



Liyan Liu · Hongxin Zhang · Jiaxin Liu · Shangching Liu · Wei Chen · Jiaju Man

## Visual exploration of urban functional zones based on augmented nonnegative tensor factorization

Received: 6 July 2020 / Revised: 30 September 2020 / Accepted: 2 October 2020 / Published online: 16 January 2021  
© The Visualization Society of Japan 2021

**Abstract** A city contains a variety of different urban functions with referring to the purpose of land use to support the diverse needs of urban residents, such as residence, working and recreation. Exploring urban functional zones is a critical task, which provides valuable applications for business site selection, transportation management and urban planning. It is well known that location information and human mobility semantics both are significant factors for identifying urban functional zones, just considering one factor is not effective. However, most of existing techniques capture the transformation of functional zones and interpret the results based solely on location or mobility semantics, and lack the capacity to deal with the multifaceted features of urban data. To tackle these problems, in this paper, we propose an interactive visual analytics system for effectively exploring urban functional zones based on spatio-temporal OD data and Points of Interest data. We first adaptively partition the territory into region units based on adaptive blue noise sampling method, extract POI feature matrix (location information) as a prior knowledge, model multidimensional spatio-temporal OD data as a tensor for addressing multifaceted features and propose an augment tensor-based algorithm that enables users to simultaneously combine mobility semantics and inherent location information for identifying functional zones. In addition, we design a set of visual encodings to better understand and interpret the results in a visual and intuitive manner. This system has been demonstrated using two case studies with a real-world dataset of HangZhou city and domain-expert interviews.

**Keywords** Tensor factorization · Visual analytics · Urban functions · Spatio-temporal data

---

L. Liu · J. Man  
State Key Laboratory of HPC&SIP (MOE of China) and College of Mathematics and Statistics, Hunan Normal University, Changsha, China  
E-mail: liyanliu\_hunnu@qq.com

J. Man  
E-mail: manjiaju@hunnu.edu.cn

H. Zhang (✉) · J. Liu · S. Liu · W. Chen  
State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou, China  
E-mail: zhx@cad.zju.edu.cn

J. Liu  
E-mail: jiaxinliu@zju.edu.cn

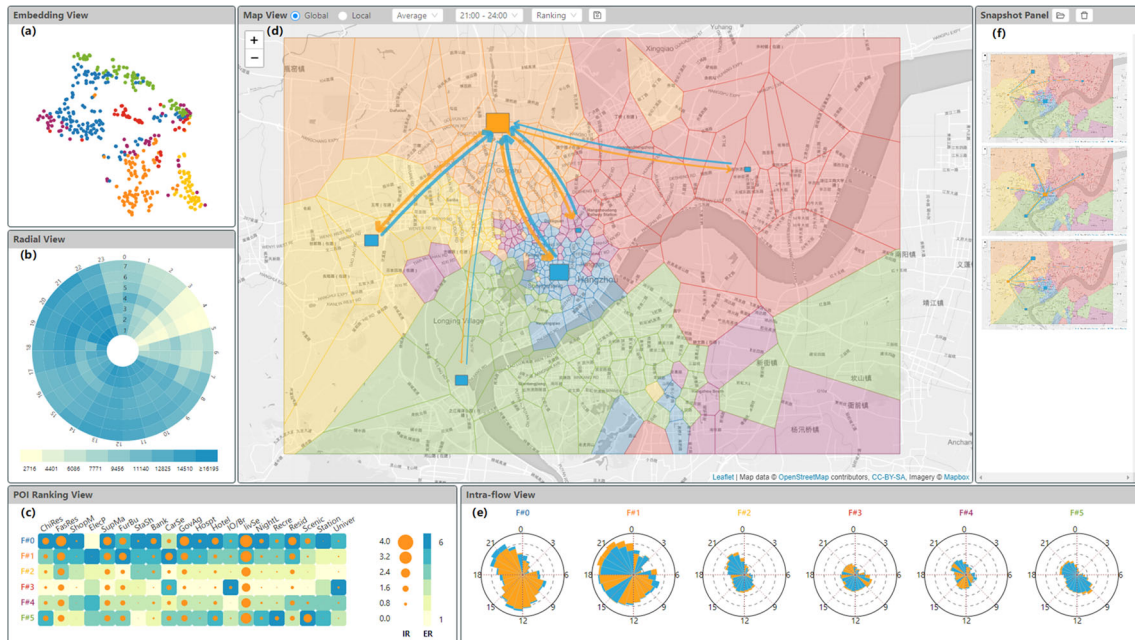
S. Liu  
E-mail: osmallfrogo.hchs@gmail.com

W. Chen  
E-mail: chenwei@cad.zju.edu.cn

## 1 Introduction

With the rapid development of a city, different functional zones gradually emerging within the modern city to serve its citizens' various requirements, such as business districts, residential areas, recreational blocks and educational zones. These functional zones may be artificially designed by urban planners or spontaneously formulated according to people's actual demands and dynamically change with the acceleration of urbanization and modern civilization. The exploration of urban functions is a fundamental research, which has the ability to reveal the physical and social characteristics of the city. Meanwhile, understanding zone functions not only provides numerous valuable applications such as business site selection, transportation management and urban planning, but also helps to solve a variety of urban-related problems such as increasing traffic capacity and enhancing location-based service. Yet despite its importance, very little work has been done to identify the dynamic changes of urban functions combined with human mobility semantics and inherent location information, and interpret functional zones via designing an interactive visual analytics system, which is our focus in this paper.

There has been prior work studying urban functions segmentation, most traditional techniques mainly adopt various kinds of clustering algorithms to classify functional zones (Adrienko and Adrienko 2011) Karlsson (2007). In recent years, the advent of location acquisition technologies and mobile computing, including GPS devices on vehicles and in mobile phones, a variety of large-scale urban data related to human mobility have been accumulated. As a result, data-driven approaches based on various urban data, such as taxi trajectory (Ferreira et al. 2013), mobile phone data (Di Lorenzo et al. 2016) and location-based social networks (LBSNs) (Von Landesberger et al. 2016), which provide an unprecedented opportunities to explore human mobility patterns. Wu et al. (2017) proposed MobiSeg system to investigate people's activity patterns for interactive region segmentation based on heterogeneous mobility data. Zhou et al. (2018) proposed a visual analysis system for exploring of urban functions based on spatio-temporal taxi OD trips. Nevertheless, these studies only considered people's movement behaviors without considering geographic location information to explore urban functions. Yuan et al. (2012, 2015) discovered functional zones in a



**Fig. 1** The visualization interface facilitates the exploration of urban functional zones via six major views, namely, **a** the embedding view of a given area illustrates region units as points in the embedding space. **b** The radial view shows the total flow volume through seven days. **c** The POIs ranking view presents the POI configuration to further compare and interpret functional zones. **d** The global map view provides the overview of the distributions of functional zones in geospatial as well as displays the spatio-temporal traffic flow graphs among functional zones to visualize human activities in the global mode. And it can display the magnitude of local region units' traffic volume in each functional zone in the local mode. **e** The intra-flow view shows the dynamic change of pickup/dropoff volume within functional zone over time to discover the differences among different zones. **f** The snapshot panel keeps the records of exploration for further comparison

city based on a topic-modeling-based approach, and using both location and mobility semantics, which are the most relevant to our work. However, it is hard for users to understand or interpret the results of urban functional zones without an interactive visual analytics system and a prior knowledge of these regions.

