Representing Urban Functions through Zone Embedding with Human Mobility Patterns

Zijun Yao[1], Yanjie Fu[2], Bin Liu[3], Wangsu Hu[1], Hui Xiong[1]*

[1]Rutgers University
[2]Missouri University of Science and Technology
[3]IBM Thomas J. Watson Research Center
{zijun.yao,wh251,hxiong}@rutgers.edu, fuyan@mst.edu, bin.liu1@ibm.com

Abstract

Urban functions refer to the purposes of land use in cities where each zone plays a distinct role and co-operates with each other to serve people's various life needs. Understanding zone functions helps to solve a variety of urban related problems, such as increasing traffic capacity and enhancing location-based service. Therefore, it is beneficial to investigate how to learn the representations of city zones in terms of urban functions, for better supporting urban analytic applications. To this end, in this paper, we propose a framework to learn the vector representation (embedding) of city zones by exploiting large-scale taxi trajectories. Specifically, we extract human mobility patterns from taxi trajectories, and use the "co-occurrence" of origin-destination zones to learn zone embeddings. To utilize the spatio-temporal characteristics of human mobility patterns, we incorporate mobility direction, departure/arrival time, destination attraction, and travel distance into the modeling of zone embeddings. We conduct extensive experiments with real-world urban datasets of New York City. Experimental results demonstrate the effectiveness of the proposed embedding model to represent urban functions of zones with human mobility data.

1 Introduction

A city consists of a variety of zones providing different functions to support diverse demands of urban residents, such as working, recreation, and residence. Studying the urban functions of city zones provides indispensable information which is useful in solving many urban challenges, therefore plays a critical role in urban analytics. Recent years, the advent of sensing technologies and mobile computing has accumulated a variety of data related to human mobility in urban areas. As a result, data-driven approaches have been increasingly applied to explore urban functions of cities.

While the literature has shown promising effectiveness of analyzing massive positioning data for urban exploration

---

* Contact author.

[Cranshaw *et al.*, 2012; Yuan *et al.*, 2012; 2015; Cheng *et al.*, 2011], there are limited studies aiming to provide an integrated and principled approach to the representation learning of city zones in terms of urban functions. In this paper, we aim to propose an effective solution to learn the distributed and low-dimensional embeddings of city zones. Zones with similar urban functions are geometrically closer in the embedding space. Using zone embeddings, we are able to identify functional regions of cities which consist of several zones with similar functions. Furthermore, many analytic models can use these extracted representations as enhanced inputs.

Generally, there are two critical challenges toward learning effective zone representations of urban functions: (i) how to infer urban functions: through intra-zone human activity or inter-zone human mobility; (ii) how to effectively exploit human mobility patterns containing spatio-temporal characteristics for learning zone embeddings.

First, by providing intelligence to profile human activity categories (e.g., shopping) within city zones, check-in data in location-based social networks (LBSNs) have been widely used for urban analytics. However, LBSNs have reliability issues due to (i) biased check-ins in entertaining related point-of-interest (POI) categories; (ii) sparse check-ins outside hot areas or time periods; and (iii) significant discrepancies between user check-in and actual mobility [Wang *et al.*, 2016]. More importantly, by analyzing intra-zone human activities, it is difficult to obtain association information between any two zones, which is essential in zone embedding learning for capturing "contexts". On the contrary, human mobility data (e.g., taxi trajectories) reveal an important association between two zones through the dynamic origin-destination mobility flow mostly related to function dependencies (e.g., commuters usually travel to an office zone around 9 a.m. from residential or transportation zones). In addition, human mobility data generally have a better coverage of unpopular areas and time periods. Therefore, analyzing human mobility has a potential to learn urban functions through zone embedding.

Second, an effective framework is highly needed to learn zone embeddings with human mobility patterns across city zones. For this purpose, we bring in the idea of word2vec [Mikolov *et al.*, 2013a; 2013b], which is originally a natural language processing (NLP) model for word semantic learning. In word2vec, the embedding of a word is learned from

its co-occur words which appear nearby in sentences. Therefore, two semantically similar words are likely to share similar vector representations. For example, "queen" and "king" would be close in the embedding space because both co-occur with the same nearby word "kingdom" frequently. Similarly, by analyzing a zone based on its "co-occur" zones between which human mobility patterns exist (e.g., a residential zone and an office zone "co-occur" when people travel from one to another), we can learn zone embeddings by analyzing association strength of every zone pair. With zone embeddings, zones with similar urban functions are geometrically close, whereas zones with different functions are distant.

Based on the above idea, a potential solution is to treat the origin-destination pair of zones in a human mobility pattern as a co-occurrence of two zones for zone embedding. However, urban functions are also jointly reflected by mobility direction and departure/arrival time. In other words, the embedding method should be able to take into account "leaving for" and "arriving from" at "different time" for modeling a zone co-occurrence. To that end, we define a set of human mobility events which contain zone, time, and status of mobility patterns, to serve as embedding "contexts" of target zones, for building directed and time-aware zone co-occurrences to include spatio-temporal characteristics. In addition, during the learning of zone embeddings, we give different importance to different co-occurrences by calculating the travel demands of origin-destination zone pairs based on destination attraction (e.g., total mobility pattern arrivals) and travel distance (e.g., average mobility pattern length) information.

