Research Article

Fusing Mobile Phone and Travel Survey Data to Model Urban Activity Dynamics

Chao Yang,¹ Yuliang Zhang ,¹ Xianyuan Zhan,² Satish V. Ukkusuri,³ and Yifan Chen⁴

¹The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai, China

²JD Intelligent City Research, Beijing, China

³Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA

⁴Institute of Transport Studies, Department of Civil Engineering, Monash University, Melbourne, Australia

Correspondence should be addressed to Yuliang Zhang; 3zyl@tongji.edu.cn

Received 28 November 2019; Revised 17 April 2020; Accepted 24 April 2020; Published 28 August 2020

Academic Editor: Giulio E. Cantarella

Copyright © 2020 Chao Yang 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.

A key issue to understand urban system is to characterize the activity dynamics in a city—when, where, what, and how activities happen in a city. To better understand the urban activity dynamics, city-wide and multiday activity participation sequence data, namely, activity chain as well as suitable spatiotemporal models, are needed. The commonly used household travel survey data in activity analysis suffers from limited sample size and temporal coverage. The emergence of large-scale spatiotemporal data in urban areas, such as mobile phone data, provides a new opportunity to infer urban activities and the underlying dynamics. However, the challenge is the absence of labeled activity information in mobile phone data. Consequently, how to fuse the useful information in household survey data and mobile phone data to build city-wide, multiday, and all-time activity chains becomes an important research question. Moreover, the multidimension structure of the activity data (e.g., location, start time, duration, type) makes the extraction of spatiotemporal activity patterns another difficult problem. In this study, the authors first introduce an activity chain inference model based on tensor decomposition to infer the missing activity labels in large-scale and multiday activity data, and then develop a spatiotemporal event clustering model based on DBSCAN, called STE-DBSCAN, to identify the spatiotemporal activity patterns. The proposed approaches achieved good accuracy and produced patterns with a high level of interpretability.

1. Introduction

Cities, which are complex sociotechnical systems, are home to more than fifty percent of the world population [1]. Highly complex urban dynamics and inherent activity participation pattern, human mobility, and multitude of trips across various modes naturally emerge from the daily life of millions of people in urban areas. A fundamental research question in urban studies is to characterize the activity dynamics in a city—when, where, what, and how activities in a city happen. Activity participation is a key driving force for the human mobility patterns in cities. Indeed, people travel because of the desire or the need to participate in a particular activity, such as working, shopping, and entertainment. Understanding when and where

these activities occur at a finer resolution is a key step in understanding the city dynamics. Previous studies focus on investigating human mobility pattern using data from mobile phone [2], social media [3, 4], smart card [5], and GPS devices [6]. Due to the lack of information about the purposes behind these movements, these studies often ignore the interplay between selection of destinations for different activity purposes and mobility dynamics [7]. To better understand the dynamics and functional characteristics of cities as related to activity participation, one should explore the underlying mobility patterns by mining large volumes of data and model the city-scale activity participation. The understanding of the city-scale activity dynamics has a wide range of applications, such as urban and transportation planning (activity-based modeling [8]),

information diffusion [9], activity and location recommendation [10], and targeted advertising [11].

To develop predictable activity models at the city level, three key challenges should be addressed: coverage, representativeness, and scale. Traditional approaches for urban activity modeling rely on household travel survey, which is known to be costly and have limited coverage. Some researchers also suggest that travel survey has recall bias as the respondents tend to forget details of activities other than major activities such as home and work activities [12, 13]. Recent studies focus more on inferring city-scale activity participation using large-scale check-in data from location-based services [14], GPS data [15], and smart card [16]. However, these data sources also suffer from a series of representativeness issues. For example, frequent check-in users are only a small group in the population. Moreover, it is found that the check-in activities are more likely to be discretionary activities (eating, shopping, entertainment, etc.) rather than more frequent but regular activities (e.g., home and work) [7]. Similarly, GPS and smart card only capture specific user groups (e.g., taxi users or metro system), which are insufficient to capture city-wide activity participation pattern. Compared with the previous urban data sources, mobile phone data are massive in scale and have large spatiotemporal coverage and population representativeness, which is ideal for city-wide inference possible. However, a main challenge is that mobile phone data only contains spatiotemporal information of the users, without related activity information. This requires us to develop methods to fuse information from multiple data sources for city-wide activity inference. This approach allows us to take the advantage of both the detailed activity information from datasets with limitations (i.e., travel survey) and the good coverage, representativeness, and scale offered in the mobile phone data.

Another challenge is how to model activity dynamics (spatiotemporal activity pattern) at the city level. Activity event (participation) data have three dimensions, which are location (including spatial information, such as longitude and latitude), time (including activity start time and duration), and characteristic (including activity type). The multidimension structure of the activity data makes the spatiotemporal activity pattern recognition a difficult problem. It requires building a model to combine multidimension information of the activity event data and calculate the proximity between activity events.

In this study, we focus on two tasks to analyze urban activity dynamics: (1) inferring the unobserved activity information from the mobile phone data; (2) based on inferred activity information, extracting the spatiotemporal patterns of the activities. Two models are proposed to tackle these tasks, namely, activity chain inference model and spatiotemporal activity pattern recognition model.

1.1. Activity Chain Inference Model. We develop a novel approach based on tensor-based collaborative filtering framework to infer large-scale individual-based activity chains by fusing mobile phone data and travel survey data.

The proposed approach consists of two submodules: (1) a rule-based model to identify home and work activities, as home and work are highly regular activities and can be accurately inferred from multiday individual spatiotemporal trajectories; (2) a tensor-based collaborative filtering (TCF) model to fuse information from travel survey data and mobile phone data and infer noncommuting activities, which makes it possible to fully utilize the hidden information in the mobile phone data.

1.2. Spatiotemporal Activity Pattern Recognition Model. We develop a spatiotemporal event clustering algorithm based on DBSCAN (STE-DBSCAN) to solve the activity clustering problem. By defining the distance between activity events (*activity event distance*), multidimension information of the activity data is carefully taken into consideration. The proposed STE-DBSCAN algorithm is fast and can be easily integrated into other big data mining tasks.

Our contributions can be summarized as follows:

- (i) A new method for understanding urban activity dynamics using big data: with good coverage, representativeness, and scale of mobile phone data, we can find more plentiful and complicated patterns of urban activity.
- (ii) A new data fusion approach for city-wide activity information gathering, which can utilize the accurate ground truth activity information from travel survey together with the ubiquitous spatiotemporal information in mobile phone data.
- (iii) A flexible solution for spatiotemporal activity pattern recognition, which can sensitively balance the effects of time and space.

The following sections review the related works, introduce the data and the urban context feature extraction, describe the methodologies of the paper, present the experiment results, and conclude this work.

