Discovering urban functional zones from biased and sparse points of interests and sparse human activities

Wen Tang [a],[*],[1], Alireza Chakeri [b], Hamid Krim [a]

[a] *Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27606, USA*
[b] *Valassis Digital, 3020 Carrington Mill Blvd, Morrisville, NC 27560, USA*

ARTICLE INFO

ABSTRACT

With rapid development of socio-economics, the task of discovering functional zones becomes critical to better understand the interactions between social activities and spatial locations. In this paper, we propose a framework to discover the real functional zones from the biased and extremely sparse Point of Interests (POIs). To cope with the bias and sparsity of POIs, the unbiased inner influences between spatial locations and human activities are introduced to learn a balanced and dense latent region representation. In addition, a spatial location based clustering method is also included to enrich the spatial information for latent region representation and enhance the region functionality consistency for the fine-grained region segmentation. Moreover, to properly annotate the various and fine-grained region functionalities, we estimate the functionality of the regions and rank them by the differences between the normalized POI distributions to reduce the inconsistency caused by the fine-grained segmentation. Thus, our whole framework is able to properly address the biased categories in sparse POI data and explore the true functional zones with a fine-grained level. To validate the proposed framework, a case study is evaluated by using very large real-world users GPS and POIs data from city of Raleigh. The results demonstrate that the proposed framework can better identify functional zones than the benchmarks, and, therefore, enhance understanding of urban structures with a finer granularity under practical conditions.

1. Introduction

The development of the network infrastructures and cellphone hardwares led to massive amount of mobile data with geographical information. In order to leverage the knowledge of such huge mobile data, it is critical to develop a digital map and its functional zones. Functional zones not only help individuals to understand a complex city, but also suggest a scientific urban planning. Additionally, functional zones are also useful for business sites management and tourism pattern discoveries, since they reveal the intention of human activity trajectories. It also answers the key questions of where and why people go visiting. On the other hand, the urbanization and modern civilization recently occur at an unprecedented rate. Hence, the task of functional zones discovery becomes essential in order to understand the interactions between social characters and spatial locations.

To address this problem, two main strategies have recently been developed in the literature. One is to reveal and visualize the complex spatial interactions between different functional zones (Guo, 2009; Yan & Thill, 2009), where any spatial interaction is quantified on the basis of the trips associated with a location and a timestamp of origin/destination. Zhong et al. (2014) defined a weighted network, where nodes are urban areas, edges are travel records and their weights are the number of trips. They then detected the graph centralities and clustered graph community structures to obtain the hubs, centers and borders. In Demsar et al. (2014) and Liu et al. (2015), the authors used community detection methods to identify the functional regions. Peng et al. (2012) applied Non-negative Matrix Factorization (NMF) to estimate different patterns and clustered them into areas. However, the quantitative analysis of spatial interactions did not carry enough information to explore the functional zones.

Another strategy is to introduce POIs to jointly combine them with the spatial interactions in order to discover functional zones (Gao et al., 2017; Liu et al., 2015; Wang et al., 2018; Yuan et al., 2018, 2015). The authors in Gao et al. (2017) and Yuan et al. (2018, 2015) applied the Latent Dirichlet allocation (LDA) based topic model to POIs and human activity behaviors, such as, user check-in data and taxi trajectories. Han et al. (2015) clustered the bus smart card data and identified different

---

* Corresponding author.
 E-mail addresses: wtang6@ncsu.edu (W. Tang), chakeri@mail.usf.edu (A. Chakeri), ahk@ncsu.edu (H. Krim).
 [1] Student.

Latent Region Represenation Learning

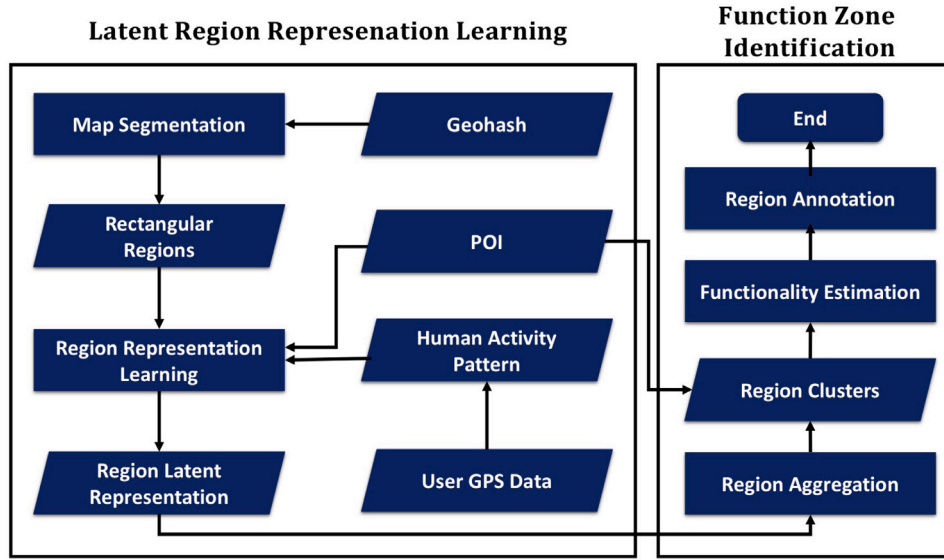**Function Zone Identification**

Fig. 1. Framework of functional zones revealing.

functional zones using POIs. Wang et al. (2018) used NMF method in Boarnet and Crane (2001) to identify functional regions based on POI data and taxi trajectory data. They then quantitatively analyze the spatio-temporal information for annotation.

However, all above methods neither leverage a real intrinsic structure between POIs and human activity patterns, nor deal with extremely sparse and biased real-world data. Additionally, the spatial information is also failed to be considered when the regions are aggregated. Generally, the methods in the literature assumed that the data follow some fixed probability distributions. However, our real-world data is highly sparse and less than 1% of the regions carry POI information. This is insufficient to estimate such parameters of the distributions. Also, compared to their POIs where they are generally unbiased and contain all different categories, in our dataset, only the commercial related POIs are provided. In order to cope with such biased and sparse information, we propose a novel functional zone identification framework to fuse the information between POI and human activity patterns. Instead of using probability assumptions, we introduce and transfer an intrinsic interaction structure between POIs and human activity patterns that was discovered in Wang et al. (2017) with unbiased data to bring the extra knowledge to alleviate the biased effects, and cooperates with matrix factorization to address the high sparsity, without any non-negative constraints. Moreover, the prior spatial knowledge is also used for region aggregation to confront against the bias issue. A novel optimization method is also proposed to learn the latent regions representation based on such intrinsic interaction structure, which is applicable to biased/unbiased and sparse/non-sparse data. To reduce the bias influence, CRF is applied to functional zones clustering to preserve both spatial information and their discriminative representations. The difference of normalized POI is then computed for functional zones annotation to emphasize the significance of different functionalities. The very large real-world users GPS and POIs data of city of Raleigh are used to evaluate our framework and demonstrate a better functional zones identification and a finer granularity than the benchmarks. The proposed framework is summarized in Fig. 1.

Our main contributions can be summarized as:

(1) A novel latent region representation learning method is proposed based on intrinsic interaction of unbiased POIs and human activity patterns to overcome the biased data.
(2) To exploit the prior spatial information and alleviate both bias effects and the imbalance POIs of map segmentation, CRF is used

as an unsupervised clustering method, which is regularized by the spatial relationship.
(3) A novel and efficient functional zone detection framework is proposed for both biased/unbiased and sparse/non-sparse data. This achieves a better functional zones identification and a finer granularity than the benchmarks based on real data.