In this paper, we propose an interactive visual analytics system, which supports the exploration of urban functional zones. Firstly, we employ the adaptive blue noise sampling to partition urban area into a set of small region units with semantics. Based on partition region units, we extract POI feature matrix from POI data utilizing the term frequency-inverse document frequency (tf-idf) (Salton et al. 1975) method and model spatio-temporal OD data as a three-dimensional tensor. And then we propose a novel augmented nonnegative tensor factorization simultaneously considering both human mobility semantics and location information to identify the urban functional zones rather than leverage the solely one factor. Subsequently, a series of insightful visual encodings are designed enabling users to quickly analyze and compare the results of urban functional zones from multiple perspectives and different levels of details. Finally, with these techniques, we develop a visualization framework to visualize and explore urban functional zones and their attributions. A set of convenient user interactions are provided to help users to conduct a high-efficiency exploration and analysis the properties of urban functions. Case studies based on real-world datasets and interviews with domain experts are conducted to demonstrate the effectiveness of our proposed system. To the best of our knowledge, this system is the first to employ augment based-tensor method with combining mobility semantics with location information to explore urban functional zones.

The major contributions of our study are as follows:

- We explore urban functional zones utilizing both spatio-temporal OD data and POI data and adaptive partition urban area into region units.
- We introduce a novel augmented nonnegative tensor factorization to identify functional zones combining with human mobility semantics and inherent location information.
- We propose an interactive visual analytics system to empower domain experts to analyze and compare urban functions from multiple perspectives and different levels of details.

The rest of this paper is organized as follows: the related work about is summarized in Sect. 2. Section 3 presents the analysis tasks and the system overview. The urban functions extraction is detailed in Sect. 4. The visual interface is further introduced in Sect. 5. Case studies as well as domain-expert interviews are described in Sect. 6 and finally we present the discussion and draw our conclusions in Sects. 7 and 8, respectively.

## 2 Related work

In this section, we review techniques that are most relevant to our work, including the tensor decomposition, visual analysis of urban data and explore of functional zones.

### 2.1 Tensor factorization of urban data

Tensor factorization which was first introduced by Hitchcock (1927) is important tool for dealing with tensor data. It can be regard as higher-order extensions of the matrix singular value decomposition (SVD) (Klema and Laub 1980). Two major tensor factorization we are familiar with, one is known as the Tucker factorization and the other is known as the canonical polyadic (CP) decomposition (a.k.a., PARAFAC or CANEDCOMP). Tucker factorization was introduced by Tucker (1963) and later redefined by in Levin (1965) and Tucker (1966); Tucker et al. (1964), which attempt to present a tensor into the product of some low-order matrices and core tensors. Later, Carroll and Chang (1970) and Harshman (1970), respectively, proposed CANDECOMP and PAEAFAC decomposition, which decompose a tensor into a sum of rank-one tensors. Compared with the matrix-based approaches such as matrix factorization (Lee and Seung 1999, 2000), the tensor-based method has the capability to model multifaceted data and directly cope with the correlations among multiple dimensions. Therefore, it have been successfully applied for heterogeneous spatio-temporal data analysis. For example, Cao et al. (2018) capture the semantics of streaming ST data based on CP decomposition. Liu et al. (2019a) apply successive rank-one CP decomposition methods to support automated partitioning and multidimensional pattern extraction on spatio-temporal data. Sun and Axhausen (2016) use a probabilistic tensor factorization framework to understand urban mobility patterns.

## 2.2 Visual analysis of urban data

With the increasing availability of location acquisition technologies, more and more urban data related to human behaviors has become available. It is well known that a variety of useful information is hidden in large-scale urban data. Therefore, visual analysis of urban data is widely studied and effectively provides many valuable applications, such as traffic simulation (Wang et al. 2018), route recommendation (Zhu et al. 2019), urban planning (Deng et al. 2020), crowd pattern discovery (Xu et al. 2017; Mingliang et al. 2019; Xu et al. 2015; Lv et al. 2019) and events exploration (Yingcai et al. 2018). Liu et al. (2017) proposed SmartAdP, an interactive visual analytics system that support selecting the optimal billboard locations based on large-scale taxi trajectories. Liu et al. (2019b) combined temporal and geospatial aggregation of traffic data for creating legible traffic flow maps and employed interactive topic modeling to capture latent semantic information. Chen et al. (2018) designed and implemented a VAUD system that supports the visualization, querying and exploration of urban data. Zhu et al. (2019) proposed a web-based visual exploration system based on the mobile trajectory data that supported the comprehensive analysis of human mobility, location functionality and traffic assessment. Different from the prior studies, our research focuses on explore urban functional zones and reveal the mutual influence between urban locations and mobility behavior by utilizing visual analytics techniques.

## 2.3 Exploration of urban functional zones

The study of urban functions is an important research topic in the fields of GIS and urban planning in a long time for supporting policy making, resource allocation and related research. Early studies mainly rely on remote-sensing techniques, classic theory, clustering algorithms, long-term observation and so on. Deng et al. (2009) and Seto and Fragkias (2005) conducted comparative analyses of satellite images of cities to classify regions based on land coverage. Vatsavai et al. (2011) thoroughly compared classification techniques based on high-resolution remote sensing images. Karlsson (2007) gave an overview of related works which were mainly based on clustering algorithms. Recently, the topic-modeling-based methods are widely applied to explore urban functional zones. For example, Wu et al. (2017) proposed an interactive visual analytics system, Mobiseg, which supported the exploration of people's movement activities to segment the urban area into regions sharing similar activity patterns. Yuan et al. (2012, 2015) creatively explored urban functional zones using modified Latent Dirichlet Allocation (LDA) model (Blei et al. 2003), which was first one that simultaneously combined location semantics of a region and mobility semantics. Yao et al. (2018) proposed a framework to lean the vector representation of city zones by exploiting large-scale taxi trajectories. Gao et al. (2017) developed a statistical framework to help discover semantically meaningful topics and functional regions based on LDA topic modeling. Zhou et al. (2018) proposed a visual analysis system for exploration of urban functions based on spatio-temporal taxi OD trips. In addition, Cai et al. (2019) extracted the dynamic semantics of urban spaces based on tensor factorization.

In our work, we present an interactive visual analytics system that lays emphasis on the exploration of urban functional zones. Comparison of most previous research, the biggest differences are as follows. First, our work partition the territory into region units with semantics, which is superior to the grid-based division methods. Second, our work utilizes combined human mobility and as prior knowledge of location information to capture the transformation of urban zones. Third, we model spatio-temporal OD data as a tensor to tackle the multifaceted features and propose an augmented tensor-based algorithm which has the ability to combine mobility semantics and location information. Although Yuan's method (Yuan et al. 2012, 2015) and Gao's method (Gao et al. 2017) are also combined human mobility and geographical location information, but our work captures functional zones based on tensor-based algorithm rather than based-topic-modeling and design a visual analytics system can help users to better understand or interpret urban functional zones.

## 3 Problem characterization

In this section, we firstly describe the types of data used in this paper. Subsequently, we summarize analytical tasks identified from the interviews of domain experts, and then we provide a overview of our system workflow.

### 3.1 Data abstraction

The taxi trajectory data and POI data of HangZhou city are used in this paper to demonstrate the effectiveness of our system. The detailed information is described as follows:

The taxi OD data are provided by the Hangzhou bureau of transportation. The data consist of a series of GPS records of over 20,000 taxi with around 916 million records per day. Each sampling point of a trajectory contains a record of taxi's ID, pickup/drop-of timestamp, GPS coordinates (latitude and longitude) and status (occupied/vacant). The time range of the dataset is from April 1–7 2015. Moreover, we mainly focus on the urban area of HangZhou, where the longitude is in [119.9638, 120.4373] and the latitude is in [30.1329, 30.4093].

The POI data are collected from Baidu API and contain 37,844 points in the city. Each POI is denoted by its ID, category and GPS location.

### 3.2 Task analysis

In order to comprehensive and efficient the exploration the urban functional zones, we discuss with the domain experts in the form of structured interviews over a 4-month period. One domain expert is an analyst from an urban planning bureau in HangZhou, China (E1). The other expert is a professor with domain knowledge in urban data visual analytics and visualization (E2). Based on the discussions with domain experts, we compile a list of analytical tasks as follows.

