu; xxy@bjtu.edu.cn

Received 8 January 2021; Revised 11 April 2021; Accepted 8 May 2021; Published 19 May 2021

Academic Editor: Chung-Cheng Lu

Copyright © 2021 Chao Yu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Urban mobility pattern recognition has great potential in revealing human travel mechanism, discovering passenger travel purpose, and predicting and managing traffic demand. This paper aims to propose a data-driven method to identify metro passenger mobility patterns based on Automatic Fare Collection (AFC) data and geo-based data. First, Point of Information (POI) data within 500 meters of the metro stations are captured to characterize the spatial attributes of the stations. Especially, a fusion method of multisource geo-based data is proposed to convert raw POI data into weighted POI data considering service capabilities. Second, an unsupervised learning framework based on stacked auto-encoder (SAE) is designed to embed the spatiotemporal information of trips into low-dimensional dense trip vectors. In detail, the embedded spatiotemporal information includes spatial features (POI categories around the origin station and that around the destination station) and temporal features (start time, day of the week, and travel time). Third, a density-based clustering algorithm is introduced to identify passenger mobility patterns based on the embedded dense trip vectors. Finally, a case of Beijing metro network is used to verify the feasibility of the above methodology. The results show that the proposed method performs well in recognizing mobility patterns and outperforms the existing methods.

1. Introduction

The number of urban residents is increasing significantly, and human mobility is becoming unpredictable and complex, posing major challenges to public safety and health (such as the COVID-19 epidemic). In recent years, urban mobility pattern recognition has become a hotspot due to its ability to reveal resident life routines, assist in transportation planning, estimate and manage travel demand, predict passenger travel purposes, and provide location-based services [1–5]. As an important part of urban transportation, the metro system has increasingly become an indispensable choice for urban residents. Therefore, studying metro passenger mobility patterns is essential for analyzing urban mobility characteristics.

Fortunately, the continuous development of digitalization has provided strong support for urban planning and transportation services. Currently, large-scale spatiotemporal travel-related data provide the possibility for the analysis of passenger mobility patterns. From the perspective of the types of raw data, the recognition of urban mobility patterns can be divided into two categories, namely, researches based on trajectory data and that based on AFC data. The former is mainly meant to reproduce the movement track of residents through GPS data, social media data, or mobile phone signaling data to identify mobility patterns [6–12]. Unlike this, the latter often uses the tap-in or tap-out data of passengers to describe the travel process in order to realize the analysis of travel patterns [1, 13–20]. However, there are some shortcomings in trajectory data. First,

trajectory data are often obtained when the mobile phone user turns on the positioning function, which means that the behavior of the user to turn on or off the positioning function has a direct impact on the collection of trajectory data. Second, the accuracy of trajectory data depends on the reliability of positioning technology. In fact, most positioning methods often have unavoidable errors, especially in densely populated areas or underground multistory buildings, resulting in blurred trajectories. Conversely, an individual trajectory identified by AFC data is error-free at the spatial level of stops and stations [15]. Admittedly, AFC data cannot pinpoint the specific activity location of passengers. However, it is possible to use the land-use data around the station to infer the possible activity locations of passengers, because passengers often complete the displacement before tap-in or after tap-out by walking [21].

It is undeniable that trajectory data and AFC data have their own advantages and disadvantages in identifying passenger mobility patterns. For metro managers and operators, AFC data are relatively accurate and easily available. Using AFC data to analyze passenger mobility patterns and behavior characteristics can significantly improve the metro service level. This paper aims to propose a data-driven approach to explore the possibility of AFC data in inferring passenger mobility patterns. In the existing research on mobility patterns, the tap-in timestamp, tap-out timestamp, and travel time are usually fused to mine the temporal characteristics. However, the discovery of spatial features usually stays at the station level. The common method is to characterize the latent spatial characteristics by dividing the stations into several different clusters, which makes it difficult to infer the specific mobility patterns of passengers. In view of the above analysis, AFC data are selected to extract passenger travel information. In addition, multisource geobased data are captured to provide the necessary land-use information to realize passenger mobility recognition. In this paper, each AFC travel record is processed by an unsupervised method into a low-dimensional vector containing spatiotemporal features. There are two advantages. First, the concrete spatial information and temporal information being transformed into abstract vector forms are convenient for large-scale processing by computers (for example, similarity calculation). Second, vectorization can extract the characteristics of travel records to the maximum extent while saving storage space to explore the internal mechanism of passenger mobility [7].

The contribution of this paper is threefold. First, a multisource data fusion method is presented. This method adds the residential area information provided by the housing trading platform and the building information provided by the geographic information service to the raw POI data to convert the raw POI data into weighted POI data considering service capabilities. It avoids the drawbacks of using POI numbers to quantify land-use characteristics in the existing works [21]. Second, an unsupervised deep learning framework based on SAE is proposed to embed the spatiotemporal information of passenger travel, so as to realize the conversion of a passenger travel record into a low-dimensional dense vector. In this framework, the self-

encoding is utilized to realize the embedding of spatiotemporal information without the labeled data and supervised training, which can extract the features of travel records more comprehensively than existing methods [22, 23]. Third, a density-based clustering algorithm is used to identify passenger mobility patterns. It can generate the number of clusters according to the data distribution without manually specifying the number of clusters, avoiding the human intervention of existing methods [7, 24].

The structure of this paper is as follows. In Section 2, the existing studies on mobility pattern recognition are classified and summarized. In Section 3, the methodology of this paper is introduced in detail, including an overview of the method and three main steps, namely, the fusion of multisource geobased data, embedding spatiotemporal semantics in trip records, and mobility pattern recognition based on the embedded vectors. In Section 4, a case based on the Beijing metro network is introduced to verify the effectiveness of the proposed method, and the results of the case study are compared with existing methods. Besides, potential applications based on passenger mobility pattern recognition are explained. Finally, the paper is summarized and discussed in Section 5.

2. Literature Review

Passenger mobility pattern recognition aims to discover the identifiable travel categories formed by passengers in the long-term travel history, such as working, going home, entertainment, etc. Existing research has revealed that urban mobility exhibits a high degree of regularity in time and space [7, 25]. This allows us to discover the daily routines and social state of travelers through mobility analysis. To do this, many methods have been proposed in the existing work. Macroscopically, these methods can be classified into two categories, namely, empirical models and data-driven models.

Intuitively, the empirical method is to quantitatively analyze passenger behavior by features or thresholds of the known activity categories. The abovementioned features and thresholds tend to be artificially designated by researchers or experts. For example, a rule was established by [18] that the cardholder's first tap-in station or the last tap-out station can be considered as his/her potential home location. An algorithm based on "center point" is proposed to infer cardholder's exact home location based on multiple potential locations. The effectiveness of this method is verified by a case of Beijing metro system, in which 88.7% of passengers' home locations were successfully inferred. Similarly, a passenger's home location was determined to be the most visited location between 7 pm and 8 am on weekends and weekdays, as suggested by [11]. It was presented by [9] that a passenger's home and work place are the most visited and second most visited locations. Although the above assumptions can help infer the passenger's home and work locations to a certain extent, they are not universal. The rules are often subjective, and their application effects rely heavily on the domain knowledge of experts or scholars [23]. Furthermore, the empirical method is incapable of

discovering new mobility patterns, resulting in the inability to keenly estimate the changing trend of urban mobility with the increase in population and the complexity of the urban transportation network.

In order to avoid the above shortcomings, data-driven methods have emerged. As mentioned in Section 1, large-scale datasets provided more possibilities for mobility analysis. In the past few years, a variety of datasets have been used to describe urban mobility, such as mobile phone signing data, GPS data, media data, AFC data, sociodemographic data, and census and administrative data [1, 26, 27]. Faced with such diverse datasets, many data-driven methods have also been proposed by researchers to mine passenger mobility patterns. For instance, multi-objective Convolutional Neural Network (CNN) was designed to infer the social demographic attributes and mobility features of passengers based on media data and land-use data [28]. Support vector machines (SVM) were introduced to divide passenger travel data into several types, and the passenger purpose was analyzed according to the characteristics of each type using sociodemographic data [8]. This method was applied to data from a large number of Californian families. The application results showed that this method performed better than the traditional multinomial logit models. Moreover, smart card data can also be utilized to construct land-use function complementation indices to improve the performance of the classic gravity model in analyzing the human mobility between different types of areas in the city [29]. The case of Shenzhen metro showed that these indices were effective tools to reveal the mechanism of spatial interaction and had a significant effect on improving the prediction of spatial flow and travel distribution. The naive Bayes probability model was improved to observe the continuous long-term changes in the attributes of metro passenger trips using AFC data and census data [13]. The verification results of real cases showed that 86.2% of passengers' travel purpose can be estimated. A data-driven robust method using AFC data and the General Transportation Feedback Specification (GTFS) was designed to infer the most likely movement trajectory of each passenger [20]. The use of GTFS data reduced many assumptions about the passenger travel process in previous studies (the threshold assumptions of transfer travel time, time window assumptions for selecting vehicles and journeys, threshold assumptions for waiting and boarding time, etc.). This method was used in the analysis of passenger travel trajectories in Minnesota and proved to be superior to traditional trajectory inference methods. Besides, to recognize the patterns of passengers' variation over time and analyze the spatial heterogeneity of the dynamic space around the metro stations, an eigendecomposition method was proposed [30]. In this work, the datasets were decomposed into a combination of principal components and eigenvectors, where the principal components represent the common pattern of passenger movement, and the corresponding elements in the eigenvectors mean the attributes of metro stations. The above method was verified in the case of the Shenzhen metro system and proved to be effective in improving urban planning. A method based on the Hidden Markov Model (HMM) was addressed to infer the sequence of passenger activities, and the model parameters were

calibrated using Baum–Welch algorithm based on land-use data around the stations [31]. The abovementioned data-driven methods excavated the rules of passenger mobility from different aspects, but there are still shortcomings of high computational cost and poor interpretability.

In recent years, various types of topic models have gradually become the mainstream methods for the analysis of urban mobility patterns [6, 9, 23]. In these studies, mobility pattern recognition was regarded as a topic mining problem in the field of natural language processing (NLP). In the model, each passenger was treated as an article, each trip record of the passenger is processed as a word in the article, and the previous and subsequent trips of a certain trip were considered as the context of the current trip. Correspondingly, passenger mobility pattern recognition can be understood as mining several topics in the corpus composed of multiple articles. For example, a multi-directional probabilistic factorization model based on tensor decomposition and probabilistic latent semantic analysis (PLSA) was proposed, which used a simple latent semantic structure to describe the multi-directional mobility characteristics of passengers involved in high-order interactions [16]. The multi-directional mobility analysis of urban residents in Singapore verified the practicality of the model. A Bayesian n-gram model was constructed to predict the location and time of individual passenger activities, and its prediction result was expressed as an ordered set of passenger potential activities, which contains the location and time of each activity [32]. On this basis, a spatiotemporal topic model based on Latent Dirichlet Allocation (LDA) was presented to classify passenger activities into several topics to realize mobility pattern recognition [23]. The above method was verified by the travel data of more than 10,000 users of the London Underground in 2 years, and the results showed that the median accuracy of travel prediction could reach 80%. The obtained passenger mobility patterns could well reveal the temporal and spatial attributes of work-related and home-related activities. Unfortunately, the abovementioned researches only analyzed mobility from the perspective of temporal characteristics, without considering spatial information, which makes the results poor in interpretability. Considering spatial features, methods based on word vector were introduced for exploring mobility patterns. For example, a *habit2vec* method was proposed by [7] to embed a passenger's current visit to a POI type during a time slice. Besides, the inbound flow, the outbound flow, and the surrounding POIs were used as elements to construct the target station vector suggested by [21]. In this work, it was worth noting that the Term Frequency–Inverse Document Frequency (TF-IDF), which was an indicator in the NLP field, was applied to quantify categories of the target station. Nevertheless, it is unreasonable to determine station categories only by the frequency or TF-IDF of different categories of POI around the station due to the significant difference in service capabilities of different categories of POI. For example, although a residential area and a cafe are both displayed as POIs on the map, the service capacity of the former is obviously greater than that of the latter. Therefore, a POI needs to be weighted according to its

service capability to be meaningful in describing passenger mobility.