2. Related Works

2.1. Activity Inference. There are many works in the literature using data fusion on activity inference. Chen et al. [17] used daytime activity center and nighttime activity center to identify the home and work location and used distance constraint with temporal constraint to get the activity chain of mobile phone users. Phithakkitnukoon et al. [18] used “activity-aware map” to estimate the most likely activity related to a specific location and then developed a model to describe the activity type of users. This is an example of combining mobile phone data and point-of-interest (POI) data. Shen and Stopher [19] used the National Household Travel Survey (NHTS) data in US to build a trip purpose imputation model from GPS devices data. In this case, the trip purposes were inferred based on rules obtained from the NHTS data rather than directly observed by the GPS devices data. Allahviranloo and Recker [15] also used the travel survey data and GPS devices data to mine activity patterns. Kusakabe and Asakura [16] combined travel survey data and

transit smart card data to infer trip purposes for metro users. Moreover, Langlois et al. [20] established a linkage between multiweek activity sequences of public transport users and sociodemographic attributes using data fusion. Alsger et al. [21], Chen et al. [22], and Soares et al. [23] inferred trip purpose for taxi GPS traces, public transport, and GPS device traces, respectively, using machine learning method to extract features and identify patterns. Wang et al. [24] also used smartphone GPS data to infer trip purpose and travel mode, but they used smaller sample size (16 volunteers) and high-density data (GPS data), which is hard to be applied in the city level. Almost all these works focus on finding rules which can be used in one of the data sources or building links between two or more kinds of datasets. As discussed previously, the datasets considered in these works all contain certain drawbacks, which limits their applications to approximate the actual individual activity pattern. For example, some models need high-precision data or plenty of attributes including demographic information, which are difficult to acquire. We want to find a unified model which can take full use of labeled activity data (i.e., travel survey data) and other spatiotemporal data (i.e., mobile phone data) with less limitation.

2.2. Collaborative Filtering Method. Collaborative filtering is a data mining framework which is widely used to infer useful information from multisource partially observed data. Most of the work on collaborative filtering focused on the traditional problems involving only label and data features [25, 26], but for data with complex structures, e.g., space-time-value data, these methods are no longer applicable. The recent introduction of tensor-based collaborative filtering models has provided a new solution for complex structured data. Tensor decomposition (TD) is a powerful tool for the analysis of multiway data (the dimension of data is at least three) in neuroscience, signal processing, computer vision, data mining, etc. [27]. Rendle et al. [28] used three-dimensional tensor to recommend tag (a list of words used to describe an item) for users. Zheng et al. [10] built a user-location-activity tensor to recommend locations and activities, but they did not consider the temporal factor and the spatiotemporal correlation. More and more researchers use tensor decomposition in transportation domain recently. Wang et al. [29] used sparse GPS trajectories to estimate path travel times using partial trajectory data from multiple data sources. Ran et al. [30] and Chen et al. [31] used TD to estimate traffic volume and speed, respectively. As tensor decomposition enables flexible handling of multiway, sparse, and partial datasets, it serves as an ideal solution to our activity chain inference problem.

2.3. Spatiotemporal Pattern Recognition. Many studies also have been done on this topic. Kumar et al. [32] used statistical techniques to capture the traffic patterns. Ma et al. [33] used k -means combined with principle component analysis and entropy index to extract driving patterns of taxis. In addition, clustering methods are widely used for spatiotemporal pattern recognition, such as spatiotemporal

(ST) events clustering [34, 35], georeferenced time series clustering [36], and ST trajectory clustering [37]. However, few methods focus on multiple (more than 2) types of activity event clustering with consideration of activity participation information like location, start time, and duration. Activity type, start time, duration, and location are basic information used in activity-based modeling approach (transportation planning model), epidemic spreading, activity and location recommendation, target advertising, etc.

3. Data and Urban Context Feature Extraction

Two datasets from Shanghai are used in this study, which are travel survey data and mobile phone data (detailed statistics are shown in Table 1). The travel survey data contains 139,195 residents' trip details including trip departure and arrival time, location of origin and destination, trip mode and purpose (activity category), and demographic attributes (age, gender, income, etc.). Activity categories in travel survey data include work, school, shopping, entertainment, business, pick-up, and home. Work and school activities are merged as work activities, because they share similar spatiotemporal activity participation patterns; i.e., both have events at fixed time and places with relative fixed durations. Moreover, as pick-up activities are not typical activities that happen at fixed locations, these activities are not considered in this study. Mobile phone data covers mobile phone signaling details including the event (i.e., call, SMS, location change, power off) time and the relevant base station ID and user ID. After matching the base stations to their locations, mobile phone user's location at the time of a given event can be approximated. The spatial resolution of mobile phone signaling data depends on the service radius of each cellular tower, which varies with different telecom companies in different areas; e.g., in China it is about 100 to 500 meters in urban area and 400 to 10,000 meters in the suburbs [38].

One mobile phone user has 432 records per day.

In addition, we extracted a set of urban context features for activity inference, including time, geolocation (latitude and longitude), point-of-interest (POI), and road network information. POI data includes the location information of special points in the map, such as school, shop, and gas station. The POI data is collected via API of Baidu Map in 2012. The POI data is used to model the urban function of a region, and the road network data is used to model the accessibility of a region. A virtual grid reference is constructed by dividing the map into square cells of size 500×500 meters (considering that the positioning accuracy of base station in the mobile phone data is about 200 meters in downtown). The POI and road network information within each cell are extracted. The POI dataset contains 486,615 POIs and are divided into 8 categories, which are (1) schools, (2) companies, offices, banks, and ATMs, (3) mails and shopping malls, (4) restaurants, (5) gas stations, vehicle service locations, parking areas, and transportation facilities, (6) residences, (7) entertainment and living services, and (8) hotels, as suggested by [7]. We count the number of POIs of each category in every cell and create 9 POI features as the relative proportions of POIs of each category in this cell, as

TABLE 1: Description of two datasets.

Data source	Data size	Period	Useful information
Travel survey data	60,032 households	One day during Sep. 18–25, 2009	Activity chains with time and location details
	139,195 residents	99,565 residents on weekday	
	266,043 trips	39,630 residents on weekend	
Mobile phone data	3,094,297 users	May 01, 2015–May 31, 2015	Spatiotemporal trajectories
	41,468,071,691 records		

well as the relative abundance of POIs (number of POIs in the cell divided by the maximum number of POIs among all cells). The roads in Shanghai road network are divided into 3 categories based on road types, which are freeway, major road, and local road. Again, we create 4 features for road network information, including the relative length proportion of each road type in the cell, as well as the total road lengths of the cell divided by the maximum cell road lengths. POI distribution and hierarchical road distribution are shown in Figure 1.

4. Methodologies

The activity chain inference model contains three steps: (1) extracting the spatiotemporal trajectories to form a travel-stay chain for each mobile phone user; (2) identifying the home and work activities; (3) inferring noncommuting activities using a tensor-based collaborative filtering approach that fuses information from both travel survey and mobile phone data. Once the activity information is inferred, we build an activity spatiotemporal pattern recognition model to extract the spatiotemporal activity patterns at city level through clustering.