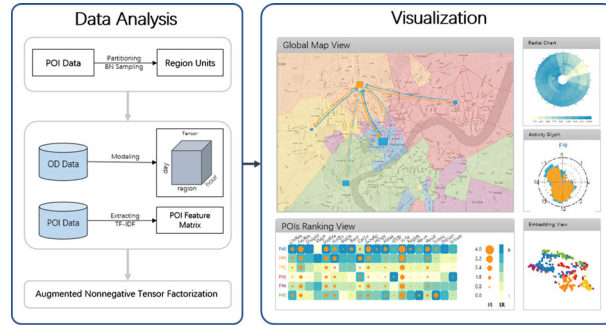
**R.1 Identify the functional zones.** Different from uniform-grid-based strategy, we need to segment the urban area of a city into local region units in a more reasonable manner? How to capture the inherent location information? How to extract the mobility semantics of the multidimensional spatio-temporal OD data? How to employ effective method with combining human mobility semantics and location information to identify the functional zones?

**R.2 Visualize and Interpret the functional zones from different aspects.** How to design meaningful Glyph for visualizing the variation of pickup/dropoff volume over time in each functional zone and the variation of traffic flow graph over time among functional zones to help users understanding activity patterns? How to present the ranking of POIs among functional zones to further interpret the results?

**R.3 Facilitate inter-region comparisons.** What are the differences and similarities among local region units? How do you provide a series of user interactions to compare the differences between urban region units? Specifically, (1) users enable to explore the proportion of POIs between region units, (2) users have the ability to discover the difference the pickup/dropoff volume over time in different region units.

### 3.3 System overview

Figure 2 illustrates an overview of the proposed system. The workflow of our visual analytics system for the exploration of urban functional zones mainly consists of two stages: the data analysis stage and the visualization stage. To be more specific, the data analysis stage includes four module. Firstly, we adaptively partition the territory into local region units through an adaptive blue noise sampling method. Secondly, we capture inherent geographical location information based on TF-IDF method. Thirdly, we model multi-dimensional spatio-temporal OD data as a three-dimensional tensor, and the data schema is  $(region \times day \times time) \rightarrow traffic\_volume$ , describing the pickup/dropoff volume in different regions on a specific date and time. Finally, we propose a novel augmented nonnegative tensor factorization (A-NTF) algorithm to identify urban functions with combing human mobility semantics and location information (**R1**). The visual exploration stage consists of a series of views and provides a user-friendly interactions to help users understand and interpret the urban functional zones. The global map overview has the ability to visualize the results of urban functional zones, display the spatio-temporal traffic volume graphs within functional zones as well as show the magnitude of local region units' traffic volume in each functional zone (**R2**). The POIs ranking view, the intra-flow view and the radial view are employed to further present differences and similarities between different functional zones (**R3**). The system we designed offers an intuitive understanding and interpretation urban functional zones in an interactive manner (**R2**, **R3**).



**Fig. 2** The pipeline of our proposed visual system, which is comprised of two components including data analysis and visualization

## 4 Urban functions extraction

In this section, we first perform territory partition to get local region units and extract POI feature matrix and then describe how to model multi-dimensional spatio-temporal OD data as a tensor. Subsequently, we propose a novel augmented nonnegative tensor factorization algorithm with combining human mobility semantics and inherent location information to identify functional zones of city.

### 4.1 Adaptive partitioning the territory

In order to explore urban functional zones, the important step is territory partition with dividing the urban area of a city into local region units. An effective partition method should be able to preserve the distribution of data as much as possible. According to our common sense, the more prosperous area contains the more urban functions. Therefore, we need to divide the area into smaller region units. By analyzing the characteristics of OD data and POI data, we discover that the spatial coverage of OD data is limited. And compared with different POI categories, we find that data of Chinese restaurant have a much better spatial coverage, see Fig. 3a. We implement an adaptive blue noise sampling method based on Poisson disk (Bridson 2007) to reduce the scale of POI data and get sampling points with preserving the distribution of data as shown in Fig. 3b. And then we employ the sampling points as seed cells to compute a Voronoi tessellation of the urban area on the basis of Delaunay triangulation. The resulting Voronoi cells correspond to region units. As illustrated in Fig. 3c, the area with higher Chinese restaurant density is divided into more region units. Intrinsically, it provides us with a more natural and semantic segmentation of urban space than the grid-based partition methods, and region units carry the social-economic functions.

We use adaptive blue noise sampling rather than other clustering algorithms, the main reason is that it can meet the needs of users to adaptively adjust the number of sampling points and can achieve the sampling without deviation, minimum distance and maximum coverage. Specifically, a Poisson disk radius of  $r$  at the trail location  $p$  is built adaptively.  $r$  generalizes the distance constraint in Poisson disk sampling for a point. Kernel density estimation (KDE) is employed to estimate distribution of POI points. Given  $m$  Chinese restaurant points  $P = \{p_1, p_2, \dots, p_m\}$ . Mathematically, the density at location  $p$  is computed as follows:

$$f(p) = \sum_{i=1}^m \frac{1}{mr^2} K\left(\frac{d_{i,s}}{h}\right), l \quad (1)$$

where  $d_{i,p}$  is the distance from  $p_i$  to the  $s$ -th region,  $h$  is the bandwidth and  $K(\cdot)$  is a kernel function which determines the smoothing degree of the reconstructed density field. A Gaussian kernel is used in our approach. Then the radius  $r$  of poisson disk is defined as  $af(p)$ .  $a$  is a user-defined parameter, which is able to control the sampling rate.

### 4.2 Extracting POI feature matrix

It is well known that there is an inseparable relationship between geographic location information and urban functional zones, such as a region is a high probability considered as a business area if there are large number of shopping malls. The allocation of POI uncovers the potential socioeconomic activities of a



**Fig. 3** **a** The data of Chinese restaurant are drawn on the map in the form of a scatter chart. **b** The sampling points based on adaptive blue noise sampling method using Chinese restaurants category. **c** Region units generated on the basis of Delaunay triangulation using the sampling points as Voronoi cells

region. Therefore, geographic location information is an important factor to help explore urban functional zones.

In order to capture geographic location information, we calculate the distribution of POIs for each region to get POI feature matrix. For each region  $r_i, i = 1, 2, \dots, R$ , we first count the number of POIs in each POI category and then measure the importance of a POI in a local region unit based on the Term Frequency-Inverse Document Frequency (TF-IDF) (Salton and Yang 1973) method. To be more specific, for a given region  $r_i$ , we calculate the POI feature vector,  $Y_i = (y_{i1}, y_{i2}, \dots, y_{iM})$ , where  $y_{ij}$  represents the TF-IDF value of the  $j$ -th POI category and  $M$  is the number of POI categories. The value of TF-IDF  $y_{ij}$  is given by:

$$y_{ij} = \frac{f_{j,r_i}}{N_{r_i}} \times \log \frac{|R|}{f_{j,R}}, \quad (2)$$

where  $R$  is the number of region,  $f_{j,r_i}$  equals the number of POIs belonging to the  $j$ -th category which appears in region  $r_i$ , and  $N_{r_i}$  is the number of POIs located in region  $r_i$ .  $f_{j,R}$  denotes the number of regions in which the  $j$ -th category appears in  $R$ . Thereafter, the POI feature matrix is denoted by the matrix  $Y \in R_+^{R \times M}$ , where the  $i$ -th row of the matrix  $Y$ , is regarded as geographic location information for region  $r_i$ . The matrix  $Y$  is used as a prior knowledge to embed into the model that we introduce below.

### 4.3 Modeling spatio-temporal OD data as a tensor

A tensor, denoted as  $\chi$ , is a multi-way array or multi-dimensional matrix. Tensors are obviously generalization of vectors and matrixes, for example, a scalar is a zero-order tensor, a vector is a first-order tensor, a matrix a second-order tensor and tensors of third-order and higher are called higher-order tensors. The order of a tensor is the number of dimensions, also known as ways or modes. Generally, we denote an N-way nonnegative tensor  $\chi \in \mathbb{R}_{\geq 0}^{I_1 \times I_2 \times \dots \times I_N}$  has N modes with the dimensionality of  $I_1, I_2, \dots, I_N$ , respectively.  $\mathbb{R}_{\geq 0}$  represents that all the elements of  $\chi$  contain nonnegative values, which commonly applies to situations when data represent number of observed instances.