The rest of the paper is organized as follows: Section 2 introduces the preliminaries and the details on the latent region representation learning procedure. Functional regions aggregation and identification methods based on function estimation are discussed in Section 3. The results are shown in Section 4. And more discussion of properly using our method is discussed in Section 5. Finally, Section 6 concludes the paper.

2. Latent region representation learning

In order to discover different functional zones, we segment our city into small rectangle regions, which are considered as the basic region units. Based on these region units, a novel algorithm is proposed to learn latent semantic representation of each region unit that captures the interaction information between POIs and human activities.

2.1. Map segmentation

To speed up the map segmentation procedure, the well-known Geohash geocoding system is adopted to segment the map into equal-sized regions. geohashes express the girded regions by short alphanumeric strings. The greater precision is obtained with a longer string and a larger Geohash level. For example, two $156.5 \text{ km} \times 156 \text{ km}$ rectangle areas of Geohash level 3 can perfectly cover the city of Raleigh, whereas it can also be accurately represented by 64 $39.1 \text{ km} \times 19.5 \text{ km}$ regions of Geohash level 4 as shown in Fig. 2. In total, there are 12 levels of Geohashes to present different area granularity. Geohash level 6 is used in our experiments.

2.2. Region feature learning

2.2.1. Preliminaries

Definition 1 (*Trajectory*). A trajectory of a user is a sequence of time stamped location points, i.e., $Traj = (lp_1, \ldots, lp_n)$, where $lp_i = (lat, lon, t)$, represents the latitude and longitude at time t_i, $\forall i = 1, \ldots n$.

Fig. 2. Examples of Geohash level 3 and Geohash 4 segmentation. 2 big Geohash level 3 regions cover 64 Geohash level 4 regions.

Definition 2 (*Distance and Time Period*). The geospatial distance between two points lp_i and lp_j is defined as $D(lp_i, lp_j)$. Time period $TP(lp_i, lp_j) = |lp_i.t - lp_j.t|$ is the time interval between two points.

Definition 3 (*Human Activity*). A human activity h_i of a user is defined such that the user stays in a geographical region within a given Dr distance over a given time period Tr. h_i^j is decided by a set of consecutive points $L = \{lp_k, lp_{k+1}, \ldots, lp_m\}$. In particular, when for all $k < i \le m$, $D(lp_k, lp_i) \le Dr$, $D(lp_k, lp_{m+1}) > Dr$ and $TP(lp_k, lp_i) \le Tr$, $h_i = (lat, lon, t_a, t_l)$, where $t_a = lp_k.t$ is the arrival time and $t_l = lp_m.t$ is the corresponding leaving time. Also, $h_i.lat = \frac{1}{|L|}\sum_{i=k}^{m} lp_i.lat$ and $h_i.lon = \frac{1}{|L|}\sum_{i=k}^{m} lp_i.lon$ are respectively the latitude and longitude of the human activity geospatial location.

Definition 4 (*Human Activity Trajectory*). A human activity trajectory is a sequence of human activities, i.e. $HT = \{h_1 \to h_2 \cdots \to h_q\}$ where h_i is a human activity.

Definition 5 (*Human Activity Information*). A human activity information A is a triplet consisting of the origin region R_o, the destination region R_d and departure time t_l or arrival time t_a. Leaving human activity information is defined as $A_L = (R_o, R_d, t_l)$, and arriving human activity information is then defined as $A_A = (R_o, R_d, t_a)$. For each human activity h_i in the human activity trajectory HT, a map f has been used to map the geospatial location to different Geohash regions, i.e., $R_i = f(h_i.lat, h_i.lon)$. The leaving and arriving time are respectively obtained by $t_l = h_i.t_l$ and $t_a = h_i.t_a$.

Definition 6 (*Bag-of-POIs*). Let assume that there are p different POI categories, and p_j shows the number of the jth POI category in the map. Then, the bag-of-POIs for the ith region unit is defined as $poi_i = (p_1^i, \ldots, p_p^i)$, $\forall i = 1, \ldots, p$.

Since the bag-of-POIs uniquely represents different functional zones, it can be used to represent each region to reserve the functional information and obtain the POI information matrix P as in the following:

Definition 7 (*POI Information Matrix*). Let each column of this matrix be the Bag-of-POIs of ith region unit, i.e.

$$P = \begin{bmatrix} poi_1^T & poi_2^T & \ldots & poi_p^T \end{bmatrix}.$$

Definition 8 (*Human Activity Pattern (HAP) Matrix*). Based on each human activity information $A_L = (R_o, R_d, t_l)$ and $A_A = (R_o, R_d, t_a)$, let each element of matrix $C_k = [c_{ij}^k]$ represents the number of A_L (or A_A)

from ith region unit to jth region unit, during the time period k. Then we define the HAP matrix as

$$T = \begin{bmatrix} C_1, C_2, \ldots, C_s \end{bmatrix}^T$$

Each time period is defined as one hour. 24 time periods are considered in one day. That is, $t_l \in [0:00, 1:00), \ldots, t_l \in [23:00, 24:00)$. In our experiment, both T based on A_L and A_A are concatenated together.

2.2.2. Latent region representation learning

In order to discover the functional zones, we need to explore the discriminative representations of various functional zones. One useful functional zone related information is POIs, as different POIs (such as fast food, coffee bar, shopping mall) attract individuals for different purposes. In addition, more POIs translates into higher probabilities for their corresponding functionalities. As a result, Bag-of-POIs were first introduced in our framework in a naive way.

However, POIs are not only the determining factor of the region functionalities. The region functionality is also decided by the intention of individuals trajectories. Chen et al. (2015) discovered that various region functional zones have their corresponding HAPs varies in different time periods. This implies that HAPs which reflect semantic meanings of Human intentions, another important factor for functional zones. HAPs is therefore, also included into our framework to enrich the knowledge of the region functional zones discovery.

In contrast to the conventional methods that proposed on the balanced and sufficient datasets, when the POIs data is very sparse, it is insufficient for small number of biased POIs to precisely estimate or learn the parameters of assumed distributions in Gao et al. (2017) and Yuan et al. (2018, 2015) and the variables in Liu et al. (2015) and Wang et al. (2018). We hence decompose the POI and region features into latent dense representations and propose an intrinsic interaction between the latent POIs and HAPs in our method. The intrinsic interaction that we proposed is inspired from Wang et al. (2017), as they revealed that 80% of HAPs can be successfully predicted by POI information. That is, POIs information can almost entirely reconstruct HAPs. Since this discovery was also fully data-driven and leveraged on the unbiased data, such intrinsic knowledge can be directly transferred into our framework to cope with the POI bias. Additionally, since region functionalities are critical to affect human travel behaviors and infers traffic pattern, a transformation between region representation and HAPs should be also incorporated.

Consequently, as shown in Fig. 3, the POI information matrix P is decomposed into two latent and dense spaces U and V. And HAP matrix T is transformed by matrix Q into a semantic latent space Z to reduce the dimensions and preserve useful information. According to the transferred intrinsic knowledge leveraged from unbiased data, the transformed HAP Z is also reconstructed by some important latent POI features U that are gleaned from the original POI information in a latent space. The coefficients matrix A that regularized by L_1 norm are used to select the important latent POI features in U for each region. Meanwhile, the latent HAP representation Z also influences the latent region representation V that is also derived from POI information in the same latent space. Besides, to avoid the lack of complete POI information, an indicator mask is also used to deal with the sparsity of POI matrix.