Along these lines, in this paper, we present a novel human mobility based zone embedding framework to represent urban functions with distributed and low-dimensional vectors. Specifically, we develop a zone embedding model to exploit origin-destination information from massive human mobility patterns. We take into account mobility direction and departure/arrival time to model spatio-temporal co-occurrence of zones, and jointly incorporate destination attraction and travel distance to give different co-occurrences different importance in the embedding learning. Finally, we conduct extensive experiments with real-world urban datasets of New York City to show the effectiveness of the proposed method.

2 Problem Statement

We first provide some preliminary concepts in our work, then we proceed to the problem statement of zone embedding learning.

Definition 1 *(**Human mobility pattern**) Given a taxi trip, we extract the human mobility pattern with the following information: (i) origin-destination (O-D) pair of zones; (ii) time of departure (e.g., taxi passenger pick-up) and arrival (e.g., taxi passenger drop-off); and (iii) travel distance.*

Definition 2 *(**Human mobility event**) Given a human mobility pattern, we extract two human mobility events: one for the departure and one for the arrival with the following information: (i) event occurrence zone; (ii) event occurrence time; and (iii) event status: a departure or an arrival.*

Given a set of city zones $Z = \{z_1, z_2, ..., z_N\}$ and a set of taxi trips. Each trip contains the passenger travel with the in-

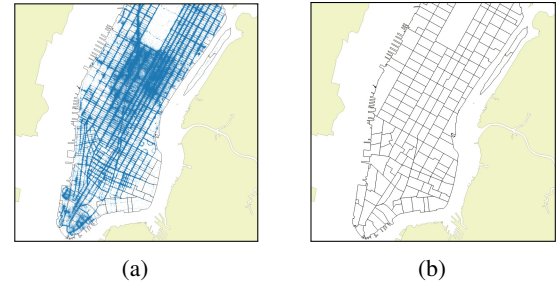(a)   (b)

Figure 1: (a) Locations of taxi pick-up in New York City (NYC). (b) City zones of NYC.

formation of locations and timestamps for the departure and the arrival. From taxi trips, we extract a set of zone level human mobility patterns $P = \{p_1, p_2, ..., p_m\}$, where each mobility pattern $p = (p.z_O, p.z_D, p.t_O, p.t_D)$ includes the zone of origin $p.z_O$ and destination $p.z_D$, as well as the time of departure $p.t_O$ and arrival $p.t_D$. Each time is converted from a timestamp to a $\langle timeslot, daytype \rangle$ combination. We define $\{ts_1, ts_2, ..., ts_J\}$ to be a set of time slots of a day (e.g., 24 hours), and $\{wd, we\}$ to be a set of day types: weekday and weekend. Figure 1 illustrates the pick-up locations of taxi trips, and the city zones in urban areas.

The objective is to learn the distributed and low-dimensional embeddings of city zones based on the spatio-temporal human mobility patterns for representing their urban functions in a city.

3 Methodology

In this section, we first review the word2vec model originated from semantic analysis. Then we show how to learn zone embeddings by incorporating spatio-temporal human mobility patterns. Last, we present the model specification.

3.1 Word Embedding

A fundamental observation in word embedding literature is that semantically similar words often have similar "contexts" (i.e., the words appear around them) in sentences [Levy and Goldberg, 2014; Mikolov *et al.*, 2013b]. By modeling the association strength of each word pair based on the frequencies they co-occur within a small context window (e.g., a window size of 5 means 5 words ahead of and 5 words behind the target word are considered as contexts), the embedding v_w of word w can be learned by

$$v_w{}^\mathsf{T} v'_c \approx \text{PMI}(\mathcal{D})_{w,c}, \tag{1}$$

where v'_c is the embedding of *context word* c.

PMI$(\mathcal{D})$ is a $|W| \times |W|$ pointwise mutual information (PMI) matrix calculated by word co-occurrence frequencies in corpus $\mathcal{D}$ with vocabulary size $|W|$. The PMI value in the $\langle w, c \rangle$ entry is computed as

$$\text{PMI}(\mathcal{D})_{w,c} = \log\left(\frac{\#(w,c) \cdot |\mathcal{D}|}{\#(w) \cdot \#(c)}\right), \tag{2}$$

where $\#(w, c)$ counts the frequency that words w and c co-occur, and $\#(w)$, $\#(c)$ count the frequencies that words w and c occur independently. $|\mathcal{D}|$ is the total number of word-context pairs in the corpus.
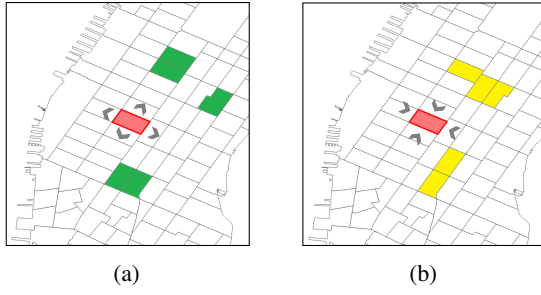
(a) (b)

Figure 2: An example of top human mobility flows of a transportation zone (red): (a) people move out to office zones (green) in 8-9 a.m., and (b) move back (yellow) in 5-6 p.m..