Most traditional data analysis techniques lack the ability to handle multi-way or multifaceted features, while tensors can meaningfully capture changes in multi-dimensional spatio-temporal characteristics of large-scale traffic data simultaneously. Thus, we model spatio-temporal OD data as a three-dimensional tensor  $\chi \in \mathbb{R}_+^{R \times D \times T}$ , where  $R$  is the number of regions,  $D$  indicates the number of days and  $T$  represents the number of time spans.

- Region dimension: we segment the entire urban areas into region units  $r = [r_1, r_2, \dots, r_R]$ ,  $R = 425$  in our implementation.
- Day dimension: we divide the duration of observation into equal day spans  $d = [d_1, d_2, \dots, d_D]$ .
- Time dimension: we divide every day into equal time spans  $t = [t_1, t_2, \dots, t_T]$ .

Correspondingly,  $\chi_{ijk}$  stores the traffic volume (e.g., total number of pickup and drop-off) at region  $r_i$  on the  $j$ -th day during hour  $k$ .

### 4.4 Augmented nonnegative tensor factorization (A-NTF)

Nonnegative tensor factorization (NTF) can decompose a tensor into the product of some low-order latent nonnegative factors and core tensors, and it considers the relationship among the multi-dimensions of data. We use the NTF method Tucker to decompose the tensor  $\chi$  to obtain the core tensor  $S \in \mathbb{R}_+^{P_1 \times P_2 \times P_3}$ , the region factor  $A \in \mathbb{R}_+^{R \times P_1}$ , the day factor  $B \in \mathbb{R}_+^{D \times P_2}$  and the time factor  $C \in \mathbb{R}_+^{T \times P_3}$ . The objective function is defined as follows:

$$\min \|\chi - S \times_A A \times_B B \times_C C\|_F^2, \quad (3)$$

where  $P_1, P_2, P_3$  are usually very small, indicating the number of latent factors, and  $\|\cdot\|$  represents the Frobenius norm of a tensor (as well as a matrix or a vector). The symbol  $\times_A$  denotes the tensor-matrix multiplication, where the subscript  $A$  stands for the direction, the core tensor whose diagonal elements represent the relative strength of the corresponding patterns.

In our study, we employ augmented NTF (A-NTF) for identifying urban functional zones with simultaneously considering both mobility semantics and POI information. In order to combine the information of both, we embed the POI feature matrix obtained in Sect. 4.2 into the factorization process of A-NTF. The objective function is defined as follows:

$$\min \|\chi - S \times_A A \times_B B \times_C C\|_F^2 + \lambda \|Y - AG\|_F^2, \quad (4)$$

where  $Y \in \mathbb{R}_+^{R \times M}$ ,  $G \in \mathbb{R}_+^{P_1 \times M}$ ,  $P_1$  denotes the number of urban functional zones,  $P_2, P_3$  denote the number of days and time reductions, respectively. And  $M$  denotes the number of POI categories,  $\lambda$  is parameter controlling the contribution of the POI feature matrix  $Y$ . In our model,  $\chi$  and  $Y$  shares matrix  $A$ , thus the results of regions segmentation that comprehensively consider both human mobility semantics and geographical location information.

Our model based on A-NTF take into account the richer context information in the spatio-temporal OD data and POI data, such as the temporal and spatial dimension in a unified manner and the location information is embedded as a prior knowledge in the process of our model decomposition. We utilize alternating least square (ALS) algorithm (Kim and Choi 2007; Cichocki et al. 2009) to solve the above optimization problem, as presented in Algorithm 1. In addition, we initialize the matrices based on NMF (Nonnegative Matrix Factorization) method to find good initial starting points, which can accelerate convergence of solutions.



**Algorithm 1** Augmented Nonnegative Tensor Factorization.**Input:** Nonnegative tensor  $\chi$ , and matrix  $Y$ , rank  $P_1, P_2, P_3$ **Output:** Nonnegative factors  $S, A, B, C$ 

- 1: Initialize  $A, B, C, G$ ;
- 2:  $A \leftarrow$  NMF initialization  $(\chi_{(1)}, P_1)$
- 3:  $B \leftarrow$  NMF initialization  $(\chi_{(2)}, P_2)$
- 4:  $C \leftarrow$  NMF initialization  $(\chi_{(3)}, P_3)$
- 5:  $G \leftarrow$  NMF initialization  $(G, P_1)$
- 6: While  $L_t - L_{t+1} > \epsilon$ ;
- 7:  $A \leftarrow A \otimes \frac{\chi_{(1)} S_A^{(1)T} + \lambda(F-AG)G^T}{AS_A^{(1)} S_A^{(1)T}}$ ;
- 8:  $B \leftarrow B \otimes \frac{\chi_{(2)} S_A^{(2)T}}{BS_A^{(2)} S_A^{(2)T}}$ ;
- 9:  $C \leftarrow C \otimes \frac{\chi_{(3)} S_A^{(3)T}}{CS_A^{(3)} S_A^{(3)T}}$ ;
- 10:  $G \leftarrow G \otimes \frac{A^T F}{A^T A G}$ ;
- 11:  $S \leftarrow \max(\chi \times_A A^+ \times_B B^+ \times_C C^+, \epsilon)$ ;
- 12:  $S \leftarrow S \otimes \frac{\chi \times_A A^T \times_B B^T \times_C C^T}{S \times_A A^T \times_B B^T \times_C C^T}$ ;
- 13: **return**  $S, A, B, C$ ;

## 5 Visual interface

We design a visual analytics system to facilitate the exploration and comparison of urban functional zones. As illustrated in Fig. 1, the visual interface of our system consists of six components, such as an embedding view, a radial chart, a POIs ranking view, a global map view, an intra-flow volume view and Snapshot panel.

### 5.1 The map view

As shown in Fig. 1d, to support an effective understanding and interpretation urban functional zones, a major view of the proposed system is a global map view which consists of the results layer, the spatio-temporal traffic flow layer and the traffic volume layer.

The results layer provides an overview of the urban area of studies as well as intuitively visual rendering the distributions of urban functional zones in geospatial. From this layer, we can clearly and quickly discover the entire urban area is segmented into six urban functional zones and the region units that could have similar functions filled with the same color. Not surprisingly, we extract the results of urban functions of segmentation through A-NTF method which semantics information is very ambiguous. The visual strategies are urgently introduced to have a comprehensive interpreting of zones' real functions from multiple perspectives. One of the strategy is provided the spatio-temporal traffic flow layer with visualizing the global human activity, which presents the traffic flow graph to reveal the variation of the human activity flows over time among different urban functional zones. When users select a time interval in the global mode, the spatio-temporal traffic flow graph is evoked. Each urban functional zone is mapped as box with different border-width. The wider border indicates that the greater traffic flow within this functional zone. We normalize the traffic flow through the sum of flow volume in the corresponding regions divided by the area of the regions to avoid the larger traffic flow due to larger area of regions. When user hovers over the box, the pickup/dropoff flows between this functional zones and others is activated. The direction of arrow represents the flow direction. The thicker line indicates the greater flow. It helps users understand global human activity rules among different functional zones. Another strategy is provided the traffic volume layer, which displays the magnitude of local region units' traffic volume in each functional zone. When users select a time interval in the local mode, the local traffic volume graph is evoked. In the same functional zones, it is clearly the darker the color is, the higher traffic volume in the region unit. In addition, when users select a functional zone by Ranking button, the ranking list for traffic flow corresponding to the local region units in the functional zone is presented on the right side of map view. When the user clicks on region unit in the ranking list, the corresponding area unit is highlighted on the map. It helps users understand local human activity rules among region units and pick up region units with high traffic flow in the functional area more quickly.

