Representation Learning for Geospatial Areas using Large-scale Mobility Data from Smart Card

Masanao Ochi
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku
Tokyo, Japan
masanao.oochi@gmail.com

Yuko Nakashio
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku
Tokyo, Japan
kuppylemonade.0d.t@gmail.com

Yuta Yamashita
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku
Tokyo, Japan
guruttosekai2011@gmail.com

Ichiro Sakata
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku
Tokyo, Japan
isakata@ipr-ctr.t.u-tokyo.ac.jp

Kimitake Asatani
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku
Tokyo, Japan
asatani@ipr-ctr.t.u-tokyo.ac.jp

Matthew Ruttley
Hearst
300 W. 57th Street, New York
NY, United States
ruttleym@googlemail.com

Junichiro Mori
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku
Tokyo, Japan
jmori@ipr-ctr.t.u-tokyo.ac.jp

Ubicomp/ISWC'16 Adjunct , September 12-16, 2016, Heidelberg, Germany
ACM 978-1-4503-4462-3/16/09.
http://dx.doi.org/10.1145/2968219.2968416

Abstract
With the deployment of modern infrastructures for public transit, several studies have analyzed the transition patterns of people by using smart card data and have characterized the areas. In this paper, we propose a novel embedding method to obtain a vector representation of a geospatial area using transition patterns of people from the large-scale data of their smart cards. We extend a network embedding by taking into account geographical constraints on people transitioning in the real world. We conducted an experiment using smart card data in a large network of railroads in Kansai areas in Japan. We obtained a vector representation of each railroad station using the proposed embedding method. The results show that the proposed method performs better than the existing network embedding methods in the task of multi-label classification for purposes of going to a railroad station. Our proposed method can contribute to predicting people flow by discovering underlying representations of geospatial areas from mobility data.

Author Keywords
Network Embedding; Auto Fare Collection; Representation Learning; Trajectory Data Mining; Spatial Databases

ACM Classification Keywords
H.2.8. [DATABASE MANAGEMENT]: Database Applications; Spatial databases and GIS []

Introduction

As the use of location-aware personal devices such as smartphones rapidly spreads, a large amount of mobility data, including GPS or cell tower logs, has been automatically accumulated [17, 8].

In addition to such personal smart devices, the deployment of recent infrastructures for public transit such as automated fare collection (AFC) systems with smart card enables to collect large volumes of mobility data including people's activities with detailed time and space information [9]. Such mobility data has been widely used to analyze characteristics of public transport systems and passenger behaviors [13, 16]. More recently, several studies on modeling and predicting people flow with the mobility data have been conducted for several purposes of city planning, disaster prevention, and advertising [3, 9, 16].

Modeling and predicting people flow in a specific area results in understanding the characteristics or demographics of the area by combining activity patterns of people with external information about the area [7]. Recent studies have analyzed transition patterns of people from one area to another using smart card data and characterized the areas or identified the segmentation of the areas [4, 14]. While these studies solely assume that an area falls into some predefined demographics based on people flow in the area, they do not consider any underlying representation of the area to be characterized. However, if we regard transition patterns of people on an area as the context of its area, we could potentially obtain the reprensetaion of the area as shown in recent studies on representation learning [1].

The basic notion of representation learning is that two entities are semantically similar if they are sharing common contexts; this is known as a distributional hypothesis in linguistics, which states that words that occur in similar contexts tend to have similar meanings [5]. That idea of representation learning has been recently expanded to a network embedding method [6, 12] that tries to solve the problem of embedding networks into low-dimensional vector spaces by assuming that two nodes are similar if they are closely connected in a network. Network embedding is useful in tasks such as visualization, node classification, and link prediction. In the case of embedding geographical areas, if we consider areas as nodes and transition patterns of people between areas as links, we can formalize the problem of embedding geospatial areas as an extension of studying the embedding of a network. Intuitively, transition patterns of people in a business district are different from those in a residential district. Therefore, we can distinguish those different types of geospatial areas by embedding a people transition network in our low-dimensional vector spaces.