4.1. Activity Chain Inference

4.1.1. Trajectory Generation and Activity Point Recognition. Due to the noise in signal, base station positioning results may jump at several nearby base stations. This situation is called ping-pong phenomenon or ping-pong handover [39]. Most existing solutions that address this issue focus on increasing hysteresis threshold (used as a spatial constraint to merge close base stations) to reduce the positioning error [39]. However, simply increasing the hysteresis threshold may lead to dropping useful information from the spatiotemporal trajectories. To provide a better solution, we propose a spatial-temporal constrained smoothing method (STCS).

STCS (shown in Figure 2) uses two thresholds to filter the trajectory. The first parameter is a spatial stay threshold δ to constrain fluctuation in spatial dimension, and the second parameter is the temporal stay threshold τ used to limit the temporal dimension. When user's position is fluctuating within a circle with radius less than δ during period τ , user is regarded as staying in this circle, and when user's signal jumps to a remote location and goes back to the previous location during time τ , user is regarded as motionless. The values of these two parameters are determined based on domain knowledge, and we want the result of activity point recognition to fit the domain knowledge of urban plan or

transportation plan. The spatial stay threshold is selected as $\delta = 400$ meters considering that the lower bound of walking trip distance is 500 m and the distance we calculate in this case is the great circle distance. τ is assigned to 30 minutes based on the analysis of travel survey data in Shanghai, which is the 5% quantile at cumulative frequency of activity duration distribution (shown in Figure 3). Based on this observation, staying at a place for more than 30 minutes is recognized as an activity stop and the travel-stay chains of all mobile phone users can be generated.

4.1.2. Home/Work Activity Identification. As home and work are highly regular activities (in general cases, residents usually stay home at night and work during daytime) and mobile phone data is long-term observation data (one month in this paper), we use a rule-based model to identify the types of these regular activities. Special cases like working all night are relatively rare, and transportation planning studies are more interested in the daytime activities. Residents' home and work activity patterns in the travel survey data are illustrated in Figure 4. We create two time intervals that capture typical home/work stay time. Home stay time is set to be during night from 8pm to 7am (next day), and work stay time is used as the common work time from 9am to 5pm. We introduce two thresholds to filter home and work activities. For home activity, the home stay time threshold τ_H is defined as 260 minutes, which captures 95% of staying-home activities (Figure 4(a)). Similarly, work stay time threshold τ_W is set as 180 minutes based on the 5% quantile at cumulative frequency of work duration (Figure 4(b)). For a given user, if his/her night stay time (from 8pm to next day 7am) at a location is larger than τ_H and this situation repeats more than 12 days (60% of weekdays in May), the location can be identified as home location. Work location is identified in a similar way. Finally, 2,193,517 users' home locations and 1,251,746 users' work locations can be identified.

4.1.3. Activity Chain Inference Using Tensor Decomposition-Based Collaborative Filtering. As mentioned before, our datasets are multidimensional (including spatial, temporal, and characteristic information), sparse (there is a lot of missing information needed to be inferred in mobile phone data), and partial (travel survey data has limited samples and recall bias), and we need a unified model which can take full use of the multisource data. Multidimensional data is often referred to as a tensor, and tensor decomposition is a standard technique to capture the multidimensional structural dependencies. By decomposing the partially observed

FIGURE 1: POI distribution (a) and hierarchical road distribution (b) (red: freeway; blue: major road; green: local road).

FIGURE 2: Illustration of STCS method handling small-scale ping-pong phenomenon.

FIGURE 3: Cumulative frequency of activity duration.

tensor, the missing data in the tensor can be replaced using the product of the decomposed results. And because of the good performance on handling multiway, sparse, and partial data [29–31], tensor decomposition is an ideal solution to our activity chain inference problem.

The user-activity-context tensor in this study is constructed using both travel survey data and mobile phone data (illustrated in Figure 5). The upper part is filled with filtered travel survey data, and the bottom part is filled with mobile phone data in which all activity labels are unobserved except home and work activities. In the upper part, the non-commuting activity labels are randomly erased (under different missing rates) and used as the test set for evaluation. The vertical dimension of the tensor represents different days of different users. The horizontal dimension represents different time slots per day. We divide the day into 30 min time intervals, resulting in an activity chain with 48 elements. Due to the low number of activities during 2am–6am and the need to reduce the dimension of the tensor, activities during this period are not considered and the dimension of this mode is reduced to 40. The third mode contains activity and spatial-temporal urban context features, which include binary activity categories (6 elements, i.e., home, work, shopping, entertainment, business, and other), POI features (9 elements), and road network features (4 elements).

Tucker decomposition and CANDECOMP/PARAFAC (CP) are the two most popular methods to decompose a tensor. The Tucker decomposition of a 3D tensor is illustrated in Figure 6. A 3D tensor $\mathcal{X}$ is decomposed as the product of a core tensor G and a set of low rank matrices

FIGURE 4: Cumulative frequency of home (a) and work duration (b) on weekday.

FIGURE 5: Construction of user-activity-context tensor.

FIGURE 6: Illumination of three-way Tucker decomposition.

(A, B, C) by each mode. The CP decomposition can be considered as a special case of Tucker decomposition when the core tensor is diagonal and each mode has the same length as others. More theoretical background on tensor algebra and Tucker decomposition can be found in Kolda and Bader [40]. Tucker decomposition typically produces solution superior to that of CP, as it can capture more correlation in each mode and does not need to determine the rank of the tensor needed in CP decomposition (an NP-hard problem). Consequently, we use the Tucker decomposition in this study.

The Tucker decomposition can be viewed as a generalized form of principal component analysis (PCA) or

matrix factorization (MF). The Tucker decomposition of a 3D tensor can be illustrated as follows:

$$\mathcal{X} \approx \mathcal{G} \times_A A \times_B B \times_C C, \quad (1)$$

where $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$ is a 3D tensor, three matrices $A \in \mathbb{R}^{I \times P}$, $B \in \mathbb{R}^{J \times Q}$, $C \in \mathbb{R}^{K \times R}$ which are usually called factor matrices. $\mathcal{G} \in \mathbb{R}^{P \times Q \times R}$ is called the core tensor which captures correlations between factor matrices. The operation " $\times_A$ " refers to tensor-matrix multiplication on the mode A . Each element of $\mathcal{X}$ in Tucker decomposition is computed by

$$x_{ijk} \approx \sum_{p=1}^P \sum_{q=1}^Q \sum_{r=1}^R g_{pqr} a_{ip} b_{jq} c_{kr}, \quad (2)$$

$$\text{for } i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, K,$$

or equivalently,

$$x_{ijk} \approx \mathcal{G} \times_A A_{i*} \times_B B_{j*} \times_C C_{k*}. \quad (3)$$

The tensor decomposition problem can be formulated as a least square problem ($\|\cdot\|_F$ represents Frobenius norm), which is minimizing the total square error between the original tensor $\mathcal{X}$ and a decomposed tensor $\mathcal{G} \times_A A \times_B B \times_C C$:

$$\min \|\mathcal{X} - \mathcal{G} \times_A A \times_B B \times_C C\|_F^2. \quad (4)$$

4.1.4. Solution Approach. There are several existing algorithms for solving the tensor decomposition problem [40]. We use an efficient algorithm based on singular value decomposition (SVD) [41] in this study. The outline of the algorithm is presented as Algorithm 1.