## 5.2 The embedding view

The embedding view shows the global pattern of urban functional zones in the embedding space. As shown in Fig. 1a, we perform the state-of-the-art methods called t-distributed stochastic neighborhood embedding (t-SNE) (Van Der Maaten and Hinton 2008), which outperforms other standard techniques such as principal component analysis (Jolliffe 1986) and multidimensional scaling (Kruskal 1964) in revealing the implicit groupings of data items in visualization application (Van Der Maaten and Hinton 2008), to project the region factor vectors into a low-dimensional space and present them as colored points. Each local region unit represents a point. The six colors can be found from this view and the color of the point is the same as corresponding the color of the functional zones. According to the colors of points, we can find the most points with the same color are clustered together which indicates the same urban functional zones might share similar semantic. Therefore, the results largely demonstrate the usability of A-NTF method used to segment urban functions.

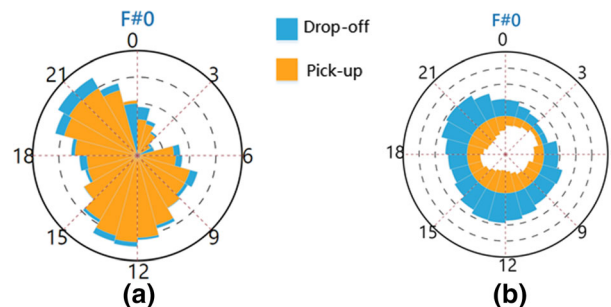
## 5.3 The radial chart

As shown in Fig. 1b, we employ the radial chart to show the total flow volume through seven days of a functional area. Every day is divided into 24 h and each unit reflects the flow volume in an hour. Based on the color bar, it is clearly that the darker the color is, the more traffic flow is. In addition, the mode of global, the total flow volume of functional zones can be showed by clicking the corresponding zones. And the mode of local, the total flow volume of local region units can be displayed by clicking the corresponding region units. Hence, it has the ability to compare the characteristics of traffic volume distribution.

## 5.4 The intra-flow view

The spatio-temporal traffic flow graph has the ability to display human activity among functional zones. We hope to design a view to show the human activity rules with each functional zone. As shown in Fig. 1e, the intra-flow view is provided to visualize human activity rules within each functional zone and help users to discover the difference among different urban functional zones.

*Glyph design.* To offer a general understanding and comparison the human mobility characteristics of urban functional zones, it is indispensable for us to provide a good glyph. We consider two aspects about the design of glyph, on the one hand, the designed glyph has the ability to display and compare human activity rules of functional zones. On the other hand, we expect to design a familiar metaphor so as to greatly improve comprehension and reduce cognitive burden on working memory. Inspired by the wheel and clock, we come up with a novel visual metaphor to display the variation of intra-flow over time. Figure 4a shows our design. In the top of text colors corresponds to the colors of encoding urban functional zones and indicates displays the intra-flow corresponding to zone. The 24 stacked bars indicate the variation of pickup/dropoff volume over time with functional zone and each stacked bar represents an hour in a day. The volume in different functional zone differs greatly, so directly mapping the volume will cause some visual errors. We normalize the data to solve this problem. The height of each bar indicates the relative size of the volume. The bar filled with yellow color represents the pickup volume of a functional zone and the bar with blue color indicates the dropoff volume. The glyph is generated in two steps: First, we calculate the value of pickup volume and dropoff volume per hour, sort the two values in an hour and select a larger value.



**Fig. 4** **a** A stacked glyph design to simultaneously present and compare the dynamic change of the traffic flow of pickup/dropoff over time. **b** The alternative glyph design

Second, we draw the larger value bar at first, then draw the smaller bar. Figure 4b presents the alternative design by using inwards and outwards bars. Although it is also easy to implement, users cannot quickly compare the magnitude of pickup/dropoff volume. Thus it undermines the efficacy of this design.

### 5.5 The POIs ranking view

As shown in Fig. 1c, the POIs ranking view presents the POI configuration in all functional zones through presenting the corresponding external ranking that rank the tf-idf value of each POI category in different functional zones as well as internal ranking that rank the tf-idf value of all POI categories in a functional zone. The ordinate is six urban functional zones and the color of text is consistent with that in map view. The abscissa lists 20 POIs categories. The external rank is represented by the color depth of each square block. The deeper color indicates the higher ranking. The internal rank is denoted by the embedding circle. The value in different POIs categories differ greatly, it's not effective to directly map the volume, thus we use the logarithm value to solve this problem. The radius of a circle encodes the logarithm value of external rank and the diameter of the maximum circle is equal to the edge length of the square block. The larger radius indicates the higher external ranked. This view together with the aforementioned two views (i.e., the spatio-temporal traffic flow graph within in the map view and the intra-flow view) empower users to understand and interpret the semantics information of functional zones.

### 5.6 Snapshot panel

Users may explore the spatio-temporal traffic flow of different functional zones at different times and the intra-flow of the local region units. The experts hope that our system has the ability to storing the exploration records to facilitate review of their existing operations. Thereby, their feedback prompted us to add the snapshot panel for storing historical records.

### 5.7 Interactions

We adopt the following interactions to support efficient exploration the urban functional zones and local region units among different view.

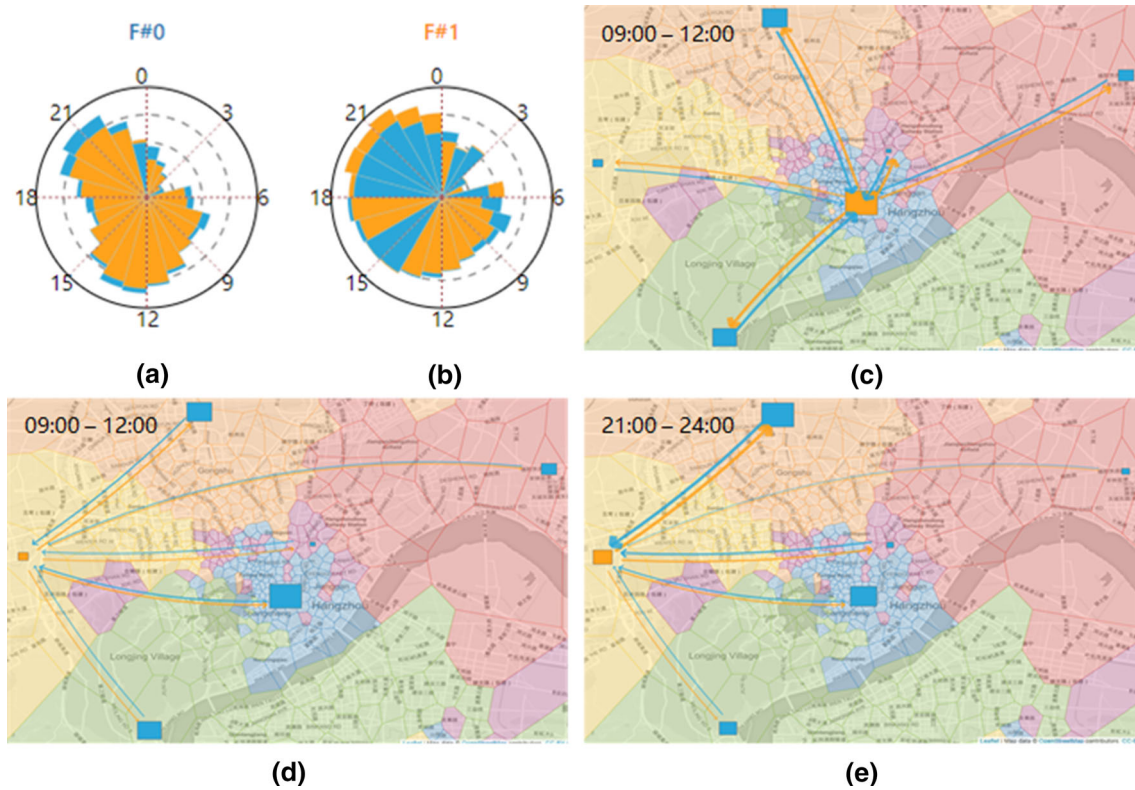
*Linking, highlight:* The system automatically connects the three view not only to support interactive pattern unfolding, but also to facilitate exploration and comparison the urban functional zones from multiple perspectives and different levels of details. Meanwhile, the highlighting feature is supported in the map view and the embedding view. In the global mode, when users double click on a region unit in map view, this region corresponds to entire functional zone is synchronously highlighted in embedding and map view. And the intra-flow view displays the results of this functional zone. Similarly, in the local mode, highlighting and showing the results of this region units.