In a nutshell, the existing works on passenger mobility is in the ascendant, but there are still defects such as high computational cost, lack of consideration of spatial features, and poor interpretability. In this paper, weighted POI is first generated through multisource geo-based data. Then, through the unsupervised learning framework based on SAE, the temporal and spatial features are simultaneously embedded into the trip vector of passengers to identify the mobility patterns. The following is the methodology of this work.

3. Methodology

The overview of the methodology is shown in Figure 1. The goal is to design an efficient method to transform trip records into standard forms that can be processed by computers, so as to simplify mobility pattern recognition into a clustering problem. After obtaining AFC records, the following three steps are required to achieve the above goal. First, a fusion of multisource, geo-based data method is proposed to weight the raw POI data and provide a basis for spatial semantic estimation. Second, a low-dimensional dense trip vector containing both spatial and temporal attributes is generated to represent the given record. Third, clustering analysis on low-dimensional dense trip vectors is addressed to distinguish between different trip clusters to realize mobility pattern recognition. Details of these three steps are described in the following sections.

3.1. Fusion of Multisource Geo-Based Data. POI is a point unit in geographic information systems to mark the location of human activity. A POI contains the POI name, category label, longitude, latitude, and land-use type information of the point unit [1]. Some existing studies infer the travel purpose of passengers through the category label of POIs around the target station. For example, when the POIs around a passenger's origin station are mostly residential and the POIs around the destination station are mostly working, the passenger's travel purpose can be considered to have a high probability of going to work [21]. Note that a POI can be a residential neighborhood, a shopping center, or a kindergarten. The service capacity of a residential neighborhood is obviously greater than that of a kindergarten. So, it is inaccurate to infer travel purpose from the number of POIs due to the difference in service capacity of different types of POIs. The goal of this section is to generate weighted POIs considering service capacity using multisource, geo-based data.

The geo-based data involved in this paper are obtained from three data sources, namely, Amap, Lianjia, and Arctiler. Among them, Amap (<https://www.amap.com/>) is a provider of digital map content, navigation, and location services solutions. It provides the raw POI data. It should be noted that Amap divides all POIs into 24 categories. For details of the classification, please refer to the website (<https://lbs.amap.com/api/webservice/download>). In this paper, from the perspective of travel purpose, these categories are integrated into 8 categories,

as shown in Table 1. In addition, some POIs that are not closely related to the travel purpose, such as public toilet and traffic light, are deleted. Besides, Lianjia (<https://www.lianjia.com/>) is a housing trading platform that can provide the neighborhood properties containing the name, housing price, property management fee, the number of buildings, and the number of households in a targeted residential neighborhood. For the residential POI in Table 1 (category 6), we use the number of households to represent its actual service capacity. Further, Arctiler (<http://www.arctiler.com/>) is a geographic information service provider that can provide the building physical properties containing the name, building category, usable area, and the number of floors of a target building. For different types of buildings, the per capita service area is stipulated by the Technical Measures for National Civil Building Engineering Design (<http://www.chinabuilding.com.cn/book-1815.html>). Therefore, we can calibrate the actual service capacity of the nonresidential POI in Table 1 by combining the building physical properties and per capita service area. With the above processing, the raw POI data have been converted into weighted POI data considering service capacity.

It should be noted that due to different data sources, the POI name may be different from the building name or the residential area name for the same point unit on the map, making data fusion difficult to achieve. Here, a data matching method is designed, as shown in Figure 2. For a given target POI, a building is selected from the Arctiler database, and the distance between the two is calculated to determine whether it matches each other. Note that it is necessary to convert the longitude and latitude of the building base outline obtained from Arctiler to that of the building base center. And then, the actual distance between the two coordinates can be calculated as follows:

$$\text{distance}(A, B) = \theta_{A,B} \cdot \frac{2\pi}{360} \cdot R_{\text{Earth}} \cdot 1000, \quad (1)$$

$$\theta_{A,B} = \arccos(\cos(A.\text{lat})\cos(B.\text{lat})\cos(A.\text{ln } g - B.\text{ln } g) + \sin(A.\text{lat})\sin(B.\text{lat})), \quad (2)$$

where $\text{Distance}(A, B)$ represents the actual distance between the two coordinate points A and B , in meters, $A.\text{lat}(B.\text{lat})$ and $A.\text{lng}(B.\text{lng})$ represent the latitude and longitude of A (B), and R_{Earth} represents the radius of the earth, which is 6371 km. All longitudes and latitudes in this paper are based on the World Geodetic System 1984 (WGS-84) coordinate system. Finally, it is judged whether the obtained distance is less than the threshold, which is set to 50 meters. If it is, the actual service capacity of the target POI is calibrated according to the per capita service area obtained from the Technical Measures for National Civil Building Engineering Design, that is, the weighted POI, otherwise, another building is selected from the Arctiler database to rematch the target POI. The data fusion process of residential POI is similar to this, and will not be repeated here. At this point, the raw POIs have been converted into weighted POIs based on multisource, geo-based data.

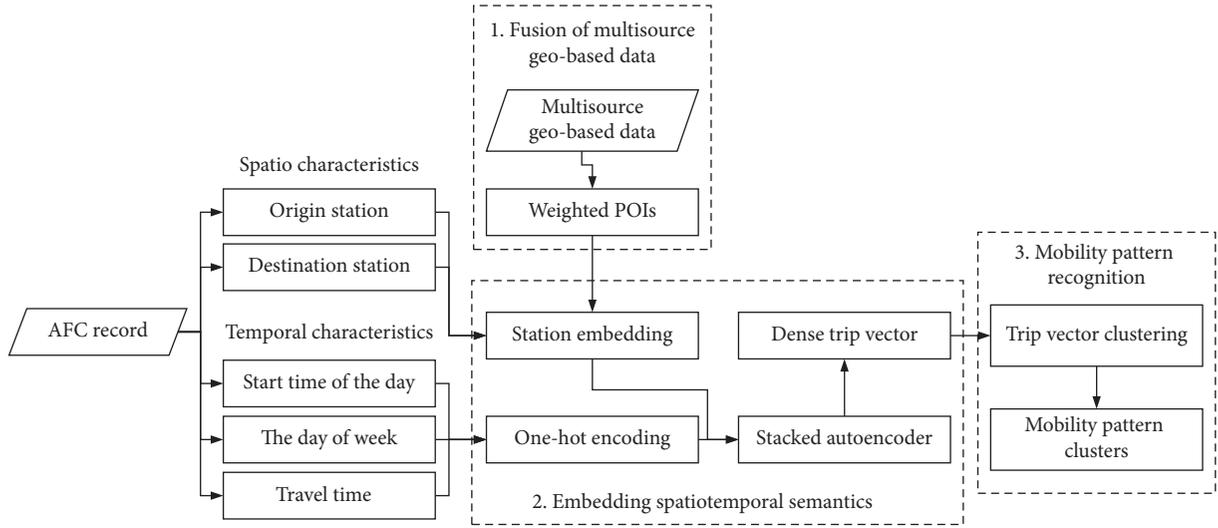


FIGURE 1: Overview of the methodology.

TABLE 1: POI categories and contents.

ID	Category	Contents
1	Entertainment	Recreation center, night club, KTV, disco, pub, game center, card and chess room, lottery center, Internet bar, recreation place, etc.
2	Working	Construction company, medical company, machinery and electronics, chemical and metallurgy, commercial trade, telecommunication company, mining company, etc.
3	Shopping	Shopping plaza, shopping center, shops, duty-free shop, convenience store, digital electronics, supermarket, plants and pet market, home building materials market, etc.
4	Transportation	Airport, railway station, passenger port, tourist routes bus station, common bus station, parking lot related, etc.
5	Education	Museum, exhibition Hall, convention and exhibition center, art gallery, library, planetarium, cultural palace, university and college, middle school, etc.
6	Residential	Hotel, residential area, villa, residential quarter, dormitory, community center, etc.
7	Hospital	Hospital, health center, clinic, disease prevention, pharmacy, medical supplies, etc.
8	Government	Governmental organization and institution, foreign embassy and consulate, representative office of international organizations, etc.

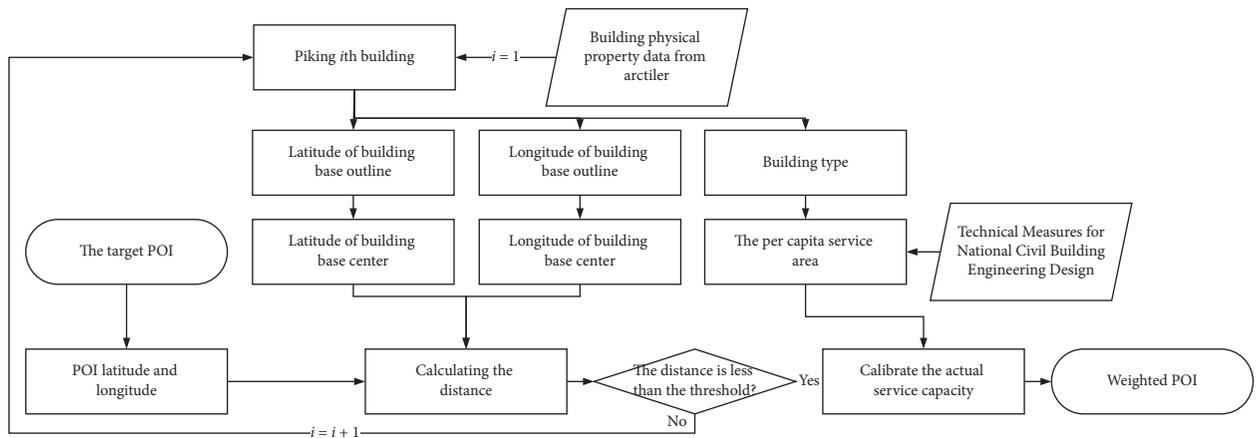


FIGURE 2: Fusion of multisource, geo-based data.

3.2. Embedding Spatiotemporal Semantics in Trip Records.

A passenger trip record R from AFC system is composed of four components, namely, the tap-in time t_{in}^R , the tap-in station s_{in}^R , the tap-out time t_{out}^R , and the tap-out station s_{out}^R . In this paper, the above four components are

transformed into five attribute vectors to describe the passenger trip. They are the origin station vector O^R , the destination station D^R , start time of the day T^R , the day of week W^R , and travel time H^R . Symbolically, a trip record R corresponds to a vector \mathbf{R} , which can be represented as

$\{\mathbf{O}^R, \mathbf{D}^R, \mathbf{T}^R, \mathbf{W}^R, \mathbf{H}^R\}$. In this section, the goal is to represent the above attributes as spatiotemporal semantics in the form of vectors for subsequent mobility pattern recognition. To do this, a SAE-based framework is built to embed spatiotemporal semantics, the structure of which is shown in Figure 3. First, weighted POIs calibrated in Section 3.1 and one-hot encoding are addressed to generate spatial/temporal attribute vectors. Subsequently, the above vectors are assembled to form a high-dimensional sparse trip vector. This method proved to be reasonable and feasible [7, 21]. It should be noted that although the high-dimensional vector contains a variety of travel information, the sparsity makes the mobility pattern difficult to be recognized effectively. To solve this problem, we train a SAE model to transform the high-dimensional trip vector into a low-dimensional dense vector to represent spatiotemporal semantics. Here are the details.

In the existing researches, the radius of the service area of a metro station is generally set as 500 meters [18, 21].