However, we can not simply apply existing network embedding methods to our problem of embedding geospatial areas. In the case of people transitioning in a large network of transportation systems such as railroads, there exist several geographical constraints on the transition. For example, a person who lives around Esaka Station goes shopping at Umeda Station but not at Tennoji Station despite the fact that people go to both Umeda and Tennnoji stations for similar purposes. The reason is that Tennoji Station is further than Umeda from Esaka Station[1]. We can assume that people usually tend to minimize their transitions depending on their purpose of activities and given available means of transportation on their current locations. If we consider geospatial areas as a network, which is connected with links of people transition patterns between areas, and try to embed the network in a low-dimensional vector space to obtain representations of the areas, we have to take into

[1] See the train route map in Figure 1

account such geographical constraints on people's transitions in a real world. In this paper, we propose a novel embedding method to obtain a vector representation of a geospatial area using transition patterns of people from large-scale data of smart cards. We describe the purpose to go to a station as a station role. Considering geographical constraints, we can collect distant stations which play similar roles in each area. It will be useful for a city planner or a marketer to find strong and weak characteristics with each area.

We conducted an experiment using smart card data in a large network of railroads in Kansai areas in Japan. We obtained a vector representation of each railroad station using the proposed embedding method and evaluated our embedding method in the task of multi-label classification for purposes to go to a railroad station.

Our contributions in this paper is two-fold:

1. We propose a novel embedding method to obtain a vector representation of a geospatial area using transition patterns of people from large-scale data of smart cards.
2. We demonstrate that our proposed method can collect distant areas where people come for the same purpose.

Related Works

Modeling the characteristics of spatial areas using mobility data
Several studies on modeling and predicting people flow with the mobility data have been recently conducted for several purposes of city planning, disaster prevention, and advertising [3, 10, 16]. Modeling and predicting people flow in a specific area results in understanding the characteristics of the area. Recent studies have analyzed transition patterns of people from one area to another using smart-card data and characterized the areas or identified the segmentation of the areas [4, 15]. While these studies solely assume that an area falls into some pre-defined demographics based on people flow in the area, they do not consider any underlying representation of the area to be characterized. In this paper, we try to obtain a vector representation of a geospatial area using the transition patterns of people from large-scale data of their smart cards, and we collect distant areas which play similar roles.

Studies on the network embedding
With recent advances in representation learning, an embedding method has been recently applied to network data to solve the problem of embedding networks into low-dimensional vector spaces [6, 11, 12]. In this paper, we can formalize our problem of embedding geospatial areas as an extension of learning the embedding of a network. When people transition in a large network of transportation systems such as railroads, there exists several geographical constraints on the transition. We can assume people tend to minimize their transitions depending on their purpose of activities given available means of transportation at their current locations. We introduce an embedding method for a geospatial area by taking into account factors such geographical constraints on people transitioning in the real world. We evaluated our proposed method with actual, large-scale, mobility data from smart cards.

Method

First, this section describes the **geographical constraints model** in which the people flow is derived from geolocation and purpose. Next, we explain how the network develops from massive people flow and the necessity of label propagation on the network. Finally, we propose a representation