Here, matrix B can be initialized by the R^B first left singular vectors from an SVD of the $J \times IK$ matrix $\mathcal{X}^{(2)}$, and C is obtained as the R^C first left singular vectors from an SVD of $(B^T \mathcal{X}^{(2)})^{(3)}$. And a three-way ($I \times J \times K$) array $\mathcal{X}$ can be unfolded as $\mathcal{X}^{(1)}$ ($I \times JK$ matrix), $\mathcal{X}^{(2)}$ ($J \times IK$ matrix), and $\mathcal{X}^{(3)}$ ($K \times IJ$ matrix). The calculation of A, B, C uses similar way of update. Using the first mode as an example, the SVD-based algorithm is updated as follows:

Step 0: Initialize B and C
 Step 1: From B , C , and $\mathcal{X}^{(1)}$, calculate $M^{(1)}$, and then update A
 Step 2: From C , A , and $\mathcal{X}^{(2)}$, calculate $M^{(2)}$, and then update B
 Step 3: From A , B , and $\mathcal{X}^{(3)}$, calculate $M^{(3)}$, and then update C
 Step 4: If convergence occurs, calculate the core G , else, go to Step 1

ALGORITHM 1: Tensor decomposition algorithm.

$$\begin{aligned} M^{(1)} &= \mathcal{X}^{(1)} (C \otimes B), \\ [A, S, V] &= \text{SVD}(M^{(1)}). \end{aligned} \quad (5)$$

The notation $\otimes$ denotes the Kronecker product and $\text{SVD}(\cdot)$ represents the SVD decomposition of a matrix. After convergent results of A , B , C are obtained, the unfolded core tensor $G^{(1)}$ (a $P \times QR$ matrix) can be computed as $G^{(1)} = A^T \mathcal{X}^{(1)} (C \otimes B)$. More theoretical background can be found in the study by Andersson and Bro [38].

The above algorithm is applicable for fully filled tensor, but our case involves a lot of missing values. To handle the missing data, we fit the decomposed tensor only for the nonmissing data and use a new loss function:

$$\min_{\text{Non-missing}} \mathcal{X} - \mathcal{G} \times_A A \times_B B \times_C C^2. \quad (6)$$

Many studies solve the above least square problem using the stochastic gradient descent (SGD) [42]. However, SGD has a very slow convergence rate and is computationally very expensive. This is because SGD needs to compute gradient and iterate over every element in the tensor. To improve computational efficiency, we use an alternative EM-based algorithm to solve the problem. The key idea is to iteratively estimate the missing values using current decomposed tensors until the missing values converge. The advantages of the EM-based algorithm are that, at every iteration, we can utilize the highly efficient SVD-based approach to decompose the tensor, and it is also computationally more stable compared with SGD. Kiers [43] has shown that directly solving the least square problem defined in (7) and using EM algorithm produce identical solutions, which justifies the effectiveness of the EM-based algorithm. For brevity, the SVD-based tensor decomposition method described previously is referred to as

$$[\mathcal{G}; A, B, C] = \text{TensorDecompose}(\mathcal{X}). \quad (7)$$

The EM-based solution algorithm can be found in Algorithm 2.

4.2. Spatiotemporal Activity Pattern Recognition. Key issues of spatiotemporal activity pattern recognition are how to measure the proximity between activity events (namely, *activity event distance*) and how to extract the similar patterns. DBSCAN (Density-Based Spatial Clustering of Applications with Noise) [44] is a clustering algorithm relying on a density-based notion of clusters and is widely used in spatiotemporal clustering [45,46]. The DBSCAN has several

attractive features: (1) it is effective in discovering clusters with arbitrary shape; (2) it does not require the predetermination of the number of clusters; (3) it can capture the main patterns and ignore the noise points automatically; (4) it can be run on a large database using suitable techniques [47]. These advantages make DBSCAN an ideal base framework for developing customized spatiotemporal clustering algorithms.

We introduce *activity event distance* to measure the similarity between activities, and further develop a spatiotemporal event clustering algorithm based on DBSCAN (STE-DBSCAN) to integrate the multidimension information of the activity data and solve our spatiotemporal activity pattern extraction problem. *DBSCAN clustering background* is presented in the appendix, and the *activity event distance* (D_a) is defined as follows:

$$\begin{aligned} D_a \left(\begin{matrix} a^1(a_s^1, a_d^1, a_l^1, a_t^1) \\ a^2(a_s^2, a_d^2, a_l^2, a_t^2) \end{matrix} \right) &= \left[(w_1 \times |a_s^1 - a_s^2|)^2 + (w_2 \times |a_d^1 - a_d^2|)^2 \right. \\ &\quad \left. + (w_3 \times D_g(a_l^1, a_l^2))^2 \right. \\ &\quad \left. + (w_4 \times D_{\text{type}}(a_t^1, a_t^2))^2 \right]^{1/2}, \end{aligned} \quad (8)$$

$$w_1 = \frac{1}{M_s} \times \delta_1, \quad (9)$$

$$w_2 = \frac{1}{M_d} \times \delta_2, \quad (10)$$

$$w_3 = \frac{1}{M_l} \times \delta_3, \quad (11)$$

$$w_4 = 1 \times \delta_4, \quad (12)$$

$$D_{\text{type}}(a_t^1, a_t^2) = \begin{cases} 1, & a_t^1 \neq a_t^2, \\ 0, & a_t^1 = a_t^2, \end{cases} \quad (13)$$

where *activity event* is denoted as $a(a_s, a_d, a_l, a_t)$; a_s, a_d, a_l , and a_t represent the start time, duration, location, and type of the activity, respectively; D_g represents the great circle distance which is the shortest distance between two points on a sphere computed using haversine formula; D_{type} represents the distance between activity types; M_s is the maximum start time difference in the datasets; M_d is the maximum duration difference in the datasets; and M_l is the

Step 0: Let $Y^{(1)} = \mathcal{X}$, fill the missing values in $Y^{(1)}$ with random values. Set $i = 1$.
 Step 1: Decompose tensor $Y^{(i)}$:
 $[\mathcal{G}^{(i)}; A^{(i)}, B^{(i)}, C^{(i)}] = \text{TensorDecompose}(Y^{(i)})$
 And
 $\tilde{\mathcal{X}}^{(i)} = \mathcal{G}^{(i)} \times_A A^{(i)} \times_B B^{(i)} \times_C C^{(i)}$
 Step 2: Let $Y^{(i+1)} = \mathcal{X}$, fill the missing values in $Y^{(i+1)}$ with values in $\tilde{\mathcal{X}}^{(i)}$.
 Step 3: Convergence check:
 $\text{SSE} = \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} I_{ijk} (y_{ijk}^{(i+1)} - y_{ijk}^{(i)})^2$,
 $I_{ijk} = 1$ if y_{ijk} is a missing value in $\mathcal{X}$;
 0 otherwise
 If $\text{SSE} < \epsilon$, terminate the algorithm, $\mathcal{X}^* = \mathcal{G}^{(i)} \times_A A^{(i)} \times_B B^{(i)} \times_C C^{(i)}$. Otherwise $i = i + 1$, go to Step 1.