*Configuration:* Users can configure our system by modulating the parameters such as the time to explore the dynamic change of the spatio-temporal traffic flow. And when users switch the mode of global or local, the traffic volume normalization mode will be change in intra-flow view.

*Zooming and scaling:* Map view and embedding view support zooming and panning to further explore local region units. Users can scroll the mouse wheel to zoom and drag the mouse to pan to have a good view of the map.

## 6 Case studies

We conduct two representative case studies with our domain experts to demonstrate the usability and usefulness of our system. Both studies are conducted based on the HangZhou city taxi OD data and POI data.



**Fig. 5** **a, b** The intra-flow of the functional zones F#0 and F#1, respectively. Two zones have a larger intra-flow volume of pickup/dropoff, but the dynamic change over time is very different. **c** The spatio-temporal traffic volume graph with the functional zones F#0 with other zones at 9:00 to 12:00, which indicates F#0 has tremendous the volume of inter-regional. **d, e** Display the spatio-temporal traffic volume graph with the functional zones F#1 with other zones at 9:00 to 12:00 and 21:00 to 24:00, which implies a large number of passengers visiting F#1 at night

### 6.1 Exploration functional zones of HangZhou

In this scenario, the experts utilize our designed system to interpret and compare the urban functional zones of HangZhou city. They first seek to overview of our system and then drill into the analysis and comparison of each functional zone.

From the results layer in map view Fig. 1d, the studied urban area is segmented into six functional zones. The experts discover most zones with similar urban functions are geometrically close, whereas zones with different functions are distant. This phenomenon is in accordance with our common senses. As shown in embedding view Fig. 1a, the experts can see that the most points with the same color are clustered, which indicates the same urban functional zones might share similar semantics. Based on observations above, they preliminary speculate that the results of urban area segmentation are reasonable and effective. And then the experts switch to several other views to quick overview the characteristics of all functional zones and the activities rules of the city. As shown in radial chart Fig. 1b, the city has more people's activity in the daytime than in the middle of the night. The whole city gradually becomes quite after around 11:00, progressively wakes up at around 6:00 and has obvious morning and evening peak hours. The traffic volume has a lower point in the day at between around 4:00 and 5:00, the experts consider the major reason may taxi shift causes this phenomenon. According to the external rank of the POIs ranking view Fig. 1c, the experts observe that the square block with deep color mainly allocate in functional zones F#0, F#1, F#5, which indicates that these functional zones are more mature and more developed areas than other zones. In additions, as illustrated in Fig. 1e, they notice that the magnitude of pickup/dropoff volume in the functional zones F#0 and F#1 is greater than other zones which implies two zones may be located in the downtown areas. Subsequently, the experts further explore each functional zone through different views and from different levels of details to have a comprehensive understanding of its real functions.

As shown in the POIs ranking view Fig. 1c, the experts observe the functional zone F#0 have a significantly higher external rankings in the commercial, restaurant, residential, hotels, medical and recreational POIs than other functional zones. They think this zone as a regional hub, because it has tremendous the volume of transportation within intra-regional and inter-regional which can reflect in the intra-flow view Fig. 5a and the spatio-temporal traffic flow graphs Fig. 5c. Compared with the volume of pickup and dropoff (Fig. 5a), the experts can see that most passengers visit this zone in the morning (5:00–9:00) and in the evening (5:00–11:00). The reason maybe is that people come here for work in the morning and entertainment in the evening. Thereby, the experts infer F #0 is dominated by the comprehensive business district or the bustling downtown area.

From Fig. 1c, the experts find that the functional zone F#1 contains the maximum number of electronic products store, furniture building materials market and car service POIs and it has the second highest rank of residential buildings, shopping mall, supermarket, hospital and so on. It shows that this zone has an adequate number of service supports the people's living. Furthermore, experts notice that F#0 and F#1 have a higher magnitude of pickup/dropoff volume as shown in Fig. 5a, b, but the dynamic change over time is very different. First, the traffic volume in F#1 are almost increasing at daytime and it is not obvious impact on traffic flow at the taxi shift time (4:00–5:00). Second, the volume of pickup is obviously greater than the volume of dropoff at night (6:00–24:00). These phenomena imply a large number of people living here. Based on these observations, the experts conclude that F #1 might be dominate by the developed residential areas.

Subsequently, the experts observe that the functional zone F#3 has the fewest POIs external rankings in most of the POI categories and the most characteristic POI categories in this zone is the entrance/Bridge, car service and university as shown in Fig. 1c. They suggest the industrial and college town might the most representative in F #3.

According to geographical location Fig. 1d, the experts notice that most region units of F#4 is located in the bounding area of the intersection between F#0 and F#1, F#2 as well as F#3. The POI configuration of this zone is similar to F#0, but in terms of the absolute quantity (compared the size of the circles), F#4 is less than F#0 as shown in Fig. 1c. Moreover, compared with Fig. 1e, the experts see that although the volume of intra-flow in this zone is lower than F#0, the dynamic change of intra-flow is similar to F#0. Therefore, the experts infer that F #4 may be dominate by the developing commercial/business/entertainment areas.

And then the experts clearly discover that the functional zone F#5 has the highest external rank of recreation and scenic and the second highest rank of Chinese restaurant, hotel, nightlife and station in Fig. 1c. In addition, Fig. 1e presents this zone has obvious morning and evening peak hours. Notably, the morning peak time is later than normal working hours, which indicates most people don't come here to work. Based on above analysis, the experts infer that F #5 is most likely dominated by the tourism and entertainment areas.

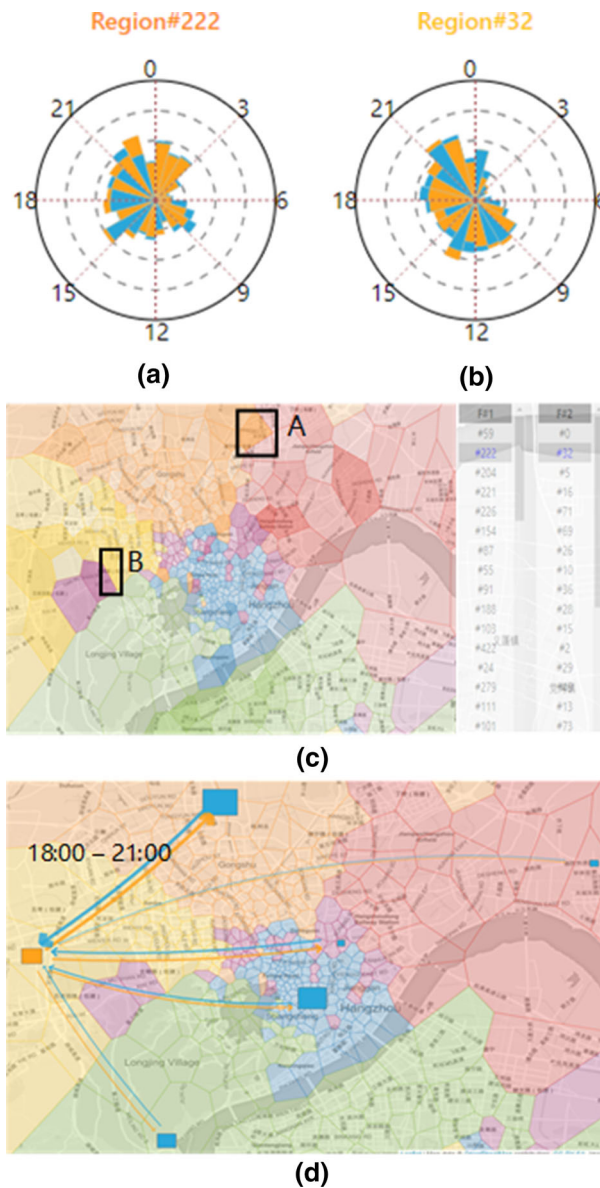
From the intra-flow view Fig. 1e, the experts find that the volume of pickup/dropoff at night is significant higher at daytime in functional F#2. In addition, they can see that the inter-regional flow F#2 as shown in Fig. 5d, e, the volume at night is larger than at daytime is further justification and a quantity passengers come from F#1 at 9:00 to 24:00. Based on the analyses, the experts speculate F #2 is mostly the entertainment/parks districts. Meanwhile, it further indicates that F#1 is dominate by the developed residential areas.