Hence, our final information fusion method for the latent region representation learning is formulated as follows:

$$
\arg\min_{\substack{U,V,Q \\ W,A,Z}} \frac{1}{2}\|I \circ (P - UV)\|_F^2 + \frac{\lambda_1}{2}\|QT - Z\|_F^2
$$
$$
+ \frac{\lambda_2}{2}\|Z - U^T A\|_F^2 + \lambda_3\|A\|_1 + \frac{\lambda_4}{2}\|V - WZ\|_F^2 \tag{1}
$$
$$
+ \frac{\lambda_5}{2}(\|U\|_F^2 + \|V\|_F^2 + \|Q\|_F^2 + \|W\|_F^2),
$$

where I is a indicator mask. $I_{i,j} = 1$ for non-empty entries and $I_{i,j} = 0$ for others, $\circ$ is an element-wise product. $\lambda_1, \lambda_2, \lambda_3, \lambda_4$, and λ_5 are tuning
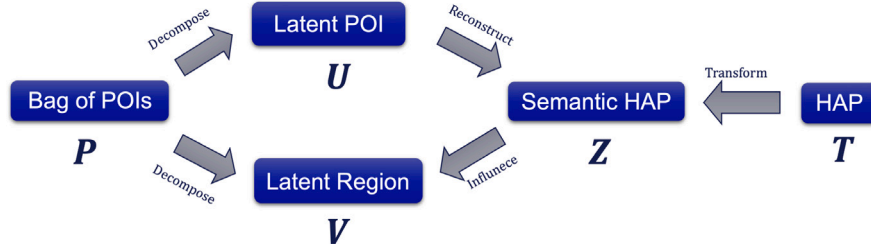
Fig. 3. Latent region representation learning model.

parameters. $P \in \mathbb{R}^{p \times r}$ is decomposed into two latent semantic matrices $U \in \mathbb{R}^{p \times k}$ and $V \in \mathbb{R}^{k \times r}$. $T \in \mathbb{R}^{q \times r}$ is transformed into its latent transition pattern matrix $Z \in \mathbb{R}^{k \times r}$ by a transform matrix $Q \in \mathbb{R}^{k \times q}$. The latent transition pattern Z receives the information from the latent POI features by using sparse coefficients $A \in \mathbb{R}^{p \times r}$ to choose a specific POI in U, and propagates the information back to the latent region feature V by $W \in \mathbb{R}^{k \times k}$ regression. The last term in Eq. (1) is the regularization for avoiding overfitting.

We, therefore, alternatively search one variable when others are fixed. The updating rule of each variable based on gradient descent are also summarized in Algorithm 1, where $\tau(\cdot)$ is the element-wise soft thresholding operator, a solution of L_1-norm and is used in many sparse learning (Beck & Teboulle, 2009; Tang et al., 2020, 2021, 2016, 2018, 2019a, 2019b).

Algorithm 1 Latent Region Representation Learning

Input: POI matrix P, HAP matrix T, and hyper-parameter $\lambda_1, \lambda_2, \lambda_3, \lambda_4,$ and λ_5

Output: Latent POI feature U, latent Region representation V, Sparse Coefficients A, transform matrix Q, latent transition pattern matrix Z, and regression matrix W

$t=1$;

while $t < maxIteration$ and $L_t - L_{t+1} > \varepsilon$ **do**

$\quad U_{t+1} = U_t - \alpha(-I \circ (P - UV)V^T - \lambda_2 A(Z - U^T A)^T + \lambda_5 U)$

$\quad V_{t+1} = V_t - \alpha(-U^T I \circ (P - UV) - \lambda_4(V - WZ) + \lambda_5 V)$

$\quad Q_{t+1} = Q_t - \alpha(\lambda_1(QT - Z)T^T + \lambda_5 Q)$

$\quad Z_{t+1} = Z_t - \alpha(-\lambda_1(QT - Z) + \lambda_2(Z - U^T A) - \lambda_4)W^T(V - WZ)$

$\quad A_{t+1} = \tau_{\alpha\lambda_3}\left(A_t - \alpha(-\lambda_2 U(Z - U^T A))\right)$

$\quad W_{t+1} = W_t - \alpha(-\lambda_4(V - WZ)Z^T + \lambda_5 W)$

$\quad \alpha = \rho\alpha$ % ρ is a decay factor

end while

return U, V, Q, Z, A, W

After solving the algorithm, the latent region representation V comprehensively fuses both information of POI and HAP and the transferred knowledge. The latent transition pattern matrix Z also contains the semantic meanings of HAP. Given U, V, Q, Z, A, W, we then can identify the functional zones based on the latent region representation V, a k-dimensional feature $(v_1^r, \ldots, v_k^r)$ for each region unit r.

3. Functional zone identification

3.1. Region aggregation

As the ground truth of region labels are not available at this moment, clustering is then naturally used to perform it to aggregate the regions. Regions in the same clusters should have similar functionalities, while those in different clusters regions should provide various different functionalities. K-means clustering method is a naive and usual way to cluster region units to functional zones. However, K-means does not consider the spatial information of each region. When the segmentation grids is very small, a big place may segmented into

different region units and carrying different number of POIs. Then the fine-grained region representations may be totally different from each other and lead to a worse clustering result by K-means. For instance, a shopping mall is segmented into 3 region units. One of them carry the majority of POIs of different stores, while another two of them only have a few stores. There is then a huge difference among POI features of those region units. Then these 3 region units are more likely clustered into 2 various clusters rather than be clustered into the same cluster. Therefore, if the spatial information is not leveraged, K-means clustering based on the region representation is not enough for the finer granularity. When the segmentation is fine-grained, many functional zones are partitioned into different small parts with different kinds of POIs and various numbers of POIs, thus introducing more variance into the same functional zones. In order to confront against such imbalance that caused by the fine-grids, we take the spatial location of each region units into account. In usual, when the granularity is fine, the 8 neighbor region units that surround this region unit should be more probably share the same cluster. Hence, the spatial location information should also be utilized in the clustering as well to improve the consistency of the clustered region units. As the region representation and neighbor clusters will both influence the current region clustered result, the current region unit representation and their neighbors clusters are connected as a small random field that conditional independent with other region units. Conditional Random Field (CRF) method is thus involved to cluster the region units, because CRF method aims to maximize a posterior (MAP) inference over each region unit representations and meanwhile also to maximize each cluster agreement between similar region units and neighbors to induce more smoothness (Krähenbühl & Koltun, 2011). Each region unit can be regarded as a k-dimensional data point that is generated by a probability distribution, such as Gaussian. And a cluster agreement penalty is also jointly used to calculate the cluster generation probability of different region units. In our paper, the Gaussian distribution is regarded as the generative distribution of each region unit, and only 8 neighbors of each region units are used in CRF to penalize the cluster agreement. The process of CRF can be formulated as follows:

(1) Let functional Zone $\mathcal{L} = \{l_1, \ldots, l_c\}$ have c different functional zones, and random variable $Y = \{y_1, \ldots, y_r\}$ denote region-unit-level clustering, such as $\{cluster_1, \ldots, cluster_r\}$ that are un-annotated and finally assigned to each region unit, and random variable $R = \{r_1, \ldots, r_r\}$ denote a feature of each region. In our experiment, the random variable R is the latent region feature V. Then the conditional probability of cluster $Y = l_K$ is assumed to follow a normal distribution:

$$P(Y = l_K | R) = N(\mu_{l_K}, \Sigma_{l_K}), \qquad (2)$$

(2) For each region unit R_i, the joint probability of cluster $y_i \in \mathcal{L}^r$ in CRF is given by:

$$P(y_i, y_j) = exp\{-\sum_{(i,j) \in N} V(i,j)\},$$

$$V(i,j) = \beta\delta(y_i, y_j) = \begin{cases} \beta & \text{if } y_i = y_j \\ -\beta & \text{else,} \end{cases} \qquad (3)$$

where N is the neighbor set of region R_i.