3.2 Spatio-Temporal Zone Embedding

We propose to learn the embedding of zones from its associated zones based on human mobility. People leave zones and arrive at zones at different time for different trip purposes, which reveals a function dependent association between zones. As shown in Figure 2, a transportation zone of Pennsylvania train station in NYC usually shows heavy mobility flows heading to office zones in the midtown in weekday mornings (8-9 a.m.), and shows huge mobility flows returning from office zones after work (5-6 p.m.). Therefore, we could define that when two zones form a origin-destination pair, a zone co-occurrence happens. By counting the number of origin-destination patterns, the frequency of zone co-occurrence can be obtained. Then zone embeddings can be learned by factorizing PMIs of zone pairs.

However, a problem of using the above simple zone co-occurrence is that the critical spatio-temporal characteristics of human mobility patterns (e.g., mobility direction and departure/arrival time) are not incorporated into the embedding framework. For example, given a mobility pattern which departs from zone A and arrives at zone B, to study the zone functions, the embedding framework needs to consider not only the co-occurrence of A and B, but also the corresponding mobility direction (e.g., A $\rightarrow$ B) and the departure/arrival time (e.g., at night).

To address this problem, we define a set of human mobility events E instead of zones to serve as embedding contexts (i.e., mobility events co-occur with zones in mobility patterns) so that the model includes zone co-occurrence, mobility direction, and departure/arrival time in the embedding learning. Specifically, a mobility event e is defined as follows:

$$e = (e.z, e.t, e.sta), \qquad (3)$$

where $e.z$ is the zone, $e.t$ is the time by time slot and day type, and $e.sta \in \{arrive, leave\}$ is the mobility status which shows the direction of the human mobility pattern. For example, a human mobility event can be defined as $\langle z_i, 12p.m.\text{-}weekday, arrive \rangle$ which means a human mobility pattern arrive at zone z_i during the time slot of 12 p.m.-1 p.m. on weekday.

From each mobility pattern $p = (p.z_O, p.z_D, p.t_O, p.t_D)$, we obtain two co-occurrences consisting of a zone as embedding target and a mobility event as embedding context:

zone	mobility event
$p.z_O$	$(p.z_D, p.t_D, arrive)$
$p.z_D$	$(p.z_O, p.t_O, leave)$

Using these directed and time-aware co-occurrences, we incorporate mobility direction and departure/arrival time of human mobility patterns to enable zone embedding.

Table 1 shows an analog between word embedding and zone embedding. In word embedding, every word serves as a target word and a context word. The association between a target word and a context word is learned from the frequency of word co-occurrences in sentences. Unlike traditional word embedding framework, city zones only serve as target "words" and mobility events only serve as context "words" in zone embedding. Accordingly, we use the co-occurrences of a zone and a mobility event in human mobility patterns to learn the pairwise zone associations.

	word embedding	zone embedding
target	word	zone
context	word	mobility event
co-occur in	sentence	mobility pattern

Table 1: Analog between word embedding and zone embedding.

3.3 Modeling Importance of Co-Occurrence

A unique characteristic of zone embedding is that different zones in contexts have different impacts based on how attractive the destination is and how far the mobility travels. Based on that, we propose to give different co-occurrences different importance for better optimizing zone embeddings.

For each origin-destination (O-D) pair of zones, we calculate the travel demand between them based on two factors: (i) destination attraction (e.g., total mobility patterns arrival at destination zones) and (ii) travel distance (e.g., average mobility pattern travel distance) with the gravity model of transportation analysis [Cascetta *et al.*, 2007]. Then we use the travel demand of the O-D zone pair to guide its importance of co-occurrence in the embedding learning with a 0-1 weight.

Specifically, we define G as a $|Z| \times |Z|$ gravity matrix ($|Z|$ is the total zone number) where row dimension means each zone as an origin, column dimension means each zones as a destination. Each row of G is a distribution of 0-1 probabilities that every destination zone z_D can attract a mobility pattern from a particular origin zone z_O, with sum of 1.