Finally, the experts find that the above results are convincing via comparing these results with a view of land use planning in Hangzhou which we contend as a form of ground truth.

## 6.2 Facilitating car-sharing sites

In this use case, we demonstrate how the experts can use our proposed system to facilitate inter-regional car-sharing sites. They mainly focus on understanding taxi demand in different functional zones to further facilitate selecting the optimal locations of the car-sharing. In recent years, the sharing cars appear with bringing the convenience of people's lives and especially become increasing popular with a group of young people. Appropriate car-sharing locations help people travel and go from one place to another easily, whereas inappropriate ones lead to waste of time and investment. To this end, the experts attempt to use our system to help them understanding the dynamic change of traffic volume at different functional zones, so as to decide when and where should be arranged.

The experts suggest that the great deal of traffic flow volume, which are likely to imply deficiencies of public transportation in such zones, thus shared cars would be needed. Furthermore, it is a dramatically different demand for car sharing at different functional zones and at different time. These areas such as the college town, residential zones, office areas, tourist attractions and recreation districts have a high demand for car sharing based on research, but different functional zones are inconsistent demand for car sharing at time aspects. On the basis of the above analysis, their criteria of selection as follows: (1) The zones have a higher proportion of selecting car-sharing as a transportation tool; (2) The local regions have a greater traffic flow within intra-regional and inter-regional. Based on these two criteria, they first seek to gain insight into the functional zones F#1 dominated by the developed residential areas and F#2 dominated by the entertainment/parks districts. Both two zones have a high demand for car sharing. Subsequently, the experts look over the spatio-temporal traffic volume to understand the traffic flow of inter-regional. As shown in Figs. 5e and 6d, they notice that the inter-regional flows between F#1 and F#2 is greater than other zones at night.



**Fig. 6** **c** The volume of local region units with each functional zone, the depth color indicates the largely traffic volume. The highlighted square A and B respectively present region#222 in functional zone F#1 and region#32 in functional zone F#2. Both regions have a greater traffic volume. **a**, **b** The intra-flow of the functional zones region#222 and region#32, respectively. **d** The spatio-temporal traffic volume graph with F#2 with other zones at 18:00 to 21:00

And then the experts hope to select local region units that have a greater local traffic volume within intra-regional at night. As shown in Fig. 6c, the experts find that region#222 and region#32 (highlighted with black square A and B) have a greater traffic volume through ranking list. Meanwhile, the experts further look at the dynamic change of the intra-flow over time, Fig. 6a, b present the intra-flow view of region#222 and region#32 through clicking on these two region units in the local mode of map view. It is clearly found that two region units have larger traffic volume at night. Finally, based on the knowledge obtain in our system, the experts make a decision that car-sharing can at the region#222 and region#32 and during 18:00 to 21:00 and 21:00 to 24:00.

### 6.3 Expert interview

The feedback of the domain experts are collected based on semi-structured interviews, which was summarized as follows:

*Visual design and interactions:* All domain experts confirmed that our system was nicely designed according to the problem domain and the characteristics of the used datasets as well as the interface was also quite user-friendly. In particular, our metaphor-based glyph design received high praises from the domain experts. They believed that the system could be easily understood by users without prior knowledge for urban areas. E1 commended “Our proposed system has the ability to effectively identify and visualize the urban functional zones, and provide an intuitive interpretation of the results for users without a prior knowledge of HangZhou”. E2 said “Our system integrate many visualization and interaction techniques, enabling me to explore and analyze the urban functions or urban region units from multiple perspectives and different levels of detail”.

*Comparison with previous work:* We asked the experts to further compare our method with previous work. Our work were familiar to Yuan’s (2012, 2015), Zhou’s (2018) and Cao’s (2018). The expert expressed our method as a pioneering study for exploring urban functions based on a A-NTF algorithm simultaneously combined human mobility and location information. E2 said “Compared with previous work, our work introduces tensor-based method which has the ability to cope with the correlations among multiple dimensions outperforms based-topic-modeling approach. And adaptive blue noise sampling method is introduced to adaptive partition the territory to get region units with strong semantics is superior to grid-based method. Besides, our designed an interactive visual analytics system which provides an intuitive visualization for users to easily understand and interpret the results of urban functional zones which has an advantage over Yuan’s work”.

*Usability and improvements:* The experts appreciated our system and found the functions provided by system quite useful. They all expressed that our system is useful in not only identifying urban functional zones and interpreting its real functions but also facilitating lots of real world applications such as business site selection. Apart from the aforementioned, our experts also provided some valuable suggestion. E1 commended, “ If system provides the POI configuration of the local region units to help users compare the proportions of POIs categories, it will be more better”.

## 7 Discussion

The case studies demonstrate the advantages of our system to explore urban functional zones based on A-NTF. The combination of the different visualizations enables uses to better understanding and interpretation functional zones form different aspects. Nevertheless, there is still space for improvement.

First, the datasets used in our system are only the spatio-temporal taxi OD data as people’s daily mobile behavior, but it may lose lots of the useful information. In the future work, we will use a variety of mobility data to more precisely explore urban functions. Second, we introduce a tensor-based algorithm A-NTF which is capable of dealing with multifaceted features. However, the A-NTF algorithm may become inefficient when the data grow (e.g., when handing data of multiple days, or with large number of region units). In the future work, several techniques can be considered to enhance the performance. For example, we can improve this method by a parallelized distributed A-NTF algorithm. Besides, we can use a more space and time efficient tensor factorization method (Wang et al. 2015) which sacrifices a certain amount of accuracy in exchange for the speed-up computation. Finally, we intend to improve our visual analytics system so that it can facilitate intra-regional car-sharing sites.

## 8 Conclusion

This paper proposes a novel visual analytics system for an insightful exploration of urban functional zones based on augmented nonnegative tensor factorization using human mobility semantics and inherent location information. We partition the territory into region units based on adaptive blue noise sampling method, which outperforms grid-based partition methods, extract inherent location information as a prior knowledge from POI data, model taxi OD data as a three-dimensional tensor and introduce a novel A-NTF method simultaneously combined with mobility semantics and location information to effectively identify urban functional zones. And then we design a set of visualization views to support a comprehensive understanding and interpretation the urban functional zones. Two case studies and expert interviews demonstrate the system. Positive feedback and in-depth insights show the effectiveness and usability of our system.

**Acknowledgements** The authors thank anonymous reviewers for their valuable comments. This research was supported by National Natural Science Foundation of China (U1909204) and the key fund of Hunan provincial education department (18A001). Wei Chen is Supported by National Natural Science Foundation of China (61772456, U1609217).