Therefore, in terms of spatial semantic, weighted POIs within 500 meters of the target station are utilized to represent the station. Define P as the set of all weighted POIs in the research area. For the tap-in station s_{in}^R and the tap-out station s_{out}^R , the weighted POIs within 500 meters can be expressed as follows:

$$p_{s_{in}^R} = \{p | \text{distance}(p, s_{in}^R) \leq 500, \forall p \in P\}, \quad (3)$$

$$p_{s_{out}^R} = \{p | \text{distance}(p, s_{out}^R) \leq 500, \forall p \in P\}. \quad (4)$$

As shown in Table 1, the weighted POIs have been divided into 8 categories, so \mathbf{O}^R and \mathbf{D}^R can each be represented as an 8-dimensional vector. The value of a weighted POI represents its service capacity, and the larger the value, the greater the probability of becoming the departure point or destination point of passengers at the station. \mathbf{O}^R and \mathbf{D}^R can be expressed as follows:

$$\mathbf{O}^R = \left\{ \frac{\sum |p_1|}{\sum |p_{in}^R|}, \frac{\sum |p_2|}{\sum |p_{in}^R|}, \dots, \frac{\sum |p_8|}{\sum |p_{in}^R|} \right\}, \quad p_i \in p_{in}^R, i = 1, 2, \dots, 8, \quad (5)$$

$$\mathbf{D}^R = \left\{ \frac{\sum |p_1|}{\sum |p_{out}^R|}, \frac{\sum |p_2|}{\sum |p_{out}^R|}, \dots, \frac{\sum |p_8|}{\sum |p_{out}^R|} \right\}, \quad p_i \in p_{out}^R, i = 1, 2, \dots, 8. \quad (6)$$

where $\sum |p_i|$ represents the sum of value of weighted POI of the i th category, $\sum |p_{in}^R|$ and $\sum |p_{out}^R|$ represent the sum of all weighted POIs within 500 meters of the tap-in station s_{in}^R and the tap-out station s_{out}^R . The order of POI categories corresponds to the row order in Table 1, namely, Entertainment, Working, Shopping, Transportation, Education, Residential, Hospital, and Government.

As for temporal semantic, one-hot coding is adopted to represent three attributes. For the convenience of expression, we divide a day into several discrete slots with a fixed interval. The metro service is not available between 0 am and 5 am. Here, the interval is set to be one hour, resulting in 19 slots in a day. So \mathbf{T}^R can be easily characterized as a 19-dimensional vector. For example, if t_{in}^R is 5:16:29 (between 5 and 6), it can be expressed as $\{1, 0, 0, \dots, 0\}$. If t_{in}^R is 22:51:33 (between 22 and 23), it can be expressed as $\{0, 0, 0, \dots, 0, 1, 0\}$. Similarly, because there are 7 days a week, \mathbf{W}^R can be represented as a 7-dimensional vector. If t_{in}^R is on Monday, it can be expressed as $\{1, 0, 0, 0, 0, 0, 0\}$. As for travel time, since most passengers travel within 240 minutes, we divide the travel time into 8 slots with the interval of 30 minutes [33]. If the travel time of R is 57 minutes (between 30 and 60), \mathbf{H}^R can be expressed as $\{0, 1, 0, 0, 0, 0, 0\}$. In summary, the trip vector $\mathbf{R} = \{\mathbf{O}^R, \mathbf{D}^R, \mathbf{T}^R, \mathbf{W}^R, \mathbf{H}^R\}$ has been represented as a 50-dimensional $(8 + 8 + 19 + 7 + 8)$ sparse vector.

We train a SAE model to extract the mixed spatiotemporal semantics of trip record R to avoid the adverse

effects of the sparsity of high-dimensional vector [34]. Essentially, the auto-encoder is an unsupervised algorithm that can automatically learn features from unlabeled data and can give a better feature description than the original data. It can be regarded as a neural network, which automatically generates an optimal coding strategy by continuously optimizing the weight parameters, resulting in the output vector being consistent with the input vector. As an extension of the classic auto-encoder, SAE is a deep neural network model constructed by stacking multiple auto-encoders, where the output of the n th layer of auto-encoder is used as the input of the $(n + 1)$ th layer of auto-encoder [35]. Structurally, SAE can be divided into two components, namely, the encoder and decoder. The former transforms the input sparse vector into a dense vector through several layers of coding, and the latter is the reverse process of the former to reconstruct high-dimensional vectors. As shown in Figure 3, the input 50-dimensional sparse vector \mathbf{R} is firstly upgraded to a 64-dimensional vector to extract abstract features, and then the dimensionality is reduced to 16-dimensional and 8-dimensional vectors layer by layer to realize the representation of dense vectors. Formulaically, the above process can be expressed as follows:

$$\mathbf{h}_{n+1} = f_a(\mathbf{W}_n \mathbf{h}_n + \mathbf{b}_n), \quad (7)$$

where \mathbf{h}_n and \mathbf{h}_{n+1} represent the output vector of the n th and the $(n + 1)$ th layer, \mathbf{W}_n and \mathbf{b}_n represent the weight parameter

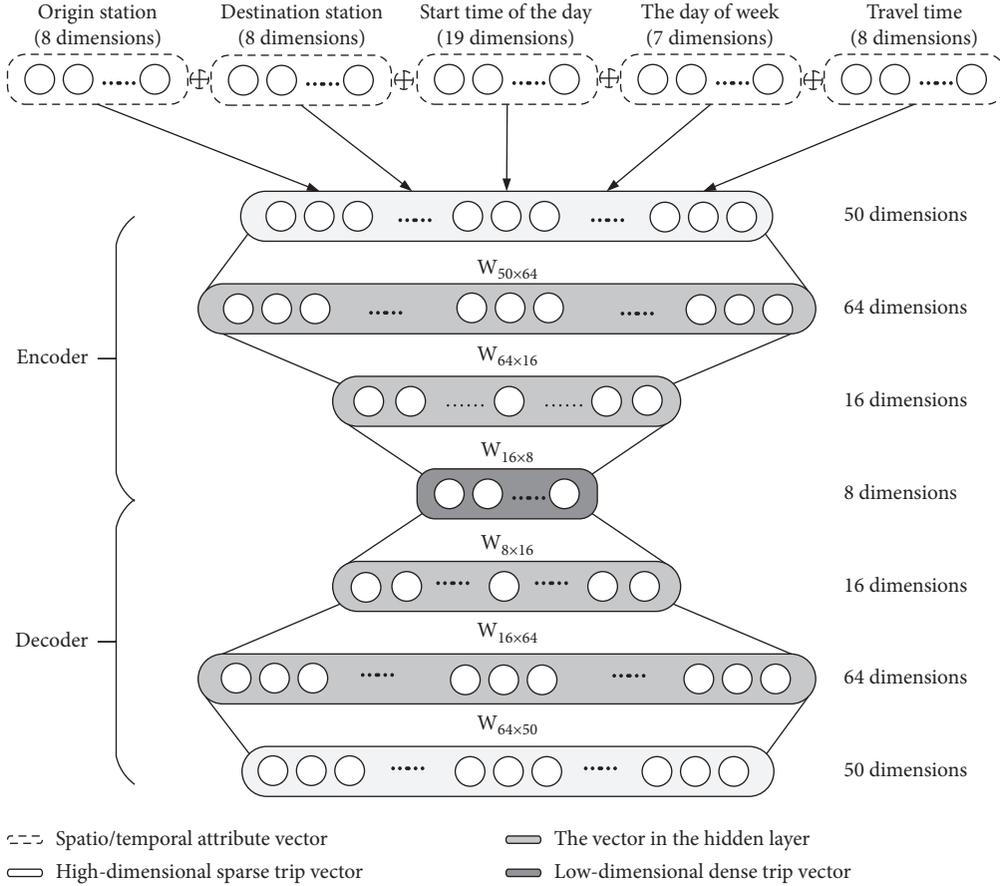


FIGURE 3: Embedding spatiotemporal semantics using SAE.

matrix and the bias from the n th layer to the $(n + 1)$ th layer, and $f_a(\cdot)$ represents the activation function, which is the rectified linear unit (ReLU) in this paper. It can be seen that the parameters that need to be estimated in the model are \mathbf{W}_n and \mathbf{b}_n . Particularly, when $n = 1$, \mathbf{h}_n is \mathbf{R} . Since the dimension of h_1 is smaller than that of h_2 ($50 < 64$), it is necessary to avoid invalid training of the weight parameters [36]. The weight parameters of this layer need to be pretrained, where the greedy layer-wise pre-training method is used. See details in reference [37]. The loss function is constructed as follows and the regularization is used in this process.

$$\text{loss} = L(x, g(f(x))) + \Omega(h). \quad (8)$$

Here, $h = f(x)$ represents the output of the encoder, and $g(f(x))$ represents the output of the decoder. Besides, $L(x, g(f(x)))$ represents the difference between x and $g(f(x))$, which can be measured by the mean square error (MSE). Further, $\Omega(h)$ represents the regularization term, which is the l_1 -norm here. Using the above procedure, the weight parameters of this layer can be initialized. As for the weight parameters of other layers, truncated normal initializer can be used.

And then, MSE is chosen as the loss function of the whole SAE. Define the dense vector as $\mathbf{R}_{\text{dense}}$ and the output reconstruction high-dimensional vector as \mathbf{R}_{rc} , then the loss function f_{loss} can be expressed as follows:

$$f_{\text{loss}} = \text{MSE}(\mathbf{R}_{\text{rc}}) = \frac{\sum_N (\mathbf{R}_i - \mathbf{R}_{\text{rc},i})^2}{N}, \quad (9)$$

where N represents the total number of trip records, \mathbf{R}_i and $\mathbf{R}_{\text{rc},i}$ represent the i th element in vector \mathbf{R} and \mathbf{R}_{rc} . As for training parameters, back propagation is used to fine-tune the parameters based on the value of the loss function. In this way, \mathbf{R} is converted to $\mathbf{R}_{\text{dense}}$.

3.3. Mobility Pattern Recognition Based on the Embedded Vectors. The goal of this section is to cluster $\mathbf{R}_{\text{dense}}$ through the cluster algorithm and achieve mobility pattern recognition through the spatiotemporal characteristics (obtained by decoder) displayed by the clustering results. It is found that passenger trajectories tend to show a high degree of temporal and spatial regularity. Passengers follow simple reproducible patterns, indicating that each individual is characterized by a significant probability to return to a few highly frequented locations [38, 39]. Since the $\mathbf{R}_{\text{dense}}$ obtained in the previous section is a dense vector with spatiotemporal semantics, we can identify mobility patterns by clustering these dense vectors. In this section, the DBSCAN algorithm, a density-based clustering method, is applied to cluster dense trip vectors. For two trip vectors (i.e., $\mathbf{R}_{\text{dense}}$) containing mixed spatiotemporal information, the distance

between them represents their spatiotemporal similarity. Additional details of the DBSCAN algorithm can be found in the study by [22]. One advantage of DBSCAN is that the number of clusters does not need to be manually specified in advance, which greatly reduces human intervention [40]. Instead, two parameters, the parameter of sample neighborhood size δ and the parameter of distance ϵ , are designed to describe the relationship between different samples to achieve clustering [31]. Here, we define a core sample to mean that there are at least δ other samples within the ϵ distance of a sample in the data set, and these samples are designated as neighbors of the core sample. For the trip vector, a core sample indicates that there are at least δ samples in the data set that have a spatiotemporal similarity less than ϵ . The flowchart of DBSCAN algorithm is shown in

$$\text{neighbor}(\mathbf{R}_{\text{dense}}) = \{\mathbf{R}_{\text{dense}}^i | d_m(\mathbf{R}_{\text{dense}}, \mathbf{R}_{\text{dense}}^x) \leq \epsilon, \mathbf{R}_{\text{dense}} \neq \mathbf{R}_{\text{dense}}^x, \forall \mathbf{R}_{\text{dense}}^x \in \mathbf{R}_D\}. \quad (11)$$