In detail, a $\langle z_O, z_D \rangle$ entry of G is calculated as

$$G(z_O, z_D) = \frac{A_{z_D} F_{z_O, z_D}}{\sum_{z \in Z} A_z F_{z_O, z}}, \qquad (4)$$

where A_z denotes the total number of mobility patterns arrive at zone z. F_{z_O, z_D} is the friction factor to serve as a cost for traveling between two zones, which is calculated based on the travel distance with a negative exponential function:

$$F_{z_O, z_D} = e^{-\beta d_{z_O, z_D}}, \qquad (5)$$

where d_{z_O, z_D} is the travel distance between zone z_O and z_D calculated by averaging the travel distances of mobility patterns. β is the parameter which is obtained by minimizing

$$\sum_{z_O \in Z} \sum_{z_D \in Z} (T_{z_O, z_D} - \hat{T}_{z_O, z_D})^2 \qquad (6)$$

with a genetic algorithm [Deb *et al.*, 2002], where T_{z_O,z_D} is the observed mobility pattern number from z_O to z_D. $\hat{T}_{z_O,z_D}$ denotes the estimated mobility pattern number calculated by $\hat{T}_{z_O,z_D} = P_{z_O} \cdot G(z_O, z_D)$, where P_{z_O} is the total number of mobility patterns leave z_O. Also, since the mobility distributions are usually different on weekday and weekend, we calculate the gravity matrices G^{wd} and G^{we} by two day types.

3.4 Model Specification

To learn the zone embeddings proposed in our framework, we minimize the following objective function over all co-occurrences of zones and mobility events:

$$\min_{V,V'} \frac{1}{2} \sum_{z \in Z} \sum_{e \in E} \left(M(z,e) - v_z^{\mathsf{T}} v'_e \right)^2 \cdot G^*(z_O, z_D), \quad (7)$$

where M denoted the $|Z| \times |E|$ matrix of positive point-wise mutual information (PPMI) which measures every co-occurrence of zones and mobility events, and $G^* \in \{G^{wd}, G^{we}\}$ denotes the gravity matrix for all origin-destination zone pairs of a particular day type.

In the first part, each value of M is computed as $M(z,e) = max\left(0, \log(\frac{\#(z,e) \cdot |\mathcal{T}|}{\#(z) \cdot \#(e)})\right)$, where $\#(z,e)$ counts the number of times that zone z and mobility event e co-occur, $\#(z), \#(e)$ count the numbers of single occurrence of z and e, and $|\mathcal{T}|$ is the observed number of co-occurrence from all mobility patterns.[1] v_z is the D-dimensional embedding of zone z, and v'_e is the D-dimensional embedding of mobility event e. We factorize PPMI matrix M into v_z and v'_e by minimize the square error for all co-occurrences.

In the second part, $G^*(z_O, z_D)$ denotes the gravity corresponding to the current co-occurrence which is retrieved by the origin-destination zones z_O, z_D and the day type $* \in \{wd, we\}$ interpreted from the current target zone and context mobility event. We use it to apply different importance on co-occurrences to enhance the embedding learning.

Parameter Estimation: Given the objective function in Equation (7), we take derivatives with respect to v_z and v'_e, and adopt gradient descent method to optimize embeddings. For experimental setup, we empirically set embedding dimension $D = 50$. Gravity matrices G^{wd} and G^{we} are calculated with $\beta^{wd} = 0.4674$ and $\beta^{we} = 0.3881$.

4 Evaluation

In this section, we empirically evaluate the performance of our proposed method. We perform all the experiments on real-world datasets of New York City (NYC). We choose NYC because of following reasons: (1) NYC contains massive population, diverse urban functions, and highly active economics. (2) Our experimental data has a sufficient coverage of NYC, therefore we are able to have convincing results.

4.1 Data Description

The first dataset is the human mobility data. We use the trip records of yellow taxi from NYC taxi & limousine commis-

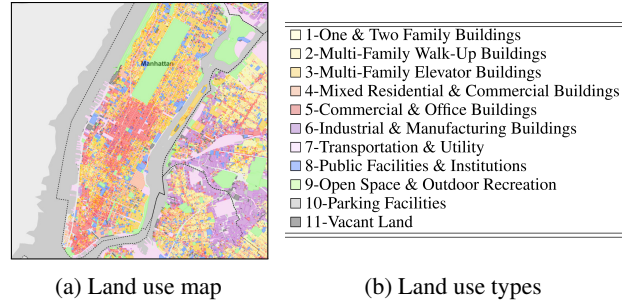| | 1-One & Two Family Buildings |
	2-Multi-Family Walk-Up Buildings
	3-Multi-Family Elevator Buildings
	4-Mixed Residential & Commercial Buildings
	5-Commercial & Office Buildings
	6-Industrial & Manufacturing Buildings
	7-Transportation & Utility
	8-Public Facilities & Institutions
	9-Open Space & Outdoor Recreation
	10-Parking Facilities
	11-Vacant Land

(a) Land use map   (b) Land use types

Figure 3: Land use of NYC for validation.

sion[2] to obtain citywide human mobility patterns. Since people in NYC seldom own cars, taxi is one of the most frequent and representative ground transportation choice. We use complete trip records of three months (June to August) in 2013. Finally we obtain 33,842,934 trips as our training data. Each taxi trip contains the locations and timestamps of a pick-up and a drop-off as well as the trip distance.