(3) The final target function is formulated as follows:

$$\arg\max_{y \in \mathcal{L}^r} \ P(r, y) = \frac{1}{F} \Pi_i P(y_i | r_i) \Pi_{i,j} P(y_i, y_j), \tag{4}$$

where F is a normalization term.

The join probability in Eq. (4) can be solved by different approaches, such as Expectation–maximization (EM) algorithm in Dempster et al. (1977).

3.2. Functionality estimation

Due to the lack of the labels, after region aggregations, functionality estimation of each cluster is then another necessary step before annotating the functional zones. Once the estimated functionalities are measured, then the functional zones can be annotated based on these estimations. For the functionality estimation, the unified POI distribution of each cluster is supposed to be adequate to carry the functionality. However, the actual number of POIs is not the same in each region units. And the number of region units in each cluster is also different. Thus, to emphasize the significant functionality of different POIs in each cluster, the maximum normalized POI distributions is adopted for comparison, which reflects the importance proportion of contributions of different POIs. For the sth cluster, the vector of POI distribution is hence formulated as follows:

$$PR^s = \left(\frac{\sum_{i \in S} p_1^i}{|S|}, \dots, \frac{\sum_{i \in S} p_p^i}{|S|} \right), \tag{5}$$

where S is a set of region units whose labels of the functional zone are l_s, and $|S|$ is the number of elements in the set S. The normalized distribution then is:

$$NPR^s = \frac{PR^s}{\max(PR^s)}. \tag{6}$$

3.3. Region annotation

Finally, region annotation is to associate the real functionalities with clusters. However, it is far from enough to name a functional zone by only using the frequency of POIs. Because many functionalities may appear in the same cluster, which leads to a hard region annotation. In our method, we use the difference of functionality to assign some semantic terms to each functional zone to understand its real function. The following conditions is utilized to annotate different functional zones:

(1) **The difference in the normalized POI distribution.** The difference G_s of sth functional zone is calculated as follows,

$$G_s = (|\mathcal{L}| - 1) NPR^s - \sum_{j \in \mathcal{L}, j \neq s} NPR^j, \tag{7}$$

where $|\mathcal{L}|$ is the number of functional zones. The positive elements of G_s is the significant POIs in the functional zone s.

(2) **Land-use maps of government.** Compared with land use or zoning map, it is more accurate for us to understand different functional zones we obtained. For example, several big industrial regions can help us fast locate such functional zones.

(3) **Human-labeling.** As some information may lack in the POI information, people know more about well-known or common places that never show up in the POI information, such as, the label of universities, which is missing in our experiments.

All the examples of the metrics of the functionality estimation and the region annotation are shown in Fig. 4. As shown in Fig. 4, it is easy to show that the magnitude of the significant functionalities changed for different clusters after normalization by comparing (a) PR and (b) NPR. To show the role of (c) G_s, the example of G_1 can directly show its improvement of significance of important POIs. In (c), after calculating the difference of G_1, and only $G_1 > 0$ is considered for annotation, it is clear that financial service (the point in the black circle) is not so
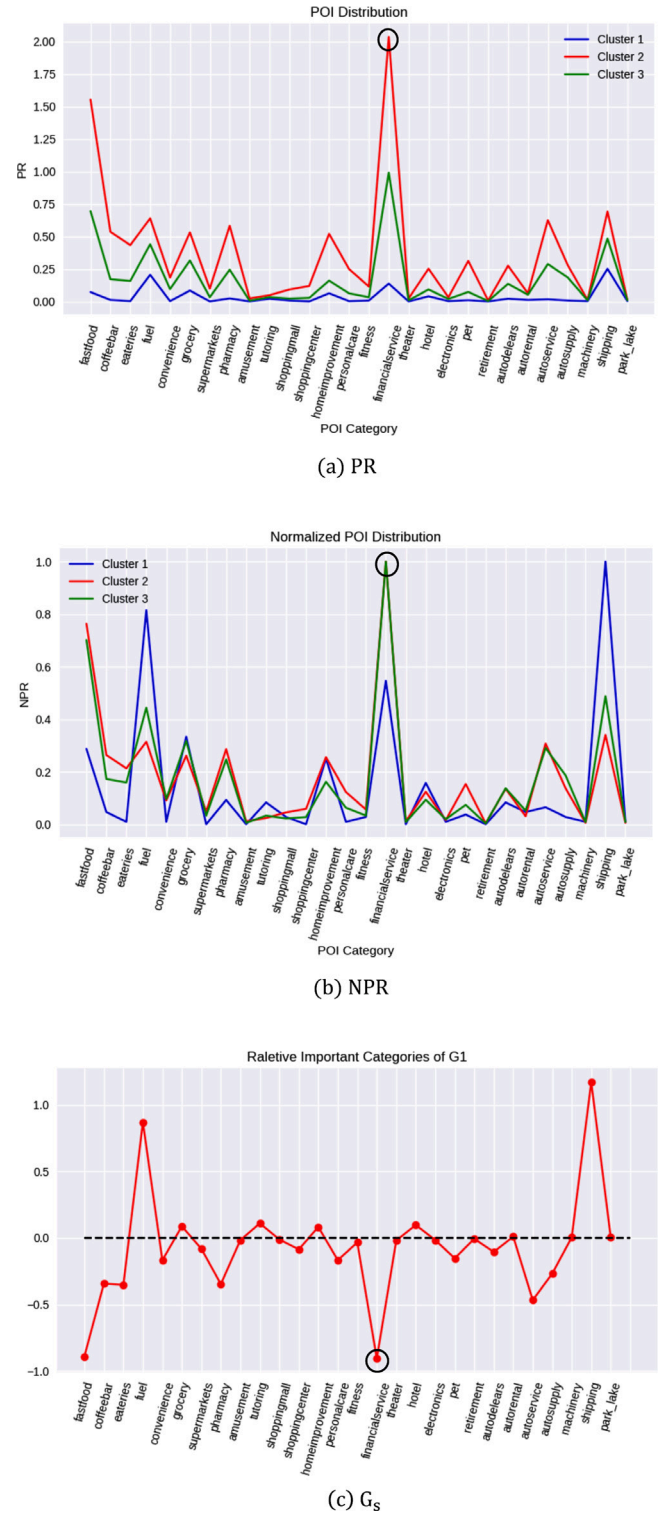

(a) PR

(b) NPR

(c) G_s

Fig. 4. Example of PR^s, NPR^s and G_s.

significant for Cluster 1 as before, when comparing with Cluster 2 and Cluster 3. Because, although there are many financial services in Cluster 1, it also appears many times in Cluster 2 and Cluster 3. Thus actually, it is not distinguished among these clusters. Therefore, the difference G_s could be used to improve the significance and distinction of the POIs for functionality annotation.

Table 1
POI category.