The condition that $\mathbf{R}_{\text{dense}}$ is the core sample can be expressed as follows:

$$\text{neighbor}(\mathbf{R}_{\text{dense}}) \geq \delta. \quad (12)$$

It needs to be clear that the values of parameters δ and ϵ need to be set in conjunction with the characteristics of the data set and the clustering target. Different values of the parameters have a significant impact on the clustering results. Here, two indicators are used to quantify algorithm performance, namely, the within-cluster sum of squared errors (SSE) and the silhouette coefficient (SC) [41]. Among them, SSE reflects the difference between different passengers who are identified as having the same mobility pattern. SSE in this paper can be calculated as follows:

$$\text{SSE} = \sum_{k=1}^K \sum_{m=1}^{M_k} \sum_{i=1}^8 (\mathbf{R}_{\text{dense},i}^m - \mathbf{R}_{\text{dense},i}^{K\mu})^2, \quad (13)$$

where K represents the number of clusters, M_k represents the number of samples in the k th cluster, $\mathbf{R}_{\text{dense},i}^m$ represents the i th element in the m th vector of the k th cluster, and $\mathbf{R}_{\text{dense},i}^{K\mu}$ represents the i th element in the center vector of the k th cluster. The smaller the value of SSE, the better the clustering performance. It means that passengers who are recognized as having the same mobility pattern have smaller identifiable differences, indicating that the pattern recognition is accurate. Besides, SC is a comprehensive index that combines cohesion and separation. Among them, the cohesion reflects the average difference between an individual passenger and other passengers identified as having the same mobility pattern. On the contrary, the separation means the smallest difference between the individual passenger and passengers with other mobility patterns. And then, SC in this paper can be expressed as follows:

$$\text{SC} = \frac{1}{|\mathbf{R}_D|} \cdot \sum_{m=1}^{|\mathbf{R}_D|} \frac{b_m - a_m}{\max(a_m, b_m)}, \quad (14)$$

Figure 4. It can be seen that the key of the algorithm is to determine whether the sample is the core sample using the two parameters (δ and ϵ). Formally, assuming that the set of all dense trip vector is \mathbf{R}_D , given two dense trip vectors $\mathbf{R}_{\text{dense}}^1$ and $\mathbf{R}_{\text{dense}}^2$, $\mathbf{R}_{\text{dense}}^1, \mathbf{R}_{\text{dense}}^2 \in \mathbf{R}_D$, the Manhattan distance is used to represent the difference in spatiotemporal semantics between them, which can be written as follows:

$$d_m(\mathbf{R}_{\text{dense}}^1, \mathbf{R}_{\text{dense}}^2) = \sum_{i=1}^8 |\mathbf{R}_{\text{dense},i}^1 - \mathbf{R}_{\text{dense},i}^2|, \quad (10)$$

where $\mathbf{R}_{\text{dense},i}^1$ and $\mathbf{R}_{\text{dense},i}^2$ represent the i th element in vector $\mathbf{R}_{\text{dense}}^1$ and $\mathbf{R}_{\text{dense}}^2$. The neighbor of the given trip vector $\mathbf{R}_{\text{dense}}$ can be expressed as follows:

where a_m reflects the degree of cohesion within a cluster, and b_m reflects the degree of separation between clusters. Specifically, for a trip vector $\mathbf{R}_{\text{dense}}^m$ ($\mathbf{R}_{\text{dense}}^m \in \mathbf{R}_D$), belonging to the k th cluster, the corresponding values of a_m and b_m can be calculated as follows:

$$a_m = \frac{1}{M_k - 1} \sum_{x=1, x \neq m}^{M_k} \sum_{i=1}^8 (\mathbf{R}_{\text{dense},i}^m - \mathbf{R}_{\text{dense},i}^x)^2, \quad (15)$$

$$b_m = \min(b_{m,k'}), \quad k' \in (1, 2, \dots, K), k' \neq k, \quad (16)$$

$$b_{m,k'} = \frac{1}{M_{k'} - 1} \sum_{x=1, x \neq m}^{M_{k'}} \sum_{i=1}^8 (\mathbf{R}_{\text{dense},i}^m - \mathbf{R}_{\text{dense},i}^x)^2. \quad (17)$$

Indeed, from the above formulation, it can be seen that $-1 \leq \text{SC} \leq 1$. If SC is close to 1, the data are well-clustered, indicating that the mobility pattern recognition is good. That is, the spatiotemporal characteristics of an individual passenger are highly similar to those of passengers in the same cluster. In contrast, passengers with different identified mobility patterns have significant differences in the spatiotemporal characteristics of travel. When SC is negative or even close to -1 , it indicates that passengers with different travel spatiotemporal characteristics are identified as having the same pattern, which is obviously not ideal. In summary, the smaller SSE and larger SC (close to 1) characterize better mobility pattern recognition results.

4. Case Study and Applications

4.1. Case Description. A case study of Beijing metro network is presented to evaluate the proposed method. A total of 176.81 million passenger travel records from September to October 2018 are acquired to identify mobility patterns. Correspondingly, the POI data in Beijing during this period is also crawled from Amap.

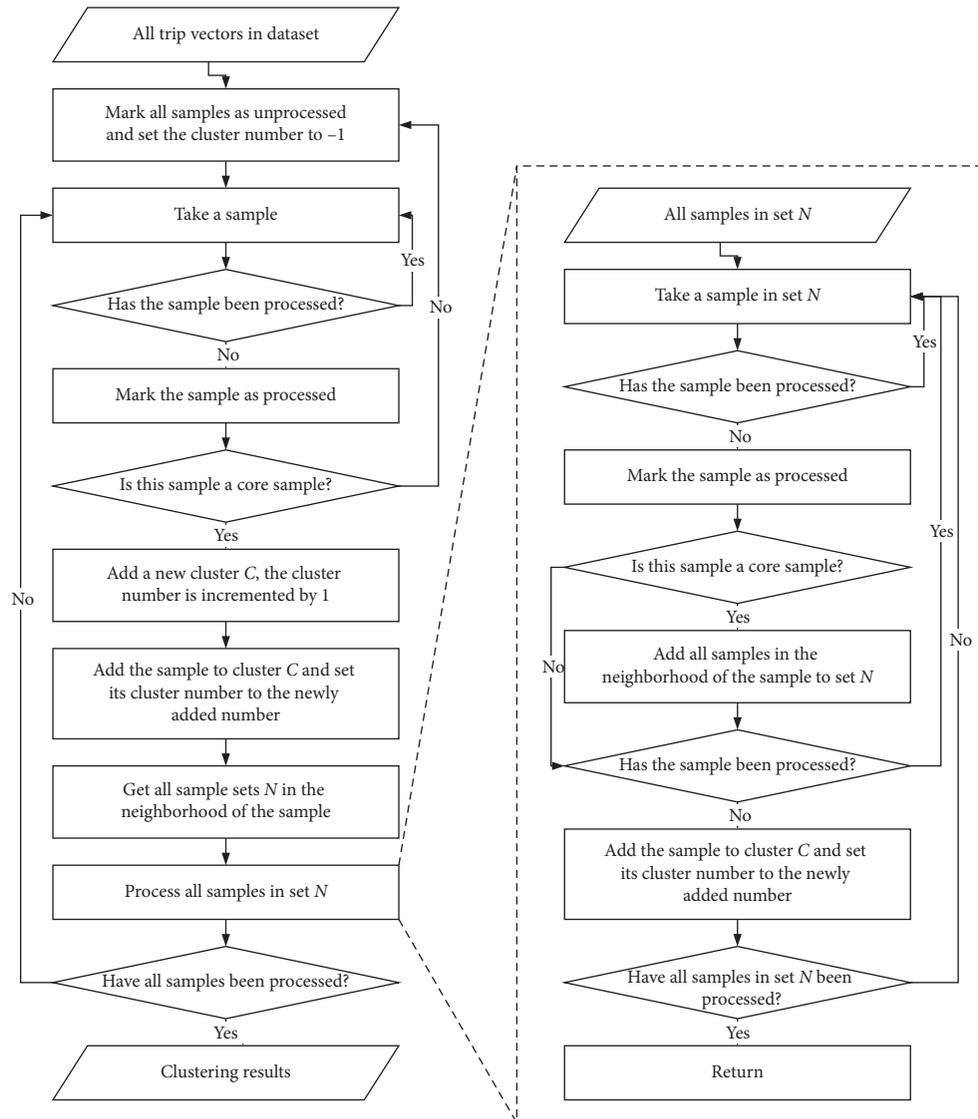


FIGURE 4: DBSCAN algorithm flowchart.

First, the weighted POIs are generated by fusing multisource data from Lianjia and Arciler using the method designed in Section 3.1. A total of 11,382 residential POIs were captured from Amap within the influence area of the metro station. Among them, 10927 residential POIs were successfully matched through the residential area data from Lianjia, indicating that the matching rate reached 96%. As for building data, a total of 6,887 buildings were captured from Amap. Among them, 6336 buildings were correctly matched through the Arciler datasets, indicating a 92% match rate. It can be found that although there are some matching failures, the matching rates were higher than 90%, which proves that the proposed method can effectively weight the original POI data into weighted POIs. Residential POIs are used as examples to illustrate the advantages of weighted POI data, as shown in Figure 5. Among them, Figure 5(a) shows the distribution of Beijing metro stations, while Figures 5(b) and 5(c), respectively, show the distribution heat map of raw POIs and weighted POIs within 500

meters of metro stations. It can be seen that the residential POIs in Figure 5(b) are more evenly distributed and have a higher density in the urban center. On the contrary, the residential POIs in Figure 5(c) are concentrated in suburban areas in an extremely uneven manner. The reason for the above difference is that the residential POIs in the central area of the city are mainly hotels, villas, and residential buildings with few floors, while that in the suburban areas are mainly high-density, multistory residential communities. Furthermore, 4 high-density residential areas can be clearly observed in Figure 5(c), which are located in the north, east, and southwest of the city. Comparing existing studies, it can be found that the above regions correspond to Changping, Tongzhou, Fangshan, and Daxing, respectively [42–44]. The above areas have similar characteristics, such as low housing prices, high housing density, and a large number of commuters living in the area. It shows that weighted POI data can more accurately reflect the categories of land use around metro stations.

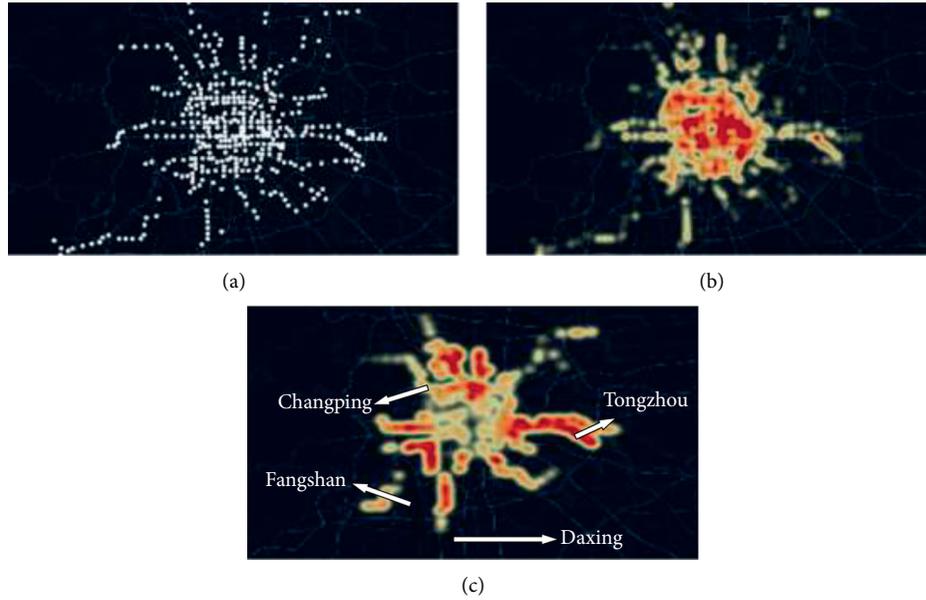


FIGURE 5: Comparison of weighted POIs and raw POIs. (a) The distribution of metro stations. (b) The distribution of raw POIs within 500 meters of metro stations. (c) The distribution of weighted POIs within 500 meters of metro stations.