ALGORITHM 2: EM-based solution algorithm.

maximum location distance in the datasets. In (9)–(12), δ_1 , δ_2 , δ_3 , and δ_4 are used to adjust the effect of spatial, temporal, and categorized information of the activity.

Finally, STE-DBSCAN (shown in Algorithm 3) can be implemented with the model inputs denoted as STE-DBSCAN($Eps, MinFts, Distance\ Matrix$).

5. Experimental Results and Discussion

5.1. Activity Chain Inference. We have conducted a series of experiments to test the proposed model named tensor-based collaborative filtering (TCF) model. In the upper part of the user-activity-context tensor, not all the survey data can be used and we make strict rules: (1) the activity chain of user has at least two activity categories; (2) the activity chain covers more than 12 hours; (3) Each activity has location information (longitude and latitude). Finally, we get 1634 activity chains from survey data. The activity inference model is implemented using the N-way Toolbox in MATLAB [48]. To evaluate the performance of the model, mean absolute percentage error (MAPE) and root mean squared error (RMSE) are used, which are computed as

$$x'_{ijk} = \mathcal{G} \times_A A_{i*} \times_B B_{j*} \times_C C_{k*},$$

$$\text{MAPE} = \frac{\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \text{abs}(x_{ijk} - x'_{ijk})}{\sum_{i \in I} \sum_{j \in J} \sum_{k \in K} x_{ijk}}, \quad x_{ijk} \text{ is non-missing},$$

$$\text{RMSE} = \sqrt{\frac{1}{N_{\text{train}}} \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} (x_{ijk} - x'_{ijk})^2},$$

x_{ijk} is non-missing entries in the tensor,

(14)

where x_{ijk} is an element in the user-activity-context tensor $\mathcal{X}$, x'_{ijk} is estimated value of x_{ijk} , and N_{train} is the number of nonmissing entries. We use the previously removed activity labels (excluding home and work activities, as they were identified before) in the travel survey data as the ground truth to evaluate the accuracy of the inference results, computed as

$$\text{score} = \begin{cases} 1, & \text{label}_{\text{Predic}} = \text{label}_{\text{True}}, \\ 0, & \text{label}_{\text{Predic}} \neq \text{label}_{\text{True}}, \end{cases} \quad (15)$$

$$\text{Total accuracy} = \frac{1}{N_{D_{\text{test}}}} \sum_{i \in D_{\text{test}}} \text{score}_i,$$

where label is the activity category and $\text{label}_{\text{Predic}}$ is obtained by selecting the category with the highest value of the activity category features in the final filled tensor. Besides, we also use F_1 score to measure the test accuracy, which is computed by harmonic mean of precision and recall as

$$F_1 \text{ score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}, \quad (16)$$

where precision is the number of correct positive results divided by the number of all positive results, and recall is the number of correct positive results divided by the number of positive results that should have been returned.

Different groups of input features have been tested before being selected as the final feature set, and different values of core tensor dimensions (P , Q , and R) have been tested to achieve the best model performance. After conducting a series of tests, we selected $P=24$, $Q=10$, and $R=14$ as the final core tensor size (the original tensor size is $1634 \times 40 \times 19$).

As ground truth activity labels of the mobile phone data are unknown to us, the activity labels in the travel survey data are randomly erased and used as the test set for validation. In our model, mobile phone data and travel survey data are fused together, and the label prediction problem is transferred to missing information imputation problem. We designed two test scenarios to evaluate the proposed framework.

Scenario 1 only uses fine grained travel survey data to construct the tensor. This scenario is designed to test the explanatory and predictive power of the model under different missing rates (different missing rates here mean different missing ratios of the known noncommuting activity labels in the travel survey data). We also compare our results with other machine learning methods which are used to infer activity types in other research works.

Step 0: $\forall m, n$ in activity event set D ,
 $\text{dist}(m, n) = D_a \begin{pmatrix} a_s^m(a_s^m, a_d^m, a_l^m, a_t^m) \\ a_s^n(a_s^n, a_d^n, a_l^n, a_t^n) \end{pmatrix}$
 Get *Distance Matrix*
 Step 1: Set values of Eps and MinPts.
 Step 2: Find and delete all noise.
 Step 3: Find all core point set C and border points B .
 Step 4: $\forall p, q$ in C ,
 set $\text{cluster_number}_p = \text{cluster_number}_q$, if $\text{dist}(p, q) \leq \text{Eps}$.
 Step 5: $\forall p$ in B ,
 find the core point q which is density-reachable from p , and set $\text{cluster_number}_p = \text{cluster_number}_q$.

ALGORITHM 3: STE-DBSCAN algorithm.

Scenario 2 combines different sizes of mobile phone data as well as travel survey data in the tensor. The sizes are tested as the multiples of the size of travel survey data. This scenario is designed to test the robustness of the model with different amount of unlabeled activity data.

Results of scenario 1 are presented in Figure 7 and Tables 2 and 3 along with the benchmark methods. Reasonable results are obtained using the limited urban context features and the partially observed data. Figure 7 shows the accuracy, MAPE, and RMSE changing under different limited situation of information. The model can stably fit about 80% of the data including activity, land use, and road network information even in the extremely missing data conditions (see the red line in Figure 7). As the input data decreases, the MAPE and RMSE drop a little bit because the data needing to be fitted becomes less. The accuracy gets worse with the reduction of the information. However, the model still can reach about 48% accuracy only using 10% of the activity data, which proves the explanatory and predictive power of the model. We also present the confusion matrix of two different missing rates in Table 3. Under the extremely missing condition, the model performance on business activity is very low because business activity does not happen at a fixed time and often happens at central business district (CBD). It is hard to distinguish business activity with shopping and entertainment sometimes because they may happen with similar urban context features. In addition, this makes it difficult for the model to capture the relationship between users, time, business activity, and urban features. However, the good news is that our model tends to regard business activity as other activity (see the bottom part of Table 3), which means that our model still has good performance if we only consider shopping, entertainment, and other activities when we regard 'business activity' as 'other activity.'