Index	POI category	Index	POI category
1	Fast food	15	Fitness
2	Coffee bar	16	Financial service
3	Eateries	17	Theater
4	Fuel	18	Hotel
5	Convenience store	19	Electronics store
6	Grocery	20	Pets/veterinary
7	Supermarkets	21	Retirement
8	Pharmacy	22	Auto dealers
9	Amusement	23	Auto rental
10	Tutoring school	24	Auto service
11	Shopping mall	25	Auto supply
12	Shopping center	26	Machinery
13	Home improvement	27	Shipping store
14	Personal care	28	Park/lake(camping site)

Table 2
POI statistics.

POI dataset	Statistics
#POI	6725
#Regions with POIs/#Segmented regions	1233/16,384
Sparsity of POI matrix	99.11%

User GPS dataset	Statistics
#Users	1,026,726
#HAP	27,930,027
#Regions with HAP/#Segmented regions	16,321/16,384
Sparsity of HAP matrix	99.97%

4. Experiments and results

4.1. Experiment settings

Our experiments are evaluated on the following datasets:

(1) Raleigh POI dataset is updated starting 2014. The category of POI has be listed in Table 1. There are 6725 POIs and 1233 regions with POI information. The sparsity of POI matrix P is 99.11%.

(2) User GPS dataset is collected during the workdays of March 15th to May 20th in 2018. In this dataset, there are 1,026,726 users GPS records that located in Raleigh. They are then transformed to 27,930,027 Human Activity Information. The number of regions with HAP is 16,321. The sparsity of HAP matrix T is 99.97%.

(3) Region Segmentation: We transformed Raleigh Map in to 1200 m × 609.4 m small rectangle regions that coding by Geohash level 6. We totally obtain 16,384 segmented regions.

The statistics of all three datasets are listed in Table 2.

4.2. Benchmarks

In our experiments, the dimension of the latent region representation is set to 10.

In order to avoid over-weighted the common but unimportant POIs, TF-IDF is applied to the POI matrix to adjust POIs. For example, people may go to some shopping center because of its cinema there, but they may go for dinner first. And the number of restaurants usually is much higher than the number of cinema. The region functionality is then apparently dominated by the restaurants if only based on the frequency of POIs. However, the true region functionality maybe the cinema. To avoid such over-weighted issues, TF-IDF is then used to adjust and balance the frequency and distinguish of POIs. TF-IDF is applied by following the settings in Wang et al. (2017) and Yuan et al. (2015). We then compared our results with the following benchmarks, and all benchmark methods build up on the POI matrix prepossessed by TF-IDF:

(1) POI Distribution: The POI distribution based method is to directly apply K-means and CRF clustering methods on the POI distribution features.

(2) Collaborative POI: In addition, to considering latent structure and avoiding missing data in TF-IDF based Bag-of-POIs, collaborative filtering methods have been used to obtain latent semantic representation of POIs (Yuan et al., 2015). In our comparison, collaborative POI is obtained by SVD decomposition and represented by the eigenvectors that are associated with the top tth maximum eigenvalues. Then, K-means and CRF clustering methods are respectively employed based on collaborative POI features.

(3) Topic Modeling: By regarding the regions as a bunch of documents and the POI is the words of each document, then the region functionalities actually play the same of the topics in the documents. As the topics determine the words generation in each documents, the region functionality also determines the POI generation in each region. In addition, the human activities are the prior knowledge of the generating distribution of the POIs, as the human activities would also influence the POIs. Based on the each topic of the documents, the documents can be clustered in the similar topic clusters. In the same way, region functional zones can be clustered by those region functionalities that computed as the topics. As a typical topic model, LDA, therefore, is used to discover the region functionalities based on the HAP and then cluster functional zones by K-means. For LDA model, we choose 10 topics with 1000 iterations, following the settings in Yuan et al. (2018).

In our experiment, we learn our latent region representation by Algorithm 1. Two clustering approaches: K-means and CRF are utilized on our latent region representations to cluster S different functional zones. We also cluster our latent semantic HAP by just using K-means to compare with Topic Modeling method (Yuan et al., 2018).

To verify our results, we compare our results with land use map of Raleigh and check whether most famous regions are assigned into their proper functional zones.

4.3. Results

4.3.1. Functional zones comparison

We aggregate 4 functional zones by hereabove methods for Raleigh city and indicated by different colors, as shown in Fig. 5(g). In order to better compare the different methods, we zoom in on the results of major parts of Raleigh and demonstrate them in Figs. 5(a)–5(f). It is noting that different colors in different figures may indicate different functional zones. The results shown in Figs. 5(a), 5(b), 5(d) and 5(e) are only based on clustering of the POI information, while the results shown in Figs. 5(c) and 5(f) are based on clustering the fusion features of both POI and HAP information. Figs. 5(a)–5(c) are all based on K-means clustering, while Figs. 5(d)–5(f) are all based on CRF clustering. Figs. 5(h) and 5(i) in the third row figures are based on the clustering of only the HAP information.

A. Quantitative Comparison

To quantitative evaluate our framework, the information retention and the unbiased degree are introduced as two direction measurements. As the dataset is extremely sparse, to avoid loss such precious information, the information retention is used to evaluate how much original dataset information is retained after the clustering. In another hand, since POIs are biased and only contain the commercial POIs, then most regions are easily clustered into the same clusters when their functionalities may totally different. Therefore, to measure how much information balanced in our dataset to avoid the bias of POIs, the unbiased degree is included to measure the how balanced the clusters are.

Like second and third-tier cities, Raleigh contains very few POIs. There are only 7.52% regions with POI information in our experiments,
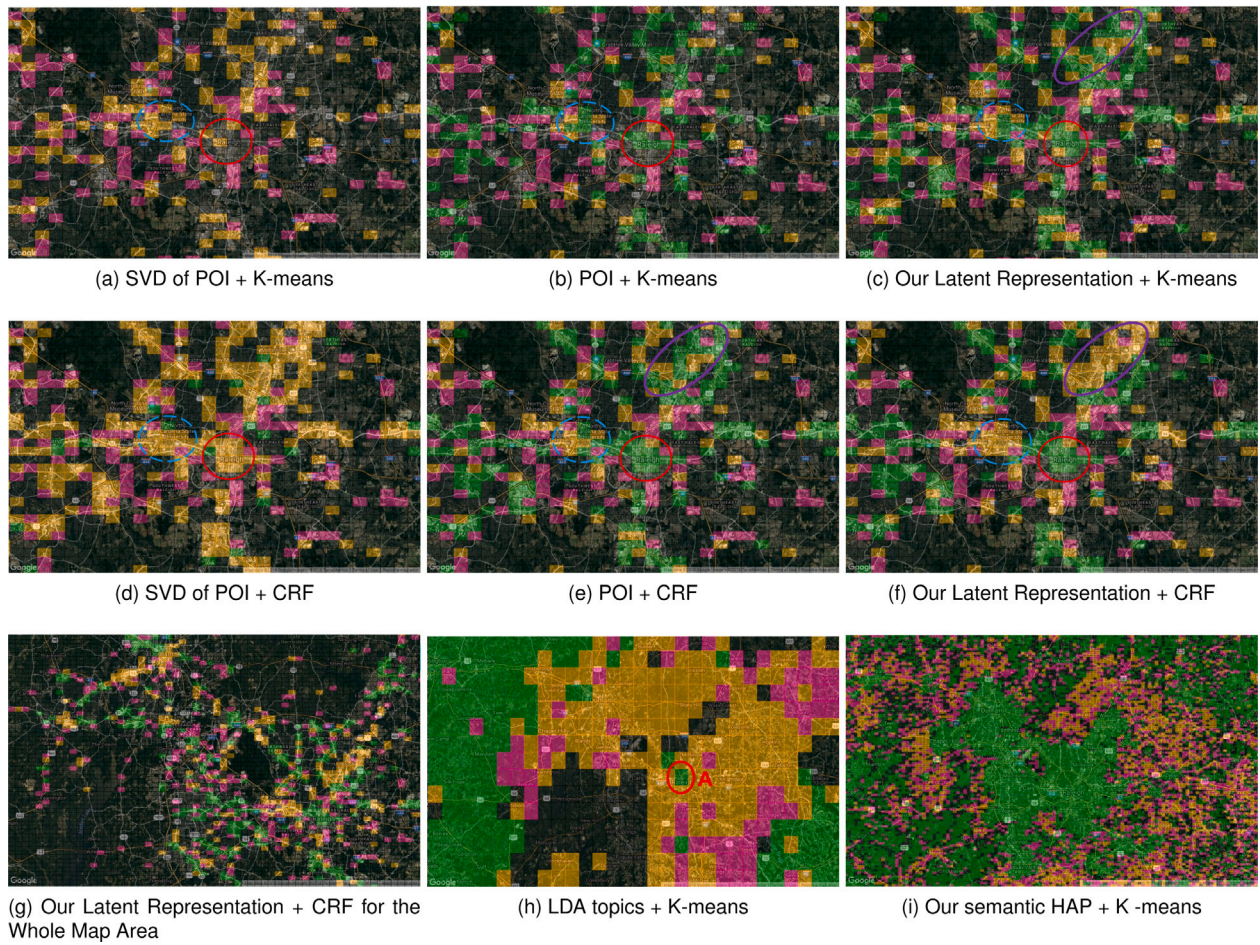
(a) SVD of POI + K-means