Second, the spatiotemporal semantics are embedded using the SAE-based framework in Section 3.2. Figure 6 shows how the MSE changes with the number of iterations when training the SAE model. It can be seen that when the number of iterations reaches 40, the value of MSE remains stable. That is, SAE can encode the spatiotemporal features of the input trip records in a stable way, transforming the high-dimensional sparse vectors into low-dimensional dense vectors.

Third, the dense trip vectors are clustered using DBSCAN algorithm to realize the mobility pattern recognition. Since the number of clusters is not manually specified but is automatically generated according to the parameters δ and ϵ , it is necessary to check the number of clusters and algorithm performance corresponding to different values of parameters. This paper aims to identify passenger mobility patterns, so the number of clusters is required not to be too large (difficult to explain the potential activities of passengers) or too small (difficult to distinguish passenger categories) in order to balance practicality and interpretability. Through pre-experiments, we found that the number of clusters decreases as δ and ϵ increase. Further, when $\delta < 8$ and $\epsilon < 7$, the number of clusters is verified to be greater than 30, which makes it difficult to accurately describe the potential activities represented by each mobility pattern. When $\delta > 18$ and $\epsilon > 10$, the number of clusters is less than 3, which is obviously not conducive for our exploration of passenger mobility patterns. Therefore, the parameter value range is determined as: $\delta \in [8, 18]$ and $\epsilon \in [7, 10]$. Figure 2 lists several results of the number of clusters and algorithm performance quantified by SSE and SC under different parameter values. It can be found that the value of SSE decreases with the increase of δ , and the influence of ϵ on SSE is limited. The relationship between SC and parameters is more complicated. Furthermore, the relationship between δ ,

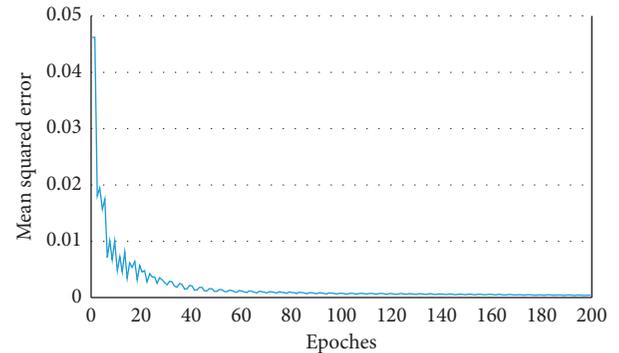


FIGURE 6: Changes in MSE during training SAE.

ϵ , and SC is shown in Figure 7. From a global perspective, SC increases with the increase of parameters δ and ϵ . When δ reaches 16 and ϵ reaches 9.5, the value of SC decreases with the increase of parameters. Combining the above two indicators, a parameter combination of $\delta = 16$ and $\epsilon = 9.5$ is selected. Herein, $SSE = 23715$ and $SC = 0.815$, showing good clustering performance.

4.2. Results Analysis. The mobility pattern is recognized using the proposed method with the above parameters. Figure 8 shows the results when $\delta = 16$ and $\epsilon = 9.5$. Each color represents a recognized mobility pattern and C1–C6 means the mobility features of cluster 1 to cluster 6. Among them, Figures 8(a) and 8(b) show the distribution of POI categories around the origin station and that around the destination station, which reveals the spatial features. The distributions of the start time of the day, the distribution of the day of week, and the distribution of travel time are presented in Figures 8(c)–8(e), respectively.

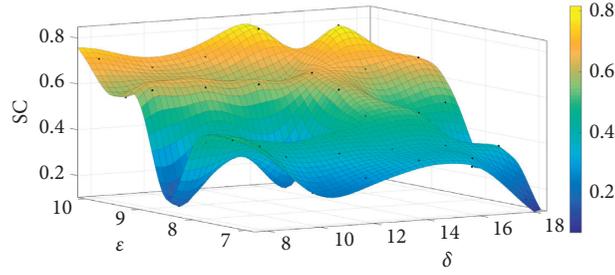


FIGURE 7: The value of SC corresponding to different parameters.

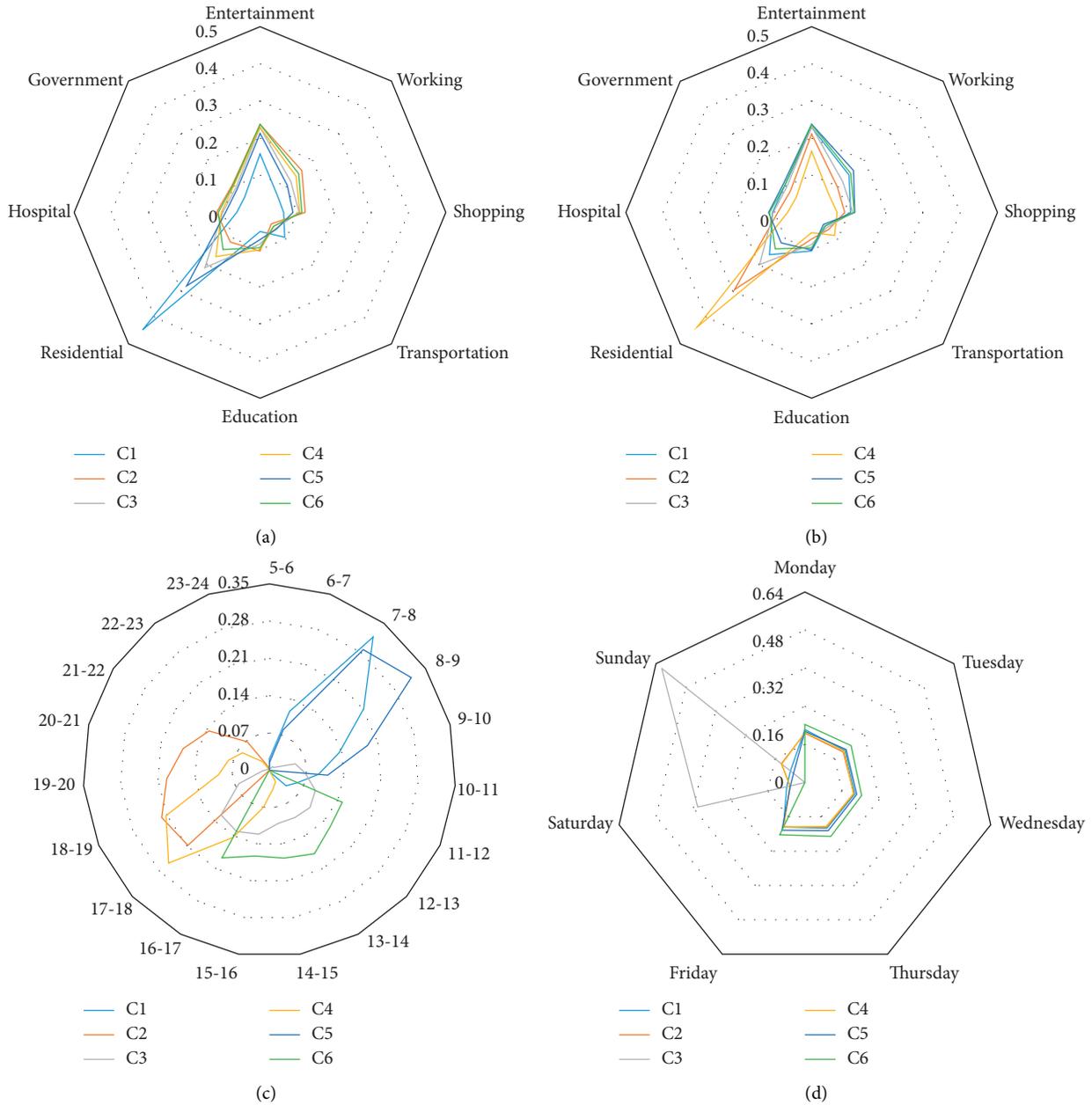


FIGURE 8: Continued.

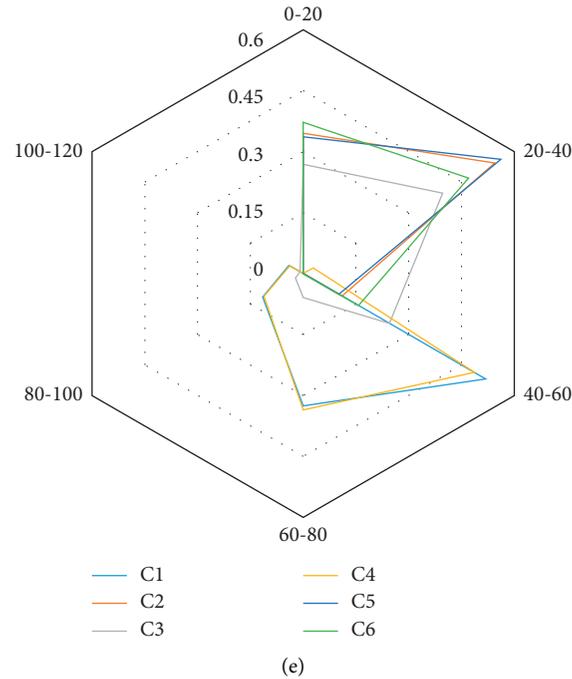


FIGURE 8: Six mobility patterns recognized by the proposed method when $\delta = 16$ and $\varepsilon = 9.5$. (a) POI categories around the origin station. (b) POI categories around the destination station. (c) The start time of the day. (d) The day of week. (e) Travel time.

The characteristics of the six mobility patterns identified above are summarized, as shown in Table 2. Among them, C1 and C5 account for 35.808% (13.716% + 22.092%), representing the work-related mobility during the workdays. More specifically, C1 reveals long-distance working mobility, where the start time is between 7 and 8 am, and travel time is mainly 40–80 min. In contrast, C5 represents short-distance working, where the start time is between 7 and 9 am (later than the start time in C1), because travelers need to spend short travel time (mainly within 40 min). It can be found that although the temporal information is in line with the typical mobility patterns of commuters, the POIs around the destination station include multiple categories (not only working), such as entertainment, working, hospital, and shopping, which characterize the various possible work places of passengers. Besides, C3 (accounting for 13.908%) shows entertainment and shopping activities that mainly take place on weekends due to the large number of entertainment and shopping POIs surrounding the destination station. The start time of this type of mobility is between 9 am and 7 pm, and the travel time is within 60 min. Correspondingly, C2 and C4 account for 34.323% (19.817% + 14.506%), revealing the home-related mobility, most of which occurs on weekdays and Sundays. It can be seen that the destination POIs are mainly residential. The difference is that the start time of the mobility represented by C2 is all after 5 pm, while that represented by C4 is mainly concentrated between 5 pm and 7 pm. In C2 and C4, the various types of POIs (entertainment, working, shopping, etc.) around the origin station represent passengers at different working locations. Finally, C6 (accounting for 15.961%) represents a kind of

mobility pattern, wherein it is difficult to directly identify the purpose of travel, where the origin location is mainly entertainment, shopping, and hospital POIs, the start time is between 11 am and 5 pm, and the travel time is within 40 min. The travel purpose of this pattern is difficult to be accurately identified, but it can be regarded as a short-distance travel that occurs during off-peak hours on weekdays.

The above analysis shows that mobility patterns related to working and home are the easiest to identify and explain, which is consistent with the conclusions of existing studies [23, 45, 46]. On the one hand, according to multidimensional temporal features, work-related mobility patterns can be divided into long-distance mobility and short-distance mobility. In this case, the number of short-distance travelers is 1.611 times (22.092%/13.716%) that of long-distance travelers, which shows that a large percentage of commuters work close to their places of residence. Nevertheless, there are still many commuters living far away from their working places, reflecting a serious imbalance between working and housing [43, 44]. On the other hand, home-related mobility patterns encompass more categories, because travelers can choose the time to go home more freely than the time to work. Taking C2 and C4 as examples, trips related to going home are clearly divided into two patterns. The start time of C2 is after 5 pm, and that of C4 is mainly between 5 and 7 pm. It can be inferred that the start time of the traveler's home trip is related to their work. In addition to working and going home, activities related to entertainment and shopping are displayed in C3. Most of them appear on weekends and their start time is between 9 am and 7 pm, which shows that passengers are more casual in choosing

TABLE 2: Characteristics of mobility patterns.