The second dataset is the zone data. We use the city zones designed by US Census Bureau[3] for zone embedding learning. Using these zones has three benefits: (i) they are professionally designed to contain homogeneous population and environment; (ii) they are separated by major road network and therefore make interpretable study results; (iii) they have relatively small areas and therefore provide relatively fine-grained analytic insights. At last we obtain 193 city zones for urban function study.

The last dataset is the Foursquare check-in data formulated by the work in [Yang *et al.*, 2015], since we compare the proposed model against the state-of-the-art baseline methods which analyze check-ins and point-of-interests (POIs) in location-base social networks (LBSNs). The dataset includes the check-in data in NYC for 10 months during 2012 to 2013. Each check-in contains the information of user ID, POI ID, location, timestamp and POI category. Finally, we have 17,009 POIs across 139 fine-grained categories, 1,046 LBSN users, and 109,073 check-ins.

4.2 Evaluation Metrics

For evaluation, we conduct k-means clustering on zone embeddings to partition the zone set into K urban function clusters. Zones with similar functions should be assigned to the same cluster. To validate clustering performance, we utilize the official land use dataset[4] of NYC as the ground-truth of zone functions. This dataset describes the land use designation of every lot (e.g., land of a building) by 11 land use types as shown in Figure 3. By aggregating all the lots at the zone level, for each zone, we obtain a vector consisting of the percentages of area for 11 different land use types. We then apply k-means clustering on land use vectors with different cluster number K for obtaining the ground-truth labels for city zones, We use the following metrics to evaluate zone clustering results of the proposed embedding method and baselines:

---

[1] We use PPMI to obtain a stable value because PMI can result in large negative values caused by log operation.

[2] http://www.nyc.gov/html/tlc/html/about/trip_record_data.shtml
[3] https://catalog.data.gov/dataset
[4] https://zola.planning.nyc.gov

• **Normalized Mutual Information (NMI)**, defined as

$$\text{NMI} = \frac{I(L;C)}{[H(L) + H(C)]/2}, \qquad (8)$$

where L is the set of true labels and C is the set of clusters. $I(L;C)$ denotes the sum of mutual information between any cluster c_i and any label l_j. $H(L)$ and $H(C)$ denote the entropy for labels and clusters, respectively. This metric evaluates the purity of clustering results from an information-theoretic perspective.

• **Adjusted Rand Index (ARI)**, defined as

$$\text{ARI} = \frac{RI - Expected\_RI}{Max\_RI - Expected\_RI}, \qquad (9)$$

which is the corrected-for-chance version of the Rand index (RI). By viewing the cluster assignments of all pair of zones as a series of decisions, we calculate the correctness by $RI = \frac{TP+TN}{TP+FP+TN+FN}$, where TP/FP denotes true/false positive and TN/FN denotes true/false negative. ARI has a score between -1.0 and 1.0 that random labeling has an ARI close to 0, and 1 stands for perfect match.

• **F-measure**, defined as

$$F_\beta = \frac{(\beta^2 + 1) \cdot Precision \cdot Recall}{\beta^2 \cdot Precision + Recall}, \qquad (10)$$

where $Precision = \frac{TP}{TP+FP}$ and $Recall = \frac{TP}{TP+FN}$. Similarly to ARI, we view the clustering result of each pair of zones as a decision, then we can have precision and recall. F-measure is the harmonic mean of precision and recall. We put more emphasis on precision than recall by $\beta = 0.5$.

• **Cluster Internal Difference (CID)**, defined as

$$\text{CID} = \frac{1}{N} \sum_{i=1}^{N} d(z_i, c_{z_i}), \qquad (11)$$

where N is the total number of zones. z_i is the land use vector of zone i. c_{z_i} is the average land use vector of i's assigned cluster. $d(x, y)$ calculates the Euclidean distance between the two vectors. Finally, we obtain the average distance from each zone's land use to its cluster's average land use. A smaller value means the zones assigned to the same cluster are more similar in terms of urban functions.

4.3 Baseline Approaches

The experimental study compares our proposed Mobility-based Zone Embedding (**ZE-Mob**) with the following approaches of zone representation for urban exploration.

• **TF-IDF (POI)**: An intuitive approach is to represent zone functions using intra-zone human activity types from LBSN data (e.g., POI category). We use Term Frequency-Inverse Document Frequency (TF-IDF) to measure the importance of different POI categories ("term") to a zone ("document"). Specifically, each zone can be represented by a $|C|$-dimensional vector where $|C|$ is the total number of unique POI categories. Each value is the TF-IDF of a category which is calculated based on the number of corresponding POIs. We apply k-means on TF-IDF vectors for zone clustering.