(b) POI + K-means

(c) Our Latent Representation + K-means

(d) SVD of POI + CRF

(e) POI + CRF

(f) Our Latent Representation + CRF

(g) Our Latent Representation + CRF for the Whole Map Area

(h) LDA topics + K-means

(i) Our semantic HAP + K -means

Fig. 5. Comparison of different region representations.

Table 3
Information retention.

	Our feature w. CRF	Our feature w. K-means	Collaborative POI w. K-means	POI w. K-means
Before clustering	7.52%	7.52%	7.52%	7.52%
After clustering	7.52%	7.52%	5.27%	5.97%
Information retention	**100%**	**100%**	70.08%	79.39%

and the sparsity of POI matrix is up to 99.11%. The black clusters in Figs. 5(a)–5(f) represent missing the POI information group. Since this group is really large, other undistinguished features are easily clustered into this missing POI information group. As shown in Fig. 5(b), only using POI distributions is not discriminative enough, many regions are regarded similar to the missing POI information one. The detailed information retention are listed in Table 3.

For the benchmark (1), after directly using k-means clustering on POI distribution, the regions containing POI information significantly decreases from 7.52% to 5.97%, which misses 20.61% information. In Fig. 5(b), it shows that only the top parts of downtown Raleigh that circled in red are not black and carrying POI information.

For the benchmark (2), in order to reduce the sparsity and increase the latent semantic meanings of POI information, SVD is then applied on POI matrix. However, with limited data in POI matrix, SVD decomposition cannot explore more latent structure. The result in Fig. 5(a) is even worse than just using the POI distribution, i.e., the benchmark (1). It also shows that most regions are clustered into 3 groups (the black, yellow and pink ones), when 4 groups should be clustered. Additionally, only 5.27% regions contain the POI information, while 5.97% regions preserve the POI information for the benchmark (1). This

loses 29.92% of all the information, which is mainly caused by a lack of discriminative features in Collaborative POI.

In Fig. 5(c), our latent region representations are fused both POI and HAP information. It shows that such latent region representations are distinguished enough to preserve all of the 7.52% POI information regions, i.e., 100% information is preserved. Most of the regions with similar functionalities are in the same clusters. For example, the major parts of the regions in the red circle are assigned a green color, which actually is the downtown area of Raleigh. And half of the regions in the blue circle are assigned an orange color, where the university area should be located. Our latent region representation is more discriminative than the other benchmarks.

To preform the unbiased degree of the clusters, $Unbiased\ Degree = 1 - silhouette\ score$, where silhouette score proposed in Rousseeuw (1987) is a measure of the cluster compactness. In contract to the general case that the lower decentralization degree, the worse, in our case, the higher decentralization degree, the better. Because that means biased POI are balanced when the knowledge is propagating between the unbiased intrinsic interaction between POIs and HAPs.

As shown in Table 4, using the simple K-means clustering on the distribution of POI obtains the lowest unbiased degree, because of the

Table 4
Unbiased degree.

Our feature w. CRF	Our feature w. K-means	Collaborative POI w. K-means	POI w. K-means
6.01%	**5.73%**	2.61%	2.28%

dominance of the missing POI regions. When the Collaborative POI feature is used, the decentralization degree is slightly increased 0.33% because its SVD decomposition seeks the latent representation of POIs and recovers the unbiased representations to some extent. However, when using our proposed feature even for K-means clustering, the decentralization degree is significantly increased by 3.12%, which is 2 times more than Collaborative POI features. That is, our proposed method did learn the unbiased information from the intrinsic structure of POIs and HAPs. And later qualitative comparison also show that our balanced region representations do reveal the true region functionalities. Moreover, when using CRF for clustering, the functionality consistence is also balanced among fine-grained regions, which results in another 0.28% more increasing of the unbiased degree than the one of K-means clustering. That is, 2.3 times more than the result of the Collaborative POI and 2.6 times more than the result of POI.

Based the quantitative comparison, our proposed framework not only retains the 100% information of the data, but also balance the region representations and clusters for true functional zone discovery.

B. Qualitative Comparison

In addition to the quantitative comparison, we also show the functional zones on the map to compare the qualities of discovering the functional zones. The land-use map is also used for more accurate comparison.

Comparing with the results of K-means shown in Figs. 5(a)–5(c), the results of CRF in Figs. 5(d)–5(f) are more consistent. Most of the regions are assigned the same colors. The blocks of regions in the same color are more numerous than those of using K-means.

For the benchmark (2), although the CRF result based on Collaborative POI in Fig. 5(d) is much better than K-means one, most of the regions are in the same functional zones, which lose its meanings to discover the functional zones. For example, the functionality of the downtown part that is circled in red should be different from the university area that is circled in blue.

For the benchmark (1), the CRF result is based on POI distributions. As shown in Fig. 5(e), it has a more uniform clustering result than benchmark (2), which is more meaningful. But it is still not so consistent as the results of our latent region representation shown in Fig. 5(f). In Fig. 5(e), parts of the region functionalities in the university area circled as blue are the same as those in the downtown area circled as red. Also, there are many industries located in the purple circle, which should have similar functionality as the university area, especially for workdays. However these regions in the purple circle are assigned half green and half orange in Fig. 5(e). To look at the regions in both the blue and purple circles in Fig. 5(f), they are almost assigned to orange color. Therefore, our latent region representation is still discriminative than other benchmarks. Additionally, the CRF clustering of our latent region representations has a more consistent result than the K-means clustering does as shown in Figs. 5(c) and 5(f). A more detailed comparison of land-use map and the results of both K-means and CRF are shown in Fig. 6.