ID	Proportion (%)	Spatial features		Temporal features			Possible activity
		The origin station	The destination station	The start time	The day of week	Travel time	
C1	13.716	Mainly residential POIs	Mainly entertainment, working, hospital, and shopping POIs	Mainly 7–8	Mainly weekdays	Mainly 40 min–80 min	Working (long distance)
C2	19.817	Mainly entertainment, working, shopping, and education POIs	Mainly residential POIs	After 17	Weekdays and Sundays	Within 40 min	Home (short distance)
C3	13.908	Mainly residential, entertainment POIs	Mainly entertainment and shopping POIs	9–19	Mainly weekends	Within 60 min	Entertainment and shopping
C4	14.506	Mainly entertainment, working, and shopping POIs	Mainly residential POIs	Mainly 17–19	Weekdays and Sundays	Mainly 40 min–80 min	Home (long distance)
C5	22.092	Mainly residential POIs	Mainly entertainment, working, shopping, and POIs	Mainly 7–9	Mainly weekdays	Mainly within 40 min	Working (short distance)
C6	15.961	Mainly entertainment, shopping, and hospital POIs	Mainly entertainment, shopping, and residential POIs	Mainly 11–17	Weekdays	Mainly within 40 min	Others

start time when engaging in entertainment and shopping activities. The above phenomenon is consistent with our empirical observation [18, 23]. It should be noted that the current analysis is based on the parameter settings of $\delta = 16$ and $\varepsilon = 9.5$. When the number of clusters decreases, several work-related patterns may be merged. Conversely, when the number of clusters increases, more mobility patterns may be found, but the difficulty of interpreting the pattern recognition results also increases.

It should be noted that sometimes the spatial information of the clustering results is confusing. For example, both clusters C1 and C3 have trips from residential POI to shopping POI. Nevertheless, C1 and C3 are interpreted as different potential activities (long-distance working/entertainment and shopping). This reflects the uncertainty of identifying passenger mobility patterns only through spatial information and the necessity of using spatiotemporal information jointly. For trips with the similar spatial information, temporal information can assist in inferring mobility patterns. Passengers who intend to shop are unlikely to choose to travel during the morning peak hours. They tend to choose off-peak hours to avoid crowded conditions and obtain higher travel comfort. It can be inferred that passengers in C1 who travel during the morning rush hours with shopping POIs as destinations are composed of most of the staff working in the mall and a small number of shoppers. Conversely, in C3, the potential activity of passengers traveling on weekends with shopping POIs as destinations is more likely to be shopping. When classifying a passenger's mobility pattern, the proposed embedding method can be used to embed the passenger's spatiotemporal information into a low-dimensional vector space. The distance between the embedded vector and the vector of each cluster center can be calculated to obtain the most likely mobility patterns and reduce the confusion caused by spatial information.

4.3. Sensitivity Analysis of Parameters. In this section, the sensitivity of parameters on the recognition results is analyzed. As shown in Table 3, the parameters of the clustering algorithm have a significant impact on the recognition performance. Here, we show the results when $\delta = 14$ and $\varepsilon = 10$ in Figure 9 and that when $\delta = 8$ and $\varepsilon = 10$ in Figure 10. In Figure 9, the trip vectors are divided into 3 patterns, $SSE = 23647$, and $SC = 0.793$. Obviously, it reveals the three most basic patterns of urban mobility: working, home, and others. Among them, C2 describes working-related trips, where the start time is mainly from 7 am to 9 am on weekdays, and the POIs around the origin station are mainly residential. Correspondingly, C3 represents trips related to going home, where the start time is mainly after 5 pm on weekdays, and the POIs around the destination station are dominated residential POIs. In addition, C3 represents trips that include entertainment, shopping, etc., where the start time is mainly distributed between 10 am and 5 pm on weekends. In Figure 10, the passenger trip vectors are identified as 11 clusters, $SSE = 23133$ and $SC = 0.712$. Compared with Figure 8, it can be seen that more passenger activities are identified.

Among them, C1, C9, and C11 are the three most easily explained patterns. They characterize working-related trips. In more detail, travel time of C1 is mainly 60–80 min, while that of C9 is within 40 min, and that of C11 is mainly 20–60 min. The travel time reflects the length of the journey. The three clusters C3, C8, and C10 represent home-related mobility. Their proportions are 13.606%, 7.478%, and 7.555%, respectively. In detail, the start time of C3 is mainly 5 pm–7 pm, while that of C8 is 7 pm–10 pm, and that of C10 is mainly 6 pm–8 pm. This shows that passengers are more flexible in time selection when going home. The remaining clusters represent mobility other than working and home, which are a refinement set of C3 and C6 in Table 2. Obviously, the mobility represented by these clusters is more dispersed in POI categories and more free

TABLE 3: The number of clusters and algorithm performance based on different parameters.

ID	δ	ϵ	K	SSE	SC
1	8	7	27	33136	0.466
2	8	7.5	29	33407	0.473
3	8	8	24	30895	0.439
4	8	8.5	18	34989	0.139
5	8	9	16	31018	0.621
6	8	9.5	15	31731	0.568
7	8	10	11	23133	0.712
8	10	7	24	28931	0.248
9	10	7.5	22	29168	0.38
10	10	8	19	27447	0.421
11	10	8.5	14	27925	0.361
12	10	9	13	27374	0.616
13	10	9.5	12	28828	0.621
14	10	10	9	27158	0.657
15	12	7	19	26676	0.295
16	12	7.5	17	26224	0.379
17	12	8	16	26786	0.424
18	12	8.5	11	26077	0.188
19	12	9	11	26775	0.614
20	12	9.5	10	26596	0.56
21	12	10	6	26638	0.686
22	14	7	18	25702	0.356
23	14	7.5	16	24084	0.381
24	14	8	12	25114	0.503
25	14	8.5	10	23429	0.597
26	14	9	9	23796	0.646
27	14	9.5	9	23889	0.596
28	14	10	3	23647	0.793
29	16	7	18	24333	0.348
30	16	7.5	14	22281	0.38
31	16	8	10	23117	0.502
32	16	8.5	10	23255	0.487
33	16	9	9	23760	0.646
34	16	9.5	6	23715	0.815
35	16	10	5	23568	0.596
36	18	7	17	22439	0.151
37	18	7.5	13	21892	0.364
38	18	8	10	23129	0.25
39	18	8.5	7	22538	0.507
40	18	9	7	23407	0.681
41	18	9.5	7	23436	0.629
42	18	10	6	23675	0.654

in the start time, which is in line with the diversified characteristics of weekend entertainment activities. Inevitably, as the number of clusters increases, the interpretability of the results is weakened. For example, there is no significant difference in the proportion of each category of POI around the origin station and the destination station in C4, which makes it difficult to find a known activity to explain its spatiotemporal characteristics. A feasible method is to investigate the purpose of the passengers in C4 to explain the above phenomenon. In summary, there must be a trade-off between the number of clusters and interpretability of the results.

4.4. Comparison of Methods. First, we compare the results with different vector forms. Here, sparse vectors (50 dimensions) and dense vectors (8 dimensions) are used to

identify passenger mobility patterns, respectively. We examine the performance with different vector forms when the number of clusters is 6. It should be noted that after many pre-experiments, when sparse vectors are used and the number of clusters is 6, the input parameters are $\delta = 22$ and $\epsilon = 72$. The comparison results are shown in Table 4. It can be seen that the calculation time using sparse vectors is much longer than that using dense vectors. This is because dense vectors need to consume less computing resources in the calculation process. In addition, compared to sparse vectors, using dense vectors can give better results, showing a smaller SSE and a larger SC. The reason is that the SAE-based embedding method efficiently extracts the spatiotemporal information in passenger travel records, which proves the necessity and superiority of embedding spatiotemporal semantics.

Next, we compare the performance of different methods. Here, two baseline methods are selected from the existing studies. The first one is a cluster-based method from literature [22]. Different from this paper, this method aims to mine the spatiotemporal travel patterns from the long-term historical travel database, whereas OD stations are regarded as spatial features and the timestamps of entering and exiting stations are regarded as temporal features. The second baseline method is a topic model based on LDA from literature [23]. In this model, the four features are considered to describe a passenger trip—they are the location (station), start time of day, start day of week, and the duration. It should be noted that this model is a “soft-cluster” method, where a probability distribution is used to quantify the relationship between a trip and mobility patterns.

Due to the lack of real activity labels for passenger travel records, it is challenging to quantify and compare the performance of various methods in mobility pattern recognition in terms of accuracy. One way to deal with this problem is to design a stated preference (SP) survey to determine the actual travel purpose of passengers, which can be utilized as a benchmark to calculate the accuracy of the mobility pattern recognition results [47]. Nevertheless, SP surveys often require huge manpower and material resources, especially in large-scale analysis. In this section, a compromise method is adopted to evaluate the performance of models by using the SSE calculated by equation (13) and the SC calculated by equation (14). These two indicators can measure the ability of the pattern recognition results to characterize the distribution of the data, evaluating the models without real activity labels [23]. Based on the data in Section 4.1, the number of clusters is set to 3, 6, and 11 respectively, and the above two methods are used to recognize mobility patterns. Figure 11 shows the values of the two indicators (SSE and SC) corresponding to the results obtained by different methods, in which the K represents the number of clusters. It can be found that when the number of cluster is 3 and 6, the SSE of baseline 2 and that of the proposed method are relatively small, while that of baseline 1 is larger. When the number of clusters is 11, the three methods have comparable SSE. This means a significant intra-cluster difference of identified mobility patterns when the OD stations are regarded as the spatial



FIGURE 9: Continued.

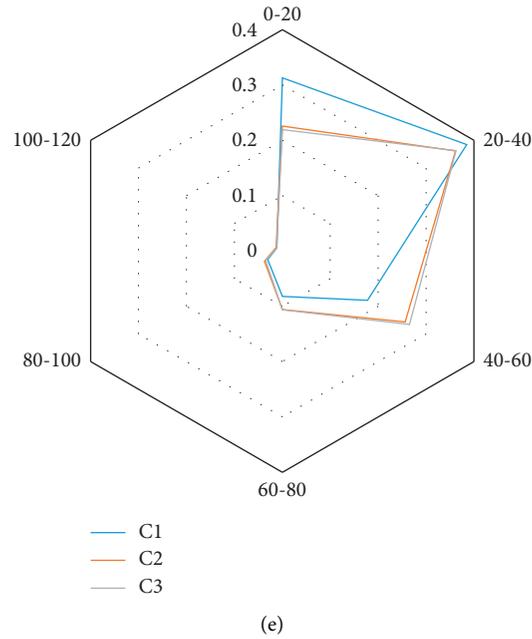


FIGURE 9: Three mobility patterns recognized by the proposed method when $\delta = 14$ and $\varepsilon = 10$. (a) POI categories around the origin station. (b) POI categories around the destination station. (c) The start time of the day. (d) The day of week. (e) Travel time.

features. Conversely, the proposed method and baseline 2 perform better in terms of SSE. Nevertheless, with the same number of clusters, the proposed method has a larger SC value than baseline 2. This shows that baseline 2 fails to distinguish the trips in different patterns well. In summary, the proposed method performs well in mobility analysis, which illustrates the necessity of using weighted POIs based on multisource data to characterize spatial attributes and the superiority of using coding-based methods to vectorize passenger trips.