We also compare our results with other machine learning methods (see Table 3 and Figure 8). It should be noted that none of the four benchmark methods can handle missing values, so we transfer the problem to classification problem. The proposed approach outperforms all the four benchmark methods using the same information. The related work we discussed before describes how to use those machine learning methods in detail. There is no solid evidence showing that the rules deriving from survey data can perfectly and directly be

FIGURE 7: Performance of the model ($P=24$, $Q=10$, $R=14$).

used in another dataset, and that is the reason we believe the framework we introduced is a better and reasonable activity inference model.

The test results of scenario 2 are presented in Figure 9. With the increase proportion of mobile phone data, the MAPE and RMSE become steady after a slight change. In addition, our model still can stably fit about 65% of the data. Larger amount of heterogeneous data makes tensor decomposition harder to reconstruct the original tensor. The convergent value of MAPE also suggests that human activity follows some underlying patterns which can be handled by the decomposed factors A, B, C and core tensor $\mathcal{Z}$, and the complexity of such patterns does not increase with the sample size. The accuracy quickly drops when the size of input phone data is twice as large as the survey data. Then, the accuracy keeps stable to about 48% even with the increase of the phone data size to 5 times the survey data. The result is coherent with what we got under extremely missing condition in scenario 1.

5.2. Spatiotemporal Activity Pattern. After inferring all activity labels over the city, we can mine the spatiotemporal activity patterns using STE-DBSCAN. We randomly select

TABLE 2: Confusion matrices of TCF results.

Actual\predicted	Shopping	Entertainment	Business	Other
Missing rate = 50%				
Shopping	1216 (0.771)	57 (0.036)	55 (0.035)	249 (0.158)
Entertainment	179 (0.110)	1203 (0.741)	68 (0.042)	173 (0.107)
Business	50 (0.094)	24 (0.045)	397 (0.749)	59 (0.111)
Other	113 (0.079)	50 (0.035)	12 (0.008)	1264 (0.878)
Overall accuracy	0.814			
Missing rate = 80%				
Shopping	1375 (0.545)	409 (0.162)	28 (0.011)	711 (0.282)
Entertainment	467 (0.180)	1428 (0.550)	1 (0.0004)	700 (0.270)
Business	77 (0.091)	81 (0.096)	23 (0.027)	667 (0.786)
Other	339 (0.099)	330 (0.096)	14 (0.004)	2739 (0.800)
Overall accuracy	0.593			
Overall accuracy after regarding business as other	0.665			

The values in parentheses are proportions.

TABLE 3: Benchmark test result.

Model	Overall accuracy	F ₁ -score			
		Shopping	Entertainment	Business	Other
TCF (missing rate: 50%)	0.814	0.776	0.814	0.748	0.857
Naïve Bayes	0.405	0.302	0.142	0.206	0.552
SVM	0.425	0.388	0.327	0.091	0.522
Decision tree (DT)	0.731	0.691	0.729	0.741	0.762
Random forest (RF)	0.773	0.749	0.765	0.790	0.795

FIGURE 8: Confusion matrices of different models. (a) TCF (50% missing rate). (b) Gaussian Naïve Bayes. (c) SVM. (d) Random forest.

FIGURE 9: Robust performance of the model ($P=24$, $Q=10$, $R=14$).

TABLE 4: Parameter selection result.

Number of clusters	Eps				
	0.05	0.06	0.07	0.08	0.09
$\delta_1 = 1, \delta_2 = 1$	218	115	71	59	36
$\delta_1 = 2, \delta_2 = 2$	242	246	178	124	90
$\delta_1 = 3, \delta_2 = 3$	316	410	204	169	152
$\delta_1 = 4, \delta_2 = 4$	262	342	403	432	199
$\delta_1 = 5, \delta_2 = 5$	229	327	359	422	421

The MinPts is set to 10, and δ_3 and δ_4 are set to 1.

FIGURE 10: Cluster number distribution related to different activity types.

100,000 inferred activity records derived from mobile phone signaling data on May 5, 2015 (Tuesday), in Shanghai to verify the performance of our spatiotemporal activity pattern recognition algorithm. To adjust the clustering results on time and space (we want to clearly identify different activity start time and duration from the clustering results, but we also want to make the activities happening at close places easily fall into the same cluster), we make the model more sensitive to time and less sensitive to space by emphasizing the time factors δ_1 and δ_2 . We also try to make Eps less to capture the pattern from the micro perspective. Finally, under the principle of finding more patterns, Eps is set to 0.06, and δ_1 and δ_2 are set to 3 (see Table 4). In addition, MinPts is set to 10 to capture more common patterns. As a result, 410 clusters are obtained. The parameters can be adjusted in any way we want to meet special applications.

We plot the spatial distribution of different clusters on the map to illustrate the patterns discovered. Different colors are used to represent different activity patterns, and average time information of the patterns is included below the map (N: cluster number; S: average start time of activity; D: average duration of activity). Cluster number distribution corresponding to different activity types is shown in Figure 10. The patterns of shopping and entertainment activities are more diversified than the patterns of home and work (regular activities), which is consistent with our daily experience.

Considering spatiotemporal information and activity type, these 410 clusters depict the urban activity patterns in fine-granularity of time and space. The results also illustrate the complexity of the urban (activity) system. And urban dynamics of human activity can be modeled by the patterns

FIGURE 11: Sample activity pattern related to work. Size of the points represents the relative activity frequency. N : cluster number; S : start time; D : duration. (a) Entertainment in the early morning. (b) Entertainment in the afternoon. (c) Entertainment in the evening. (d) Long time entertainment.

with different space and time properties. In this section, because of the space limitations, we only pick partial patterns in the downtown of Shanghai to show the performance of our methods. Work is the most important activity in human life (some work-related patterns are shown in Figure 11). It is noted that the size of the point in the figure represents the relative frequency of the activity happening at one place. Work activities in the morning, afternoon, and evening are described in Figures 11(a)–11(c), respectively. Normal work activities begin from 8am to 9am. We can see that the work activities happening in the evening have less density and are in the center of the downtown. Figure 11(d) shows the

activity pattern of high-intensity work beginning in the morning (frequently happening in the center of the city). Activity patterns related to shopping are depicted in Figure 12. The results show a large-scale phenomenon which is shopping in the early morning, noon, and twilight (shown in Figures 12(a)–12(c)). Because the activity label of eating is not included in our activity data (the activity label of eating is not included in the household travel survey data) and it is difficult to distinguish the eating and shopping activities, the results represent the mixed activities of eating and shopping in the early morning, noon, and twilight. We can also find the less-density shopping activities happening in the late

FIGURE 12: Sample activity pattern related to shopping. Size of the points represents the relative activity frequency. N : cluster number; S : start time; D : duration. (a) Work in the morning. (b) Work in the afternoon. (c) Work in the evening. (d) Long time work in the morning.

night at the shopping mall of the city (shown in Figure 12(d)). For entertainment activity, we can find large-scale entertainment activities happening in the early morning (in Figure 13(a)), which probably represent the morning exercise of the residents. Because the selected data is on Tuesday, we can see that the entertainment activities have less density in the afternoon and higher density in the evening (in Figures 13(b) and 13(c)). We can also find long duration entertainment activities happening with less

density in Figure 13(d). Note that the selected patterns represent only a fraction of the entire activities in Shanghai. All these activity patterns are included in the urban dynamics of human activity and are helpful to understand and model urban human activity. The results prove the effectiveness of the STE-DBSCAN for spatiotemporal activity pattern recognition. In addition, the good interpretability of the patterns also points to the reasonable results of the activity chain inference model.