For the benchmark (3), since (Yuan et al., 2018) demonstrated that LDA model has a similar result of functional zones as the one of DMR model, and DMR model which takes the POI information into account only change few areas but spending more time and memory for training, it is fair to choose LDA to be the benchmark for comparison. As constrained memory limitation for LDA training, we use Geohash level 5 4900 m × 4900 m grids to segment the same city map and show its results in Fig. 5(h). As shown in Fig. 5(h), the result of LDA nicely indicates that the map is segmented into 4 different parts based
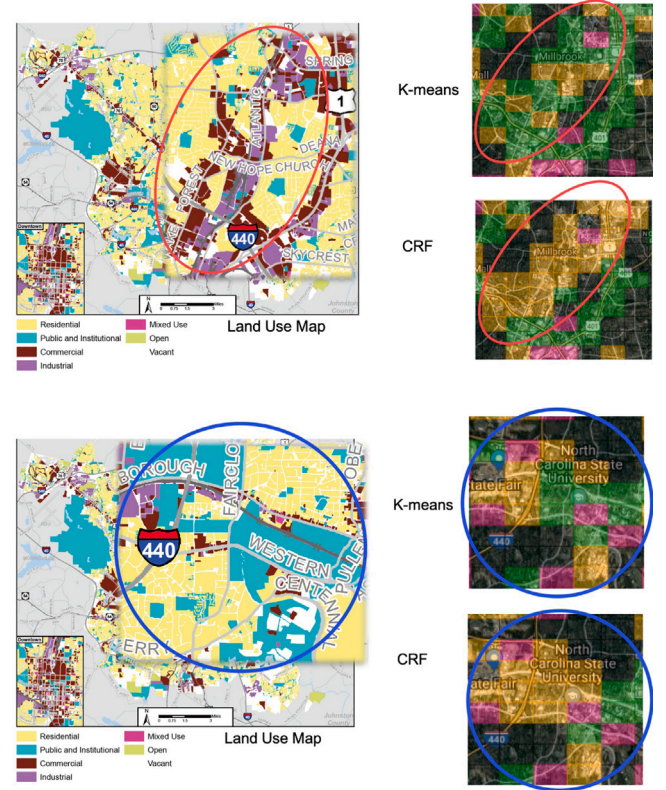
Fig. 6. Comparison of different clustering methods based on latent region representation.

on the spatial locations. However, there are still some not well clustered regions. For example, the green part that is circled in rad and labeled with "A" is the downtown area of Raleigh, and it should be a high-level-activity area. However, the activity level in other green areas in Fig. 5(h) is much lower than the ones of downtown Raleigh. That is, downtown Raleigh should not be assigned to the same group as other green areas are. Also, the orange area covers for the most part cities, while many vacant and nature places are also included in this orange area. For a fair comparison, we also used K-means to cluster our semantic HAP and show its results in Fig. 5(i). In Fig. 5(i), the center green areas are the major cities, such as, Raleigh, Durham, Chapel-Hill, Cary and Apex. The orange areas represent a middle-level activity area, and the pink areas denote a low-level-activity area. Our latent transition pattern Z is capable highlight how urban these areas are, while LDA still blurs the high-level-activity area with other level-activity areas. It worth nothing that, our proposed method is based on the Geohash level 6 (1 region unit of Geohash level 5 covers 32 region units of Geohash level 6), which has a finer granularity than LDA does.

Overall, using the CRF to cluster our latent region representations that fused both POI and HAP information achieves a superior performance than other approaches. The latent region representations are more discriminative. The CRF method achieves a more consistent clustering result. And our proposed framework supports a finer granularity with a practical implementation time.

4.3.2. Functional zones annotation

According to our functionality estimation, we calculate the normalized POI distribution NPR for each different functional zone. The difference G_s of each functional zone is then computed, and the components of $G_s > 0$ are listed by decreasing the significance in Table 5. Since areas and cities in our experiment are not modern cities, they do not have a complex city structure. We thus consider 4 different

Table 5
Map color and its significant POIs with ranked difference G_s.

Ranked	Green	G_1	Orange	G_2	Pink	G_3
1	Fast food	0.538	Financial service	0.453	Shipping	1.172
2	Financial service	0.453	Fast food	0.352	Fuel	0.871
3	Coffee bar	0.308	Auto service	0.208	Tutoring	0.110
4	Auto service	0.259	Auto supply	0.207	Hotel	0.097
5	Eateries	0.257	Pharmacy	0.114	Grocery	0.087
6	Pharmacy	0.232	Eateries	0.095	Home improvement	0.082
7	Pet	0.195	Convenience	0.092	Auto rental	0.010
8	Personal care	0.173	Auto dealers	0.055	Park/lake(camping site)	0.004
9	Home improvement	0.099	Grocery	0.041	Machinery	0.004
10	Shopping center	0.089	Coffee bar	0.035	–	–

functional zones in our experiments. The area annotations are based on the results of clustering our latent region representations by CRF, as shown in Fig. 5(f) or 5(g). As the black areas are where no POI information regions are, and most regions in black are residential, vacant and natural areas, we did not list them in Table 5. In terms of the ranked difference of significant POIs in the table, we annotate the regions as follows:

Commercial/Downtown/Entertainment Area (Green regions): The distinguished POI in this functional zone are Pet, Personal Care and Shopping center, which are much significant in this area than others. Pets, Personal Care and Shopping center all provide service and merchandise, which indicates this functional zone is probably a commercial area. The other highly ranked POIs in this area also follow the characteristics of the commercial area, such as, financial service and coffee bar. Additionally, the famous shopping malls, such as Southpoint and Crabtree are all located in this functional zone.

Working/Education Area Mixed with Commercial Area in Inner City (Orange regions): Since we only count for the workdays, this area is annotated as a mixed use area. During the working time, its function is working/education. During other times, its commercial area. The G_s of this area focus on the financial service, fast food, and auto service, which is easy to compose a commercial area. Meanwhile, looking at Fig. 6, both industrial and Public and institutions are mixed with commercial part and located in orange regions, such as NC state, Duke, Chapel-Hill and Wake tech Universities.

Working/Education Area in Outer City (Pink regions): Most significant POIs in this functional zone are shipping, fuel, hotel and auto rental, which all refers to a outer-side of a city. Additionally, as our dataset is based on the workdays, and some industrial and tutoring are also located in the outer city, so this area is implied to be a working/education area. Consequently, this functional zone is annotated as a working/Education area in outer City. Since this area include few commercial area, it is different from above mixed with commercial one. According to Fig. 5(g), it is easy to find pink area is in outer side of centralized area. Moreover, the Research Triangle Park (RTP), a large company area is also assigned in the pink area, where it locates in middle part among the triangle part that consisted by the outer side of the cities of Raleigh, Durham and Chapel Hill.

5. Discussion

For further discussion, in this section we introduce the proper conditions to use our method. In our framework, there are two presumptions to propose the whole pipeline. One is the fine-grained region units and their neighbors may share the same region functionalities. Another is the HAP and POI pattern are follow the same introduced intrinsic pattern. For coarse grained granularity, the neighbors of each region unit are not in the same clusters, and each of them has their independent functionalities. Then CRF clustering for the spatial location consistency will be harmful for the final clustering results rather than helpful. Besides, the introduced inner pattern between POI and HAP is transferred from Wang et al. (2017), whose knowledge was discovered on the traffic pattern in Beijing, China. Other city may not share the same pattern of Beijing, then the latent region presentation is not accurate for clustering any more. In our case study, we did not have any prior knowledge of Raleigh city, thus we transfer the knowledge of Beijing city as a cold start, which also performs a good result. But if this pattern are totally different, the clustering result will be affected. Since in our framework, we have a mask to deal with the sparse data, therefore it also works for dense dataset, which should not be a concern.