4.5. Applications Based on Passenger Mobility Patterns.

The ultimate goal of mobility pattern recognition is to accurately grasp the characteristics of passenger needs and assist subway operators and managers to provide passengers with high-quality travel services. As described in Section 4.2, with the help of mobility pattern recognition, the time preferences, start location preferences, and the attributes of potential activities of different types of passengers can be explored. Furthermore, when a certain passenger's historical travel data are given, his/her mobility mode type can be calculated through similarity calculation, individual travel preferences can be estimated, and demand characteristics can be clarified. Based on this, it has become possible to provide personalized services according to individual travel needs.

On the one hand, individual mobility pattern helps generate more accurate personalized passenger guidance strategies. In traditional practice, metro operators empirically recommend the route with the shortest travel time or the lowest travel cost to passengers. Nevertheless, existing studies have shown that passengers with different travel

purposes pay different attention to different factors [47, 48]. For example, commuters may be more concerned about travel time reliability. On the contrary, travelers do not have high requirements for travel time reliability but are more concerned about the comfort of travel. The identification and analysis of mobility patterns can help provide personalized guidance strategies.

On the other hand, mobility pattern recognition can be used as a powerful tool to guide business applications. Here, the applications in advertising and Mobility-as-a-Service (MaaS) design and promotion are introduced. For advertisers, it would be wise to consider the passenger demand of the station when placing advertisements at a designated station, which can be obtained through the research of this paper. Related researches have been conducted in recent years to support mobility-pattern-based advertising [21, 49]. For example, it is obvious that in stations where many commuters live in the surrounding area, recruitment and job-hunting advertisements are very competitive. Besides, as a technological innovation with the potential to revolutionise the urban mobility paradigm, MaaS is emerging and closely related to mobility pattern recognition. MaaS is a service offered to the user in a single mobile app platform, which integrates all aspects of the travel experience, including booking, payment, and information, both before and during the trip [50]. The latest research shows that understanding passengers' mobility patterns and expectations is key for designing successful MaaS technologies [51]. And then, researches also show that willingness to use MaaS is strongly correlated with age and lifecycle stage, which can be identified by the proposed method in this paper [52]. For example, young individuals who are employed full-time are most likely to use MaaS.

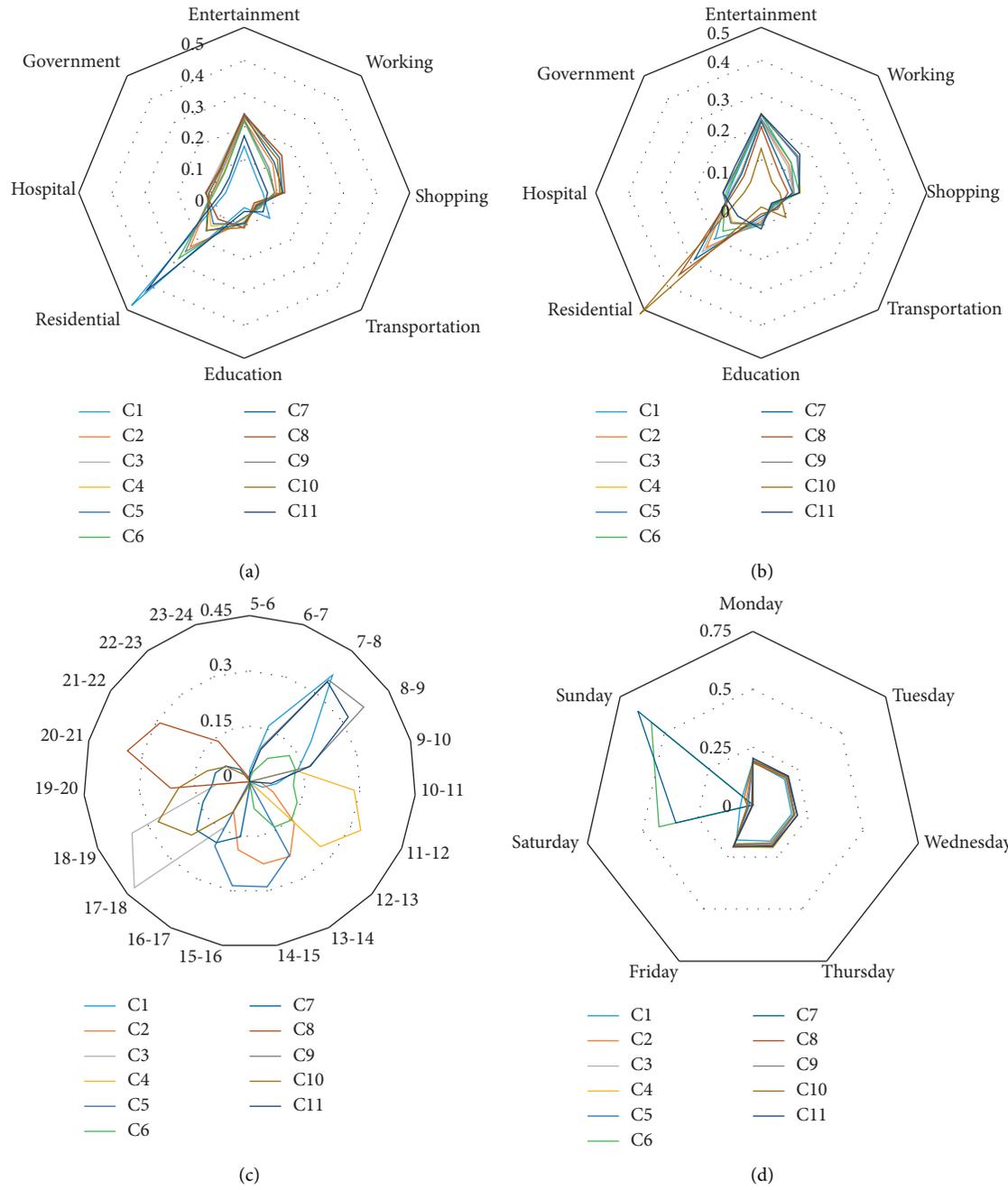


FIGURE 10: Continued.

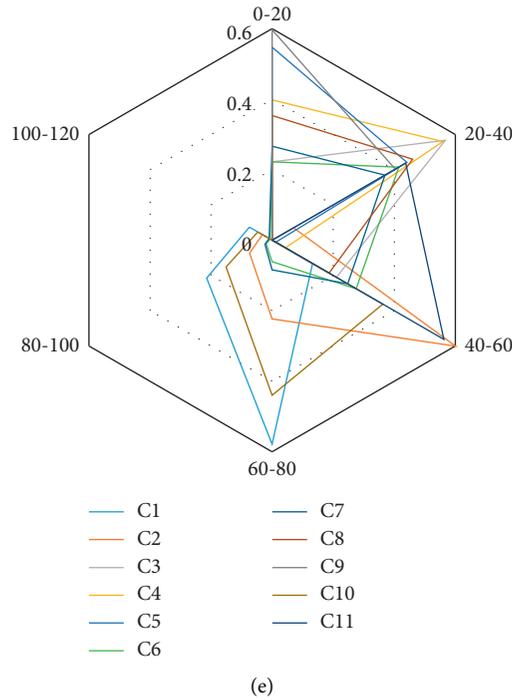


FIGURE 10: Eleven mobility patterns recognized by the proposed method when $\delta = 8$ and $\epsilon = 10$. (a) POI categories around the origin station. (b) POI categories around the destination station. (c) The start time of the day. (d) The day of week. (e) Travel time.

TABLE 4: Comparison of results with different vector forms.

Vector form	Computing time (second)	SSE	SC
Sparse vector	2674	28706	0.602
Dense vector	144	23715	0.815

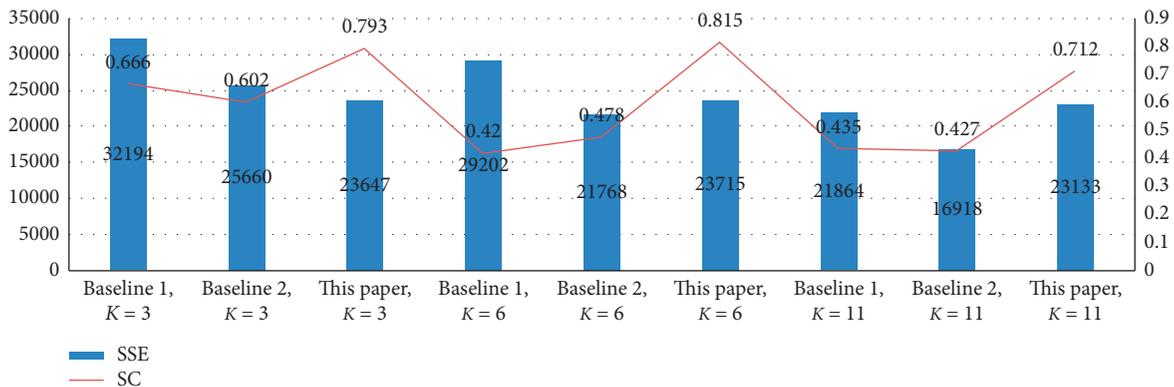


FIGURE 11: Comparison of methods.

It should be noted that mobility pattern recognition also has important applications in the prevention and control of epidemic spreading and the assessment of social and economic development. For details, please refer to references [10, 27].

5. Conclusions and Discussion

This paper presents a SAE-based unsupervised learning framework to explore the potential of AFC data in recognizing passenger mobility patterns. The proposed model

converts the travel records of passengers into trip vectors in an embedded manner to facilitate large-scale pattern recognition. Each trip vector contains spatial attributes (POIs around the origin station, POIs around the destination station) and temporal attributes (start time, day of the week, and travel time), which enhance the interpretability of the mobility analysis results. Specifically, the spatial characteristics are obtained through the fusion of multisource, geo-based data. A density-based clustering algorithm is introduced to group the trip vectors into multiple clusters to realize the

mobility pattern recognition. A case of Beijing metro network is used to verify the feasibility of the above methods. In this case, six typical mobility patterns are identified, two of which are related to working (accounting for 36.702%), three of which are related to home (accounting for 46.057%), and one of which is related to entertainment and studying (accounting for 17.242%), revealing the mobility distribution characteristics of Beijing metro passengers. Furthermore, the sensitivity analysis of the parameters is done. It is found that as the number of clusters in the results increases, the identified mobility patterns can reflect more detailed passenger activity characteristics and at the same time have greater inexplicability. The comparison with the other two baseline methods proves that the proposed method can better explore the passenger mobility patterns based on multisource data than the existing methods. This research provides a way of embedding complex, multisource, and different-dimensional spatiotemporal information into dense trip vectors, which is suitable for large-scale calculations to identify mobility patterns.

Admittedly, the proposed method still has several limitations that can be considered in the future works. First, geographic information needs to be processed more finely. This paper divides the captured POIs into 8 categories as shown in Table 1, and each category contains multiple subcategories. There may be great differences between subcategories. For example, Card & Chess Room and Camping Site are considered the same category (entertainment) in this paper. In fact, these two kinds of POIs can be treated separately as indoor entertainment and outdoor sports, which helps to discover more detailed features of passenger activities. Second, the dependence between multiple trips of a passenger needs to be considered. This paper only embeds the spatiotemporal features of the current trip into the dense vector, and does not consider the previous and subsequent trips, which limits the application of the proposed method in the generation of passenger activity chains and the prediction of trips [32]. For example, a point of view is widely agreed that, due to geographical constraints, the origin station of the current trip is likely to be the destination station of the previous trip. It means that considering the information of previous and subsequent trips to complete the embedding of the current trip has potential application value. Third, although the validity of the matching between the selected multiple data sources is acceptable, there are still some matching failures. The selection of data sources with better matching is worth exploring to improve the proposed data fusion method. This will be the focus of the future studies.