FIGURE 13: Sample activity pattern related to entertainment. Size of the points represents the relative activity frequency. N : cluster number; S : start time; D : duration. (a) Shopping in the early morning. (b) Shopping in the noon. (c) Shopping in the twilight. (d) Shopping in the late night.

6. Conclusions

In this paper, we propose two models to analyze urban activity dynamics. In the first model, we infer multiday and all-time activity chains using large-scale heterogeneous and sparse data sources from urban areas. The proposed model overcomes the limitations of existing studies and utilizes the accurate ground truth activity information from travel survey together with the ubiquitous spatiotemporal information in mobile phone data. The activity labels of individual mobile phone records can be automatically annotated, which can lead to important applications in urban transportation planning and activity/location recommendation systems as well as targeted advertising systems. This approach can also be used to infer complete

activity chains from not only mobile phone data and survey data (which we use here) but also any other type of massive partial and highly heterogeneous datasets. The dataset considered in this work is from Shanghai, China, but this method is transferable to other urban areas with similar data. The results also suggest that this approach is powerful and robust in handling missing values. The second model proposes the STE-DBSCAN algorithm, which can automatically capture detailed spatiotemporal activity patterns. The results can tell us when, where, and what activities happen in our city, and different spatiotemporal activity distributions are found at different times of the day. The discovered activity patterns also confirm the reasonability of the activity inference. Our work can be used to understand urban activity dynamics at a deeper level, which is important for urban

planning, disease control, location/activity recommendation services, etc.

Future research can explore parallelized solution approaches for the proposed models to more efficiently handle huge volume of data. Better model performance can be achieved if more informative features are introduced, such as detailed urban land use and high-precision survey data captured by GPS devices. The results will be more reasonable if we can use datasets at the same time period, as we made an assumption that the way people travel did not change too much from 2009 to 2015. In addition, it is interesting if we can model how the way people travel changes with the city development. The activity spatiotemporal patterns can be used as an input to future activity prediction problem, which has wide applications in location/activity recommendation, targeted advertising, etc.

Appendix

DBSCAN Clustering Background

Some basic definitions for DBSCAN are provided below [35]:

Definition 1 (Eps-neighborhood of a point).

The Eps-neighborhood of a point $p(N_{\text{Eps}}(p))$ is defined by $N_{\text{Eps}}(p) = \{q \in D \mid \text{dist}(p, q) \leq \text{Eps}\}$, where $\text{dist}(p, q)$ is the distance function between two points p and q , and Eps is the input parameter.

Definition 2 (core point) and *Definition 3* (directly density-reachable).

A point p is core point when its neighborhood of a given radius (Eps) contains at least MinPts of other points, where MinPts is also the input parameter. And these “other points” are said to be directly density-reachable from p .

Definition 4 (density-reachable).

A point q is density-reachable from p if there is a path $p_1, \dots, p_n$ with $p_1 = p$ and $p_n = q$, where each p_{i+1} is directly density-reachable from p_i , for $1 \leq i \leq n-1$, $p_i \in D$.

Definition 5 (density-connected).

Two points p and q are density-connected if there exists a point o such that both p and q are density-reachable from o .

Definition 6 (border points).

The border point is density-reachable from another core point but is not core point itself.

Definition 7 (noise).

The points which are not density-reachable from any other point are noise.

Data Availability

The mobile phone data and household travel survey data used to support the findings of this study have not been made available because we do not have the right to share

them with the public. However, we can provide part of encrypted data. The POI data and the road network data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was sponsored by National Natural Science Foundation of China (71171147).

References

- [1] Department of Economic, Social Affairs, and United Nations, “World Urbanization Prospects: The 2014 Revision, Highlights,” United Nations Department of Economic and Social Affairs, New York, NY, USA, 2014.
- [2] M. C. González, C. A. Hidalgo, and A.-L. Barabási, “Understanding individual human mobility patterns,” *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [3] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui, “Exploring millions of footprints in location sharing services,” in *Proceeding of the 5th International AAAI Conference on Weblogs and Social Media*, pp. 81–88, Barcelona, Spain, 2011.
- [4] A. Nikolaidou and P. Papaioannou, “Utilizing social media in transport planning and public transit quality: survey of literature,” *Journal of Transportation Engineering, Part A: Systems*, vol. 144, no. 4, Article ID 04018007, 2018.
- [5] S. Hasan, C. M. Schneider, and S. V. Ukkusuri, “Spatiotemporal patterns of urban human mobility,” *Journal of Statistical Physics*, vol. 151, no. 1-2, pp. 304–318, 2013.
- [6] X. Qian, X. Zhan, and S. V. Ukkusuri, “Characterizing urban dynamics using large scale taxicab data,” in *Engineering and Applied Sciences Optimization*, pp. 17–32, Springer International Publishing, Berlin, Germany, 2015.
- [7] S. Hasan, X. Zhan, and S. V. Ukkusuri, “Understanding urban human activity and mobility patterns using large-scale location-based data from online social media,” in *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing—UrbComp ’13*, vol. 6, 2013.
- [8] C. R. Bhat and F. S. Koppelman, “Activity-based modeling for travel demand,” in *Handbook of Transportation science*, R. W. Hall, Ed., pp. 35–61, Springer, Boston, MA, USA, 1999.
- [9] J. L. Iribarren and E. Moro, “Impact of human activity patterns on the dynamics of information diffusion,” *Physical Review Letters*, vol. 103, no. 3, Article ID 038702, 2009.
- [10] V. W. Zheng, B. Cao, and Y. Zheng, “Collaborative filtering meets mobile recommendation: a user-centered approach,” *AAAI*, vol. 10, pp. 236–241, 2010.
- [11] K. Partridge and B. Begole, *Activity-Based Advertising*, Springer, London, UK, 2011.
- [12] W. Bohte and K. Maat, “Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: a large-scale application in the Netherlands,” *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 3, pp. 285–297, 2009.
- [13] R. Kitamura, “Panel analysis in transportation planning: an overview,” *Transportation Research Part A: General*, vol. 24, no. 6, pp. 401–415, 1990.