• **TF-IDF (Check-in)**: This baseline adopts the same TF-IDF method. The difference is that we use check-in frequency instead of POI number to count category frequencies.
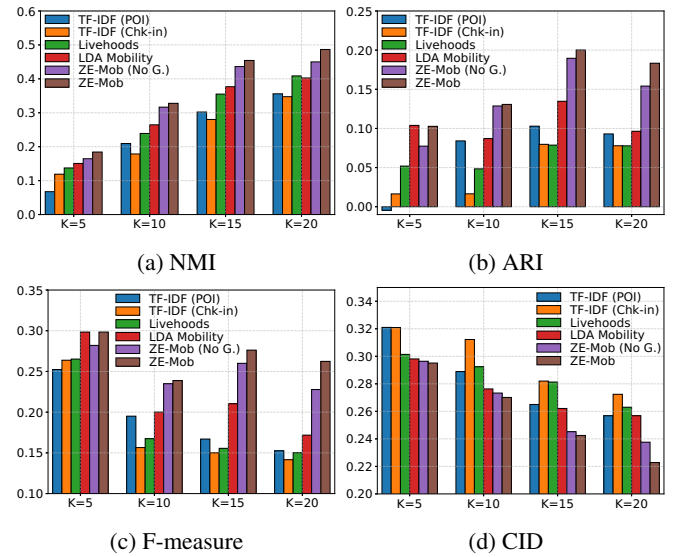

(a) NMI (b) ARI

(c) F-measure (d) CID

Figure 4: Performance comparisons on different cluster number K.

• **Livehoods**: The work in [Cranshaw *et al.*, 2012] proposes a clustering method to POIs by utilizing social similarity (e.g., check-in by the same user) and geographical proximity (e.g., distance between two POIs) in LSBN data. We aggregate POIs at the zone level and set locations of zones by zone centers to implement this baseline.

• **LDA Mobility**: The work in [Yuan *et al.*, 2012] proposes to mine human mobility data to learn urban functions of zones using latent Dirichlet allocation (LDA) model. The idea is to view zones as "document" and mobility events as "word". By having each zone as a "bag of mobility events" with observed frequencies, we learn the vectors of latent function topics. Then we apply k-means on topic vectors for zone clustering.

• **ZE-Mob (No Gravity)**: This baseline is the same with the propose method, except that we remove the gravity matrices G^*. Therefore, we do not model the importance of co-occurrence in zone embedding learning using destination attraction and travel distance of human mobility patterns.

4.4 Performance Comparisons

Figure 4 shows the Normalized Mutual Information (NMI), Adjusted Rand Index (ARI), F-measure, and Cluster Internal Difference (CID) of zone clustering results on all approaches with 4 different cluster number K: 5, 10, 15, and 20. Overall, we can see that our proposed approach ZE-Mob outperforms baseline methods on all metrics and all cluster number K.

Specifically, TF-IDF (POI) does not perform well. An important reason is that POIs in unpopular places such as residential areas are sparse. Meanwhile, since most of POIs concentrate in several categories (e.g., dinning), even the zones with sufficient POIs can not be differentiated by TF-IDF vector effectively. TF-IDF (check-in) performs the worst because check-in records are quite uneven on different POIs, which makes popular zones and unpopular zones more indistinguishable. Livehoods gives similar or better result than TF-IDF (POI) method by modeling zone affinities based on shared users and geographical distances. LDA Mobility gives better performance by learning latent urban function topics.

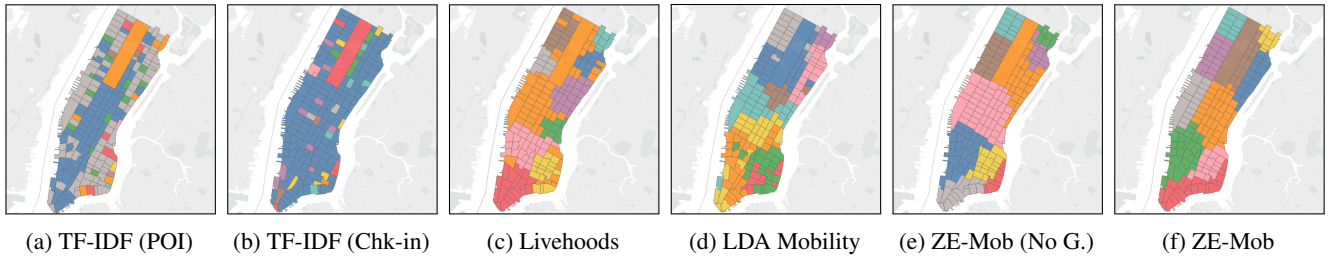| (a) TF-IDF (POI) | (b) TF-IDF (Chk-in) | (c) Livehoods | (d) LDA Mobility | (e) ZE-Mob (No G.) | (f) ZE-Mob |

Figure 5: City zone clustering ($K = 10$) for functional region identification. Each cluster is denoted by a unique color.

Last, ZE-Mob and ZE-Mob (No Gravity) achieve the best performances. Based on the better performance of ZE-Mob with gravity, we validate that the incorporation of co-occurrence importance effectively improves zone embedding.