Consequently, the proposed methodology would be more favorable for fine-grained granularity segmentation of regions. Also, our framework is a really good starting when lacking of the knowledge of the pattern in the dataset or obtaining a similar pattern between the HAP and POI.

6. Conclusion

In this paper, we proposed a framework for discovering functional zones on the basis of a novel algorithm to learn discriminative latent representations of regions and spatial clustering method. The proposed new algorithm leverages the latent effects between POIs and HAP, and successfully fuses the information of highly sparse POIs and HAP, and simultaneously enhances the discriminative ability. An intrinsic knowledge that explored from the real data are transferred in our framework to cope with biased data. Moreover, we also utilize the label consistency of CRF to preserve the spatial information when aggregating the functional zones, which unites the clustering of fine-girded regions. The difference of normalized POI distributions is employed to estimate the functionalities for a better annotation. Comparing with LDA-based methods, our method produces better results and a finer granularity with practical implementation conditions. The proposed framework can efficiently discover different functional zones and help designers to fully understand the urban structure.

CRediT authorship contribution statement

Wen Tang: Conceptualization, Methodology, Software, Validation, Writing – original draft. **Alireza Chakeri:** Conceptualization, Methodology, Writing – review & editing. **Hamid Krim:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgments

References

Beck, A., & Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, *2*(1), 183–202.

Boarnet, M., & Crane, R. (2001). The influence of land use on travel behavior: specification and estimation strategies. *Transportation Research Part A: Policy and Practice*, *35*(9), 823–845.

Chen, J., Yang, S., Wang, W., & Wang, M. (2015). Social context awareness from taxi traces: mining how human mobility patterns are shaped by bags of POI. In *Adjunct proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2015 ACM international symposium on wearable computers* (pp. 97–100). ACM.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, *39*(1), 1–22.

Demsar, U., Reades, J., Manley, E., & Batty, J. (2014). *Edge-based communities for identification of functional regions in a taxi flow network*. Department of Geodesy and Geoinformation, Vienna University of Technology.

Gao, S., Janowicz, K., & Couclelis, H. (2017). Extracting urban functional regions from points of interest and human activities on location-based social networks. *Transactions in GIS*, *21*(3), 446–467.

Guo, D. (2009). Flow mapping and multivariate visualization of large spatial interaction data. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6), 1041–1048.

Han, H., Yu, X., & Long, Y. (2015). Discovering functional zones using bus smart card data and points of interest in beijing. arXiv preprint arXiv:1503.03131.

Krähenbühl, P., & Koltun, V. (2011). Efficient inference in fully connected crfs with gaussian edge potentials. In *Advances in neural information processing systems* (pp. 109–117).

Liu, X., Gong, L., Gong, Y., & Liu, Y. (2015). Revealing travel patterns and city structure with taxi trip data. *Journal of Transport Geography*, *43*, 78–90.

Peng, C., Jin, X., Wong, K.-C., Shi, M., & Liò, P. (2012). Collective human mobility pattern from taxi trips in urban area. *PLoS One*, *7*(4), Article e34487.

Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65.

Tang, W., Chouzenoux, E., Pesquet, J.-C., & Krim, H. (2020). Deep transform and metric learning network: Wedding deep dictionary learning and neural networks. arXiv preprint arXiv:2002.07898.

Tang, W., Chouzenoux, E., Pesquet, J.-C., & Krim, H. (2021). Deep transform and metric learning networks. In *ICASSP 2021-2021 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 2735–2739). IEEE.

Tang, W., Otero, I. R., Krim, H., & Dai, L. (2016). Analysis dictionary learning for scene classification. In *Statistical signal processing workshop (SSP), 2016 IEEE* (pp. 1–5). IEEE.

Tang, W., Panahi, A., Krim, H., & Dai, L. (2018). Structured analysis dictionary learning for image classification. In *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 2181–2185). IEEE.

Tang, W., Panahi, A., Krim, H., & Dai, L. (2019a). Analysis dictionary learning: an efficient and discriminative solution. In *ICASSP 2019-2019 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 3682–3686). IEEE.

Tang, W., Panahi, A., Krim, H., & Dai, L. (2019b). Analysis dictionary learning based classification: Structure for robustness. *IEEE Transactions on Image Processing*, *28*(12), 6035–6046.

Wang, Y., Gu, Y., Dou, M., & Qiao, M. (2018). Using spatial semantics and interactions to identify urban functional regions. *ISPRS International Journal of Geo-Information*, *7*(4), 130.

Wang, M., Yang, S., Sun, Y., & Gao, J. (2017). Human mobility prediction from region functions with taxi trajectories. *PLoS One*, *12*(11), Article e0188735.

Yan, J., & Thill, J.-C. (2009). Visual data mining in spatial interaction analysis with self-organizing maps. *Environment and Planning B: Planning and Design*, *36*(3), 466–486.

Yuan, N. J., Zheng, Y., & Xie, X. (2018). Discovering functional zones in a city using human movements and points of interest. In *Spatial analysis and location modeling in urban and regional systems* (pp. 33–62). Springer.

Yuan, N. J., Zheng, Y., Xie, X., Wang, Y., Zheng, K., & Xiong, H. (2015). Discovering urban functional zones using latent activity trajectories. *IEEE Transactions on Knowledge and Data Engineering*, *27*(3), 712–725.

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science*, *28*(11), 2178–2199.
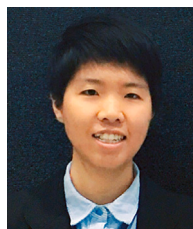
Wen Tang received one B.S. degree in Information and Computing Science from East China University of Science and Technology, Shanghai, China, in 2012. In the same year, she also received another B.S. degree in Computer Science from Fayetteville State University, Fayetteville, NC, USA. She is currently working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA. Further, she has been a visiting student at the Center for Visual Computing at CentraleSupelec, France. Her current research interests include machine learning and its applications for computer vision. More specifically, she is working on developing analysis dictionary learning methods and deep dictionary learning methods for image classification problems and its applications.

Alireza Chakeri received his B.Sc. degree in electrical engineering from Isfahan University of Technology, Isfahan, Iran in 2011. He obtained his M.Sc. degree in computer science from University of South Florida, Tampa, Florida in 2016, and his Ph.D. degree in computer science from the same institution in 2017. Alireza then joined Valassis Digital (formerly MaxPoint Interactive Inc.) in 2017 as the Staff Data Scientist. His current research interests include data mining and machine learning with focus on graph theory, game theory and statistical methods for large scale systems.

Hamid Krim received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Washington, Seattle, WA, USA, and the Ph.D. degree in electrical and computer engineering from Northeastern University, Boston, MA, USA. He was a Member of Technical Staff at AT&T Bell Labs, where he has conducted research and development in the areas of telephony and digital communication systems/subsystems. Following an NSF Postdoctoral Fellowship at Foreign Centers of Excellence, LSS/University of Orsay, Paris, France, he joined the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA, as a Research Scientist and where he was performing and supervising research. He is currently a Professor of electrical engineering in the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA, leading the Vision, Information, and Statistical Signal Theories and Applications Group. His research interests include statistical signal and image analysis and mathematical modeling with a keen emphasis on applied problems in classification and recognition using geometric and topological tools. He has served on the SP society Editorial Board and on TCs, and is the SP Distinguished Lecturer for 2015–2016.