- [14] E. Çelikten, G. L. Falher, and M. Mathioudakis, "Extracting patterns of urban activity from geotagged social data," 2016, <https://arxiv.org/abs/1604.04649>.
- [15] M. Allahviranloo and W. Recker, "Mining activity pattern trajectories and allocating activities in the network," *Transportation*, vol. 42, no. 4, pp. 561–579, 2015.
- [16] 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.
- [17] X. Chen, X. Wan, Q. Li et al., "Trip-chain-based travel-mode-shares-driven framework using cellular signaling data and web-based mapping service data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 3, pp. 51–64, 2019.
- [18] S. Phithakkitnukoon, T. Horanont, G. Di Lorenzo, R. Shibasaki, and C. Ratti, "Activity-aware map: identifying human daily activity pattern using mobile phone data," *Human Behavior Understanding*, pp. 14–25, Springer Berlin Heidelberg, Berlin, Germany, 2010.
- [19] L. Shen and P. R. Stopher, "A process for trip purpose imputation from Global Positioning System data," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 261–267, 2013.
- [20] G. G. Langlois, H. N. Koutsopoulos, and J. Zhao, "Inferring patterns in the multi-week activity sequences of public transport users," *Transportation Research Part C: Emerging Technologies*, vol. 64, pp. 1–16, 2016.
- [21] A. Alsger, 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.
- [22] C. Chen, S. Jiao, S. Zhang, W. Liu, L. Feng, and Y. Wang, "TripImputor: real-time imputing taxi trip purpose leveraging multi-sourced urban data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 10, pp. 3292–3304, 2018.
- [23] E. F. D. S. Soares, K. Revoredo, F. Baião, C. A. M. S. Quintella, and C. A. V. Campos, "A combined solution for real-time travel mode detection and trip purpose prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4655–4664, 2019.
- [24] L. Wang, W. Ma, Y. Fan, and Z. Zuo, "Trip chain extraction using smartphone-collected trajectory data," *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, pp. 255–274, 2019.
- [25] R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted Boltzmann machines for collaborative filtering," in *Proceedings of the 24th International Conference on Machine Learning*, pp. 791–798, Corvallis, OR, USA, 2007.
- [26] Y. Koren, "Factorization meets the neighborhood: a multifaceted collaborative filtering model," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 426–434, Las Vegas, NV, USA, 2008.
- [27] M. Mørup, "Applications of tensor (multiway array) factorizations and decompositions in data mining," *WIREs Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 24–40, 2011.
- [28] S. Rendle, L. B. Marinho, A. Nanopoulos, and L. Schmidt-Thieme, "Learning optimal ranking with tensor factorization for tag recommendation," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 727–736, Paris, France, 2009.
- [29] Y. Wang, Y. Zheng, and Y. Xue, "Travel time estimation of a path using sparse trajectories," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 25–34, New York City, NY, USA, 2014.
- [30] B. Ran, H. Tan, Y. Wu, and P. J. Jin, "Tensor based missing traffic data completion with spatial-temporal correlation," *Physica A: Statistical Mechanics and Its Applications*, vol. 446, pp. 54–63, 2016.
- [31] X. Chen, Z. He, and L. Sun, "A Bayesian tensor decomposition approach for spatiotemporal traffic data imputation," *Transportation Research Part C: Emerging Technologies*, vol. 98, pp. 73–84, 2019.
- [32] B. A. Kumar, L. Vanajakshi, and S. C. Subramanian, "Pattern-based time-discretized method for bus travel time prediction," *Journal of Transportation Engineering, Part A: Systems*, vol. 143, no. 6, Article ID 04017012, 2017.
- [33] Q. Ma, H. Yang, A. H. Zhang, K. Xie, and Z. Wang, "Modeling and analysis of daily driving patterns of taxis in reshuffled ride-hailing service market," *Journal of Transportation Engineering, Part A: Systems*, vol. 145, no. 10, Article ID 04019045, 2019.
- [34] T. Pei, C. Zhou, A.-X. Zhu, B. Li, and C. Qin, "Windowed nearest neighbour method for mining spatio-temporal clusters in the presence of noise," *International Journal of Geographical Information Science*, vol. 24, no. 6, pp. 925–948, 2010.
- [35] T. Pei, W. Wang, H. Zhang, T. Ma, Y. Du, and C. Zhou, "Density-based clustering for data containing two types of points," *International Journal of Geographical Information Science*, vol. 29, no. 2, pp. 175–193, 2015.
- [36] P. Zhang, Y. Huang, S. Shekhar, and V. Kumar, "Correlation analysis of spatial time series datasets: a filter-and-refine approach," *Advances in Knowledge Discovery and Data Mining*, vol. 2637, pp. 532–544, 2003.
- [37] M. Nanni and D. Pedreschi, "Time-focused clustering of trajectories of moving objects," *Journal of Intelligent Information Systems*, vol. 27, no. 3, pp. 267–289, 2006.
- [38] Z. Wang, S. Y. He, and Y. Leung, "Applying mobile phone data to travel behaviour research: a literature review," *Travel Behaviour and Society*, vol. 11, pp. 141–155, 2018.
- [39] D. Vandembroucke, B. Bucher, and J. Cromptvoets, "Geographic Information Science at the Heart of Europe," Springer Science & Business Media, Berlin, Germany, 2009.
- [40] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [41] C. A. Andersson and R. Bro, "Improving the speed of multiway algorithms," *Chemometrics and Intelligent Laboratory Systems*, vol. 42, no. 1–2, pp. 93–103, 1998.
- [42] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, "Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering," in *Proceedings of the Fourth ACM Conference on Recommender Systems*, pp. 79–86, Barcelona, Spain, 2010.
- [43] H. A. L. Kiers, "Weighted least squares fitting using ordinary least squares algorithms," *Psychometrika*, vol. 62, no. 2, pp. 251–266, 1997.
- [44] 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 KDD*, vol. 96, no. 34, pp. 226–231, Portland, OR, USA, 1996.
- [45] S. Kisilevich, F. Mansmann, M. Nanni, and S. Rinzivillo, "Spatio-temporal clustering," in *Data Mining and Knowledge Discovery Handbook*, pp. 855–874, 2009.

- [46] X. Ma, Y.-J. Wu, Y. Wang, F. Chen, and J. Liu, "Mining smart card data for transit riders' travel patterns," *Transportation Research Part C: Emerging Technologies*, vol. 36, pp. 1–12, 2013.
- [47] A. Zhou, S. Zhou, J. Cao, Y. Fan, and Y. Hu, "Approaches for scaling DBSCAN algorithm to large spatial databases," *Journal of Computer Science and Technology*, vol. 15, no. 6, pp. 509–526, 2000.
- [48] C. A. Andersson and R. Bro, "The N -way toolbox for MATLAB," *Chemometrics and Intelligent Laboratory Systems*, vol. 52, no. 1, pp. 1–4, 2000.