Data Availability

All data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Chao Yu was in charge of conceptualization, writing the original draft, providing software, and visualization. Haiying Li was concerned with methodology and project administration. Xinyue Xu was involved in conceptualization, methodology, reviewing, and editing. Jun Liu did supervision and funding acquisition. Jianrui Miao was responsible for methodology. Yitang Wang and Qi Sun contributed to data curation.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (Grant No. 71871012) and the State Key Lab of Rail Traffic Control and Safety of China (Grant No. RCS2020ZT005).

References

- [1] A. Alsker, A. Tavassoli, M. Mesbah, L. Ferreira, and M. Hickman, "Public transport trip purpose inference using smart card fare data," *Transportation Research Part C: Emerging Technologies*, vol. 87, pp. 123–137, 2018.
- [2] S. Wang, L. Li, W. Ma, and X. Chen, "Trajectory analysis for on-demand services: a survey focusing on spatial-temporal demand and supply patterns," *Transportation Research Part C: Emerging Technologies*, vol. 108, pp. 74–99, 2019.
- [3] Y. Yang, J. Liu, P. Shang, X. Xu, and X. Chen, "Dynamic origin-destination matrix estimation based on urban Rail transit AFC data: deep optimization framework with forward passing and backpropagation techniques," *Journal of Advanced Transportation*, vol. 2020, Article ID 8846715, 2020.
- [4] H. Li, Y. Wang, X. Xu, L. Qin, and H. Zhang, "Short-term passenger flow prediction under passenger flow control using a dynamic radial basis function network," *Applied Soft Computing*, vol. 83, Article ID 105620, 2019.
- [5] X. Xu, H. Li, J. Liu, B. Ran, and L. Qin, "Passenger flow control with multi-station coordination in subway networks: algorithm development and real-world case study," *Transportmetrica B: Transport Dynamics*, pp. 1–27, 2018.
- [6] K. Farrahi and D. Gatica-Perez, "Discovering routines from large-scale human locations using probabilistic topic models," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 1, pp. 1–27, 2011.
- [7] H. Cao, F. Xu, J. Sankaranarayanan, Y. Li, and H. Samet, "Habit2vec: trajectory semantic embedding for living pattern recognition in population," *IEEE Transactions on Mobile Computing*, vol. 19, no. 5, pp. 1096–1108, 2020.
- [8] M. Allahviranloo and W. Recker, "Daily activity pattern recognition by using support vector machines with multiple classes," *Transportation Research Part B: Methodological*, vol. 58, pp. 16–43, 2013.
- [9] S. Hasan and S. V. Ukkusuri, "Urban activity pattern classification using topic models from online geo-location data," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 363–381, 2014.
- [10] J. Blumenstock, G. Cadamuro, and R. On, "Predicting poverty and wealth from mobile phone metadata," *Science*, vol. 350, no. 6264, pp. 1073–1076, 2015.
- [11] L. Alexander, S. Jiang, M. Murga, and M. C. González, "Origin-destination trips by purpose and time of day inferred from mobile phone data," *Transportation*

- Research Part C: Emerging Technologies*, vol. 58, pp. 240–250, 2015.
- [12] Y. Xu, A. Belyi, I. Bojic, and C. Ratti, “Human mobility and socioeconomic status: analysis of Singapore and Boston,” *Computers, Environment and Urban Systems*, vol. 72, pp. 51–67, 2018.
- [13] T. Kusakabe and Y. Asakura, “Behavioural data mining of transit smart card data: a data fusion approach,” *Transportation Research Part C: Emerging Technologies*, vol. 46, pp. 179–191, 2014.
- [14] L.-M. Kieu, A. Bhaskar, and E. Chung, “A modified density-Based Scanning Algorithm with Noise for spatial travel pattern analysis from Smart Card AFC data,” *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 193–207, 2015.
- [15] G. Han and K. Sohn, “Activity imputation for trip-chains elicited from smart-card data using a continuous hidden Markov model,” *Transportation Research Part B: Methodological*, vol. 83, pp. 121–135, 2016.
- [16] L. Sun and K. W. Axhausen, “Understanding urban mobility patterns with a probabilistic tensor factorization framework,” *Transportation Research Part B: Methodological*, vol. 91, pp. 511–524, 2016.
- [17] M. K. El Mahrsi, E. Come, L. Oukhellou, and M. Verleysen, “Clustering smart card data for urban mobility analysis,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 712–728, 2017.
- [18] Q. Zou, X. Yao, P. Zhao, H. Wei, and H. Ren, “Detecting home location and trip purposes for cardholders by mining smart card transaction data in Beijing subway,” *Transportation*, vol. 45, no. 3, pp. 919–944, 2018.
- [19] J. B. Gordon, H. N. Koutsopoulos, and N. H. M. Wilson, “Estimation of population origin-interchange-destination flows on multimodal transit networks,” *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 350–365, 2018.
- [20] P. Kumar, A. Khani, and Q. He, “A robust method for estimating transit passenger trajectories using automated data,” *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 731–747, 2018.
- [21] D. Zhuang, S. Hao, D.-H. Lee, and J. G. Jin, “From compound word to metropolitan station: semantic similarity analysis using smart card data,” *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 322–337, 2020.
- [22] L. M. Kieu, A. Bhaskar, and E. Chung, “Passenger segmentation using smart card data,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1537–1548, 2015.
- [23] Z. Zhao, H. N. Koutsopoulos, and J. Zhao, “Discovering latent activity patterns from transit smart card data: a spatiotemporal topic model,” *Transportation Research Part C: Emerging Technologies*, vol. 116, Article ID 102627, 2020.
- [24] Z. Chen, Y. Zhang, C. Wu, and B. Ran, “Understanding individualization driving states via latent dirichlet allocation model,” *IEEE Intelligent Transportation Systems Magazine*, vol. 11, no. 2, pp. 41–53, 2019.
- [25] Z. Chen, H. Cao, H. Wang, F. Xu, V. Kostakos, and Y. Li, “Will you come back/check-in again? understanding characteristics leading to urban revisitation and Re-check-in,” *Proceedings of the ACM Interactive, Mobile, Wearable Ubiquitous Technol.* vol. 4, no. 3, 2020.
- [26] J. Wang, X. Kong, F. Xia, and L. Sun, “Urban human mobility,” *ACM SIGKDD Explorations Newsletter*, vol. 21, no. 1, pp. 1–19, 2019.
- [27] H. Barbosa, M. Barthelemy, G. Ghoshal et al., “Human mobility: models and applications,” *Physics Reports*, vol. 734, pp. 1–74, 2018.
- [28] Y. Zhang, N. Sari Aslam, J. Lai, and T. Cheng, “You are how you travel: a multi-task learning framework for geodemographic inference using transit smart card data,” *Computers, Environment and Urban Systems*, vol. 83, Article ID 101517, 2020.
- [29] M. Ren, Y. Lin, M. Jin, Z. Duan, Y. Gong, and Y. Liu, “Examining the effect of land-use function complementarity on intra-urban spatial interactions using metro smart card records,” *Transportation*, vol. 47, no. 4, pp. 1607–1629, 2020.
- [30] Y. Gong, Y. Lin, and Z. Duan, “Exploring the spatiotemporal structure of dynamic urban space using metro smart card records,” *Computers, Environment and Urban Systems*, vol. 64, pp. 169–183, 2017.
- [31] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, “DBSCAN revisited, revisited: why and how you should (still) use DBSCAN,” *ACM Trans. Database Syst.* vol. 42, no. 3, 2017.
- [32] Z. Zhao, H. N. Koutsopoulos, and J. Zhao, “Individual mobility prediction using transit smart card data,” *Transportation Research Part C: Emerging Technologies*, vol. 89, pp. 19–34, 2018.
- [33] C. Yu, H. Li, X. Xu, and J. Liu, “Data-driven approach for solving the route choice problem with traveling backward behavior in congested metro systems,” *Transportation Research Part E: Logistics and Transportation Review*, vol. 142, Article ID 102037, 2020.
- [34] F. Jia, H. Li, X. Jiang, and X. Xu, “Deep learning-based hybrid model for short-term subway passenger flow prediction using automatic fare collection data,” *IET Intelligent Transport Systems*, vol. 13, no. 11, pp. 1708–1716, 2019.
- [35] X. Yuan, B. Huang, Y. Wang, C. Yang, and W. Gui, “Deep learning-based feature representation and its application for soft sensor modeling with variable-wise weighted SAE,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3235–3243, 2018.
- [36] P. Baldi, “Autoencoders, unsupervised learning, and deep architectures,” *ICML Unsupervised and Transfer Learning*, vol. 27, pp. 37–50, 2012.
- [37] Y. Bengio, P. Lamblin, D. Popovici, and H. Larochelle, “Greedy layer-wise training of deep networks,” *Advances in Neural Information Processing Systems*, vol. 1, pp. 153–160, 2007.
- [38] M. C. González, C. A. Hidalgo, and A.-L. Barabási, “Understanding individual human mobility patterns,” *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [39] C. Song, T. Koren, P. Wang, and A.-L. Barabási, “Modelling the scaling properties of human mobility,” *Nature Physics*, vol. 6, no. 10, pp. 818–823, 2010.
- [40] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, “A density-based algorithm for discovering clusters in large spatial databases with noise,” in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, Menlo Park, CA, USA, August 1996.
- [41] P. J. Rousseeuw, “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis,” *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.
- [42] M. Zhang, S. He, and P. Zhao, “Revisiting inequalities in the commuting burden: institutional constraints and job-housing relationships in Beijing,” *Journal of Transport Geography*, vol. 71, pp. 58–71, 2018.

- [43] P. Qin and L. Wang, "Job opportunities, institutions, and the jobs-housing spatial relationship: case study of Beijing," *Transport Policy*, vol. 81, pp. 331–339, 2019.
- [44] P. Zhao, D. Liu, Z. Yu, and H. Hu, "Long commutes and transport inequity in China's growing megacity: new evidence from Beijing using mobile phone data," *Travel Behaviour and Society*, vol. 20, pp. 248–263, 2020.
- [45] H. Ballis and L. Dimitriou, "Revealing personal activities schedules from synthesizing multi-period origin-destination matrices," *Transportation Research Part B: Methodological*, vol. 139, pp. 224–258, 2020.
- [46] E. Chen, Z. Ye, C. Wang, and W. Zhang, "Discovering the spatio-temporal impacts of built environment on metro ridership using smart card data," *Cities*, vol. 95, Article ID 102359, 2019.
- [47] N. C. Iraganaboina, T. Bhowmik, S. Yasmin, N. Eluru, and M. A. Abdel-Aty, "Evaluating the influence of information provision (when and how) on route choice preferences of road users in Greater Orlando: application of a regret minimization approach," *Transportation Research Part C: Emerging Technologies*, vol. 122, Article ID 102923, 2021.
- [48] A. Ceder and Y. Jiang, "Route guidance ranking procedures with human perception consideration for personalized public transport service," *Transportation Research Part C: Emerging Technologies*, vol. 118, Article ID 102667, 2020.
- [49] H. Yin, B. Cui, Z. Huang, W. Wang, X. Wu, and X. Zhou, "Joint modeling of users' interests and mobility patterns for point-of-interest recommendation," in *Proceedings of the 2015 ACM Multimedia Conference*, Portland, OR, USA, March 2015.
- [50] M. J. Alonso-González, S. Hoogendoorn-Lanser, N. van Oort, O. Cats, and S. Hoogendoorn, "Drivers and barriers in adopting Mobility as a Service (MaaS) – a latent class cluster analysis of attitudes," *Transportation Research Part A*, vol. 132, pp. 378–401, 2020.
- [51] I. Lopez-Carreiro, A. Monzon, E. Lopez, and M. E. Lopez-Lambas, "Urban mobility in the digital era: an exploration of travellers' expectations of MaaS mobile-technologies," *Technology in Society*, vol. 63, Article ID 101392, 2020.
- [52] A. Vij, S. Ryan, S. Sampson, and S. Harris, "Consumer preferences for Mobility-as-a-Service (MaaS) in Australia," *Transportation Research Part C: Emerging Technologies*, vol. 117, Article ID 102699, 2020.