## References

- Adrienko N, Adrienko G (2011) Spatial generalization and aggregation of massive movement data. *IEEE Trans Visual Comput Graph* 17(2):205–219
- Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. *J Mach Learn Res* 3:993–1022
- Bridson R (2007) Fast Poisson disk sampling in arbitrary dimensions, p 22
- Cai L, Jun X, Liu J, Ma T, Pei T, Zhou C (2019) Sensing multiple semantics of urban space from crowdsourcing positioning data. *Cities* 93:31–42
- Cao N, Lin C, Qiuhan Zhu Y-R, Lin XT, Wen X (2018) Voila: visual anomaly detection and monitoring with streaming spatiotemporal data. *IEEE Trans Vis Comput Graph* 24(1):23–33
- Carroll JD, Chang J-J (1970) Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart–Young decomposition. *Psychometrika* 35(3):283–319
- Chen W, Huang Z, Feiran W, Zhu M, Guan H, Maciejewski R (2018) VAUD: a visual analysis approach for exploring spatio-temporal urban data. *IEEE Trans Vis Comput Graph* 24(9):2636–2648
- Cichocki A, Zdunek R, Phan AH, Amari SI (2009) Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation. Wiley, Hoboken
- Deng JS, Wang K, Hong Y, Qi JG (2009) Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanization. *Landsc Urban Plan* 92(3–4):187–198
- Deng Z, Weng D, Chen J, Liu R, Wang Z, Bao J, Zheng Y, Wu Y (2020) Airvis: visual analytics of air pollution propagation. *IEEE Trans Vis Comput Graph* 26:800–810
- Di Lorenzo G, Sbodio M, Calabrese F, Berlingerio M, Pinelli F, Nair R (2016) Allaboard: visual exploration of cellphone mobility data to optimise public transport. *IEEE Trans Vis Comput Graph* 22(2):1036–1050
- Ferreira N, Poco J, Vo HT, Freire J, Silva CT (2013) Visual exploration of big spatio-temporal urban data: a study of New York city taxi trips. *IEEE Trans Vis Comput Graph* 19(12):2149–2158
- Gao S, Janowicz K, Couclelis H (2017) Extracting urban functional regions from points of interest and human activities on location-based social networks. *Trans GIS* 21:446–467
- Harshman RA et al (1970) Foundations of the PARAFAC procedure: models and conditions for an “explanatory” multimodal factor analysis, pp. 1–84
- Hitchcock FL (1927) The expression of a tensor or a polyadic as a sum of products. *J Math Phys* 6:164–189
- Jolliffe IT (1986) Principal component analysis. Springer-Verlag, New York
- Karlsson C (2007) Clusters, functional regions and cluster policies. *JIBS and CESIS Electronic Working Paper Series* (84), 03
- Kim Y-D, Choi S (2007) Nonnegative tucker decomposition. In: 2007 IEEE conference on computer vision and pattern recognition, pp 1–8. IEEE
- Klema V, Laub AJ (1980) The singular value decomposition: its computation and some applications. *IEEE Trans Autom Control* 25(2):164–176
- Kruskal JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29(1):1–27
- Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. *Nature* 401(6755):788–791
- Lee DD, Seung HS (2000) Algorithms for non-negative matrix factorization, pp 556–562
- Levin J (1965) Three-mode factor analysis. *Psychol Bull* 64(6):442–452
- Liu D, Weng D, Li Y, Jie Bao Yu, Zheng HQ, Yingcai W (2017) Smartadp: visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE Trans Vis Comput Graph* 23(1):1–10
- Liu D, Panpan X, Ren L (2019a) TPFlow: progressive partition and multidimensional pattern extraction for large-scale spatio-temporal data analysis. *IEEE Trans Vis Comput Graph* 25(1):1–11
- Liu L, Zhang H, Liu J, Man J (2019b) Visual analysis of traffic data via spatio-temporal graphs and interactive topic modeling. *J Vis* 22(1):141–160
- Lv P, Zhang Z, Li C, Guo Y, Zhou B, Mingliang X (2019) Crowd behavior evolution with emotional contagion in political rallies. *IEEE Trans Comput Soc Syst* 6(2):377–386



- Mingliang X, Ge Z, Jiang X, Cui G, Lv P, Zhou B, Changsheng X (2019) Depth information guided crowd counting for complex crowd scenes. *Pattern Recogn Lett* 125:563–569
- Salton G, Yang C (1973) On the specification of term values in automatic indexing. *J Doc* 29(4):351–372
- Salton G, Yang CS, Yu CT (1975) A theory of term importance in automatic text analysis. *J Am Soc Inf Sci* 26(1):33–44
- Seto KC, Fragkias M (2005) Quantifying spatiotemporal patterns of urban land-use change in four cities of china with time series landscape metrics. *Landsc Ecol* 20(7):871–888
- Sun L, Axhausen KW (2016) Understanding urban mobility patterns with a probabilistic tensor factorization framework. *Transp Res Part B-Methodol* 91:511–524
- Tucker LR (1963) Implications of factor analysis of three-way matrices for measurement of change. *Probl Meas Change* 15:122–137
- Tucker LR (1966) Some mathematical notes on three-mode factor analysis. *Psychometrika* 31(3):279–311
- Tucker LR et al (1964) The extension of factor analysis to three-dimensional matrices. In: Frederiksen N, Gulliksen H (eds) *Contributions to mathematical psychology*. Holt, Rinehart & Winston, New York, p 110119
- Van Der Maaten L, Hinton GE (2008) Visualizing data using T-SNE. *J Mach Learn Res* 9:2579–2605
- Vatsavai RR, Bright E, Varun C, Budhendra B, Cheriyyadat A, Grasser J (2011) Machine learning approaches for high-resolution urban land cover classification: a comparative study. In: *Proceedings of the 2nd international conference on computing for geospatial research & applications*, p 11. ACM
- Von Landesberger T, Brodtkorb F, Roskosch P, Andrienko N, Andrienko G, Kerren A (2016) Mobilitygraphs: visual analysis of mass mobility dynamics via spatio-temporal graphs and clustering. *IEEE Trans Vis Comput Graph* 22(1):11–20
- Wang Y, Tung HY, Smola AJ, Anandkumar A (2015) Fast and guaranteed tensor decomposition via sketching, pp 991–999
- Wang H, Mingliang X, Zhu F, Deng Z, Li Y, Zhou B (2018) Shadow traffic: a unified model for abnormal traffic behavior simulation. *Comput Graph* 70:235–241
- Wu W, Zheng Y, Cao N, Zeng H, Ni B, Qu H, Ni LM (2017) MobiSeg: interactive region segmentation using heterogeneous mobility data. In: *2017 IEEE Pacific visualization symposium (PacificVis)*, pp 91–100. IEEE
- Xu M, Wu Y, Ye Y, Farkas I, Jiang H, Deng Z (2015) Collective crowd formation transform with mutual information-based runtime feedback. *Comput Graph Forum* 34(1):60–73
- Xu M, Li C, Lv P, Lin N, Hou R, Zhou B (2017) An efficient method of crowd aggregation computation in public areas. *IEEE Trans Circuits Syst Video Technol* 28(10):2814–2825
- Yao Z, Fu Y, Liu B, Hu W, Xiong H (2018) Representing urban functions through zone embedding with human mobility patterns. In: *IJCAI*, pp 3919–3925
- Yingcai W, Chen Z, Sun G, Xie X, Cao N, Liu S, Cui W (2018) StreamExplorer: a multi-stage system for visually exploring events in social streams. *IEEE Trans Vis Comput Graph* 24(10):2758–2772
- Yuan J, Zheng Y, Xie X (2012) Discovering regions of different functions in a city using human mobility and POIs, pp 186–194
- Yuan NJ, Zheng Y, Xie X, Wang Y, Zheng K, Xiong H (2015) Discovering urban functional zones using latent activity trajectories. *IEEE Trans Knowl Data Eng* 27(3):712–725
- Zhou Z, Jiajun Yu, Guo Z, Liu Y (2018) Visual exploration of urban functions via spatio-temporal taxi OD data. *J Vis Lang Comput* 48:169–177
- Zhu M, Chen W, Xia J, Ma Y, Zhang Y, Luo Y, Huang Z, Liu L (2019) Location2vec: a situation-aware representation for visual exploration of urban locations. *IEEE Trans Intell Transp Syst* 20(10):3981–3990