4.5 Identifying Functional Regions

We perform k-means clustering (K=10) on zone embeddings to identify functional regions of NYC, such as residential regions, entertaining regions, and so on. As shown in Figure 5, each cluster is denoted by a color and used to identify a functional region. For TD-IDF (POI and check-in), the functional regions are weakly identified since the city is mainly partitioned into the zone clusters with sufficient POIs and the zone clusters with sparse POIs. For livehoods, the clustering does better job than TF-IDF using the same LBSN data, because it considers geographic proximity in the clustering analysis. However, its algorithm prevents POIs with far distances to be in the same cluster, as a result, the shape of clusters can not be extended which does not fit the real situation of NYC. LDA Mobility gives better performance than all LBSN based methods by utilizing mobility data with topic modeling. We can see that it makes more reasonable identification than Livehood. However, some clusters are still mixed with each other, which make functional regions not clear enough. ZE-Mob and ZE-Mob (No Gravity) give the most satisfied functional region identification. For both approaches, we can see that upper town, middle town, lower town are clearly identified, and the lower town is correctly separated into the Financial District and the East/West Village (two famous entertaining regions). Compared with ZE-Mob, ZE-Mob (No Gravity) assigns all mid-town zones into a single cluster, but different functions exist on the east side (manufacturing region) and the west side (commercial region) referring to the land use map in Figure 3a. Also, ZE-Mob (No Gravity) incorrectly merges the 5th Avenue shopping region into the Upper East region which mainly provides a residential function.

5 Related Work

In this section, we introduce the related work from three research perspectives: urban function learning, human mobility analysis, and word embedding learning.

Urban function study is an important research topic for city planners and urban designers in a long time for supporting decision making of city development. Early studies mainly rely on classic theory, long-term observation, and case-by-case survey for investigation. The work in [Goddard, 1970] surveys the taxi flow to analyze complex linkage sys-

tem exists in center of London to study the location of activities. The work in [Putnam, 2001] discusses the change of community in a perspective of people's social interactions. More recently, a series of work [Cranshaw *et al.*, 2012; Yuan *et al.*, 2012; 2015; Kling and Pozdnoukhov, 2012; Liu *et al.*, 2017] use large-scale positioning data to perform data-driven urban function analysis.

Human mobility study has attracted many attentions in urban data analytics. Early work in [McFadden, 1974] develops a multiple dimensions to analyze people's travel demands from their traveling behavior. More recently, the availability of vehicle GPS traces has empowered many urban intelligence applications. The work in [Zheng *et al.*, 2011] uses trajectories of taxicabs to detect flawed urban planning areas, such as O-D region pairs with traffic issues. The work in [Ge *et al.*, 2010; Yuan *et al.*, 2011] analyzes spatio-temporal patterns of city taxi's pick-up and drop-off behavior and driving route to find the optimal strategy for helping taxi driver increase their cars' occupancy rate. The work in [Wang *et al.*, 2015] uses taxi, bus, and subway data to measure spatial connectivity among areas for boosting the performance of user location prediction. The work in [Yao *et al.*, 2016] utilizes taxi drop-offs to profile temporal popularity patterns of POIs for improving the performance of POI recommendation.

Word embedding learning has been studied in recent years through deep neural networks [Bengio *et al.*, 2003; Collobert and Weston, 2008]. Later, the work in GloVE [Pennington *et al.*, 2014] and word2vec [Mikolov *et al.*, 2013b; 2013a] which utilizes word dependency for semantic analysis, has shown significant improvement on major NLP tasks, such as document clustering [Kusner *et al.*, 2015] and word similarity [Levy *et al.*, 2015]. More recenltly, the work in [Levy and Goldberg, 2014] shows an equivalence of matrix factorization of a shifted PMI matrix to word2vec skip-gram.

6 Conclusion

In this paper, we presented a city zone embedding framework using human mobility patterns to represent urban functions as distributed and low-dimensional vectors. For this purpose, we exploited the human mobility patterns from massive taxi trajectories to model the embeddings with zone associations. Specifically, we developed a spatio-temporal zone embedding model by incorporating mobility direction and departure/arrival time for building directed and time-aware zone co-occurrence, and utilizing destination attraction and travel distance for calculating travel demand of origin-destination pair of zones as co-occurrence importance. Extensive exper-

iments on real-world datasets demonstrated the effectiveness of the proposed method with a consistent performance improvement over all baselines.

Acknowledgments

References

[Bengio *et al.*, 2003] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. A neural probabilistic language model. *JMLR*, 3(Feb):1137–1155, 2003.

[Cascetta *et al.*, 2007] Ennio Cascetta, Francesca Pagliara, and Andrea Papola. Alternative approaches to trip distribution modelling: a retrospective review and suggestions for combining different approaches. *Papers in Regional Science*, 86(4):597–620, 2007.

[Cheng *et al.*, 2011] Zhiyuan Cheng, James Caverlee, Kyumin Lee, and Daniel Z Sui. Exploring millions of footprints in location sharing services. *ICWSM*, 2011:81–88, 2011.

[Collobert and Weston, 2008] Ronan Collobert and Jason Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *ICML*, pages 160–167. ACM, 2008.

[Cranshaw *et al.*, 2012] Justin Cranshaw, Raz Schwartz, Jason I. Hong, and Norman M. Sadeh. The livehoods project: Utilizing social media to understand the dynamics of a city. In *ICWSM*, 2012.

[Deb *et al.*, 2002] Kalyanmoy Deb, Amrit Pratap, Sameer Agarwal, and TAMT Meyarivan. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation*, 6(2):182–197, 2002.

[Ge *et al.*, 2010] Yong Ge, Hui Xiong, Alexander Tuzhilin, Keli Xiao, Marco Gruteser, and Michael Pazzani. An energy-efficient mobile recommender system. In *KDD*, pages 899–908. ACM, 2010.

[Goddard, 1970] John Burgess Goddard. Functional regions within the city centre: A study by factor analysis of taxi flows in central london. *Transactions of the Institute of British Geographers*, pages 161–182, 1970.

[Kling and Pozdnoukhov, 2012] Felix Kling and Alexei Pozdnoukhov. When a city tells a story: urban topic analysis. In *SIGSPATIAL*, pages 482–485. ACM, 2012.

[Kusner *et al.*, 2015] Matt J. Kusner, Yu Sun, Nicholas I. Kolkin, and Kilian Q. Weinberger. From word embeddings to document distances. In *ICML*, pages 957–966, 2015.

[Levy and Goldberg, 2014] Omer Levy and Yoav Goldberg. Neural word embedding as implicit matrix factorization. In *NIPS*, pages 2177–2185, 2014.

[Levy *et al.*, 2015] Omer Levy, Yoav Goldberg, and Ido Dagan. Improving distributional similarity with lessons learned from word embeddings. *Transactions of the Association for Computational Linguistics*, 3:211–225, 2015.

[Liu *et al.*, 2017] Junming Liu, Leilei Sun, Qiao Li, Jingci Ming, Yanchi Liu, and Hui Xiong. Functional zone based hierarchical demand prediction for bike system expansion. In *KDD*, pages 957–966. ACM, 2017.

[McFadden, 1974] Daniel McFadden. The measurement of urban travel demand. *Journal of Public Economics*, 3(4):303–328, 1974.

[Mikolov *et al.*, 2013a] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[Mikolov *et al.*, 2013b] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *NIPS*, pages 3111–3119, 2013.

[Pennington *et al.*, 2014] Jeffrey Pennington, Richard Socher, and Christopher D Manning. Glove: Global vectors for word representation. In *EMNLP*, volume 14, pages 1532–1543, 2014.

[Putnam, 2001] Robert D Putnam. *Bowling alone: The collapse and revival of American community*. Simon and Schuster, 2001.

[Wang *et al.*, 2015] Yingzi Wang, Nicholas Jing Yuan, Defu Lian, Linli Xu, Xing Xie, Enhong Chen, and Yong Rui. Regularity and conformity: Location prediction using heterogeneous mobility data. In *KDD*, pages 1275–1284. ACM, 2015.

[Wang *et al.*, 2016] Gang Wang, Sarita Yardi Schoenebeck, Haitao Zheng, and Ben Y. Zhao. "will check-in for badges": Understanding bias and misbehavior on location-based social networks. In *ICWSM*, pages 417–426, 2016.

[Yang *et al.*, 2015] Dingqi Yang, Daqing Zhang, Vincent. W. Zheng, and Zhiyong Yu. Modeling user activity preference by leveraging user spatial temporal characteristics in lbsns. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(1):129–142, 2015.

[Yao *et al.*, 2016] Zijun Yao, Yanjie Fu, Bin Liu, Yanchi Liu, and Hui Xiong. POI recommendation: A temporal matching between POI popularity and user regularity. In *ICDM*, pages 549–558, 2016.

[Yuan *et al.*, 2011] Jing Yuan, Yu Zheng, Liuhang Zhang, XIng Xie, and Guangzhong Sun. Where to find my next passenger. In *Ubicomp*, pages 109–118. ACM, 2011.

[Yuan *et al.*, 2012] Jing Yuan, Yu Zheng, and Xing Xie. Discovering regions of different functions in a city using human mobility and pois. In *KDD*, pages 186–194. ACM, 2012.

[Yuan *et al.*, 2015] Nicholas Jing Yuan, Yu Zheng, Xing Xie, Yingzi Wang, Kai Zheng, and Hui Xiong. Discovering urban functional zones using latent activity trajectories. *TKDE*, 27(3):712–725, 2015.

[Zheng *et al.*, 2011] Yu Zheng, Yanchi Liu, Jing Yuan, and Xing Xie. Urban computing with taxicabs. In *Ubicomp*, pages 89–98. ACM, 2011.