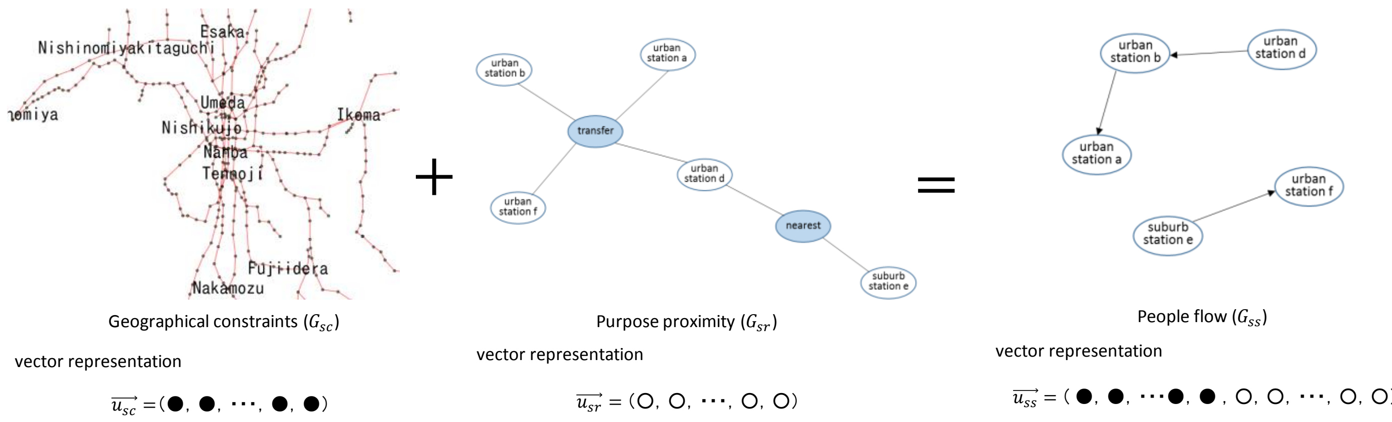

Figure 1: Schematic of the geographical constraints model and the vector representation.

learning algorithm based on this constraints model and explain it precisely.

Geographical Constraints model

We assume that people tend to minimize their transitions depending on the purpose of their activities and the available means of transportation at their current location. We model this assumption with the geographical constraints model, as shown in Figure 1. This model is composed of three components. The first one is the geographical constraints (G_{sc}), which describe the geographical proximity, such as a subway route map. In this paper, we adopt the station–company graph because we regard stations that belong to the same company as being part of same region. The second one is the purpose proximity (G_{sr}), for which we adopt the station–role graph that describes the purposes for which people go to each station. The last one is the people flow, (G_{ss}), in which we adopt the getting-on-and-off network that describes the people transition rate between two stations, with direction.

In this model, we design that the geographical proximity and the purpose proximity generate the people flow. We think that the three networks' relationship depends on the distance on the latent vector representation. There are three graphs that do not mutually share their vectors: $u_i^{\vec{ss}}$ for all vertices $v_i \in G_{ss}$, $u_i^{\vec{sc}}$ for all vertices $v_i \in G_{sc}$, and $u_i^{\vec{sr}}$ for all vertices $v_i \in G_{sr}$. We lead to the following equation from the schematic of our geographical constraints model shown in Figure 1 and these vetor representatons.

$$u_i^{\vec{ss}} = u_i^{\vec{sc}} + u_i^{\vec{sr}} \tag{1}$$

In this paper, we interpret the operator "+" as connecting two vectors and producing a new vector with dimensions that are twice as numerous as the number of dimensions of each vector, not that we add each element in the two vectors.

Learning algorithm

Based on the geographical constraints model, vector representations are acquired by the learning algorithm shown in

Table 1: Learning Algorithm

Learning Algorithm	
1:	**Input**: $G_{sc}, G_{sr}, G_{ss}, T, \rho_0, K$.
2:	**Output**: $u^{\vec{sc}}, u^{\vec{sr}}, u^{\vec{ss}}$.
3:	Initialize each vector $u^{\vec{sc}}, u^{\vec{sr}}, u^{\vec{ss}}, u^{\vec{\prime sc}}, u^{\vec{\prime sr}}, u^{\vec{\prime ss}}$.
4:	**for** $t = 1$ **to** T
5:	Sample an edge e_{ij}^{sc} from G_{sc}.
6:	Load $u_i^{\vec{sc}}$ and $u_j^{\vec{\prime sc}}$ from the corresponding part of $u_i^{\vec{ss}}$ and $u_j^{\vec{\prime ss}}$.
7:	Update $u_i^{\vec{sc}}$ and $u_j^{\vec{\prime sc}}$ using the objective function O_{sc}.
8:	Overwrite the corressponding part of $u_i^{\vec{ss}}$ and $u_j^{\vec{\prime ss}}$ with $u_i^{\vec{sc}}$ and $u_j^{\vec{\prime sc}}$.
9:	Sample an edge e_{ij}^{sr} from G_{sr}.
10:	Load $u_i^{\vec{sr}}, u_j^{\vec{\prime sr}}$ from the corresponding part of $u_i^{\vec{ss}}$ and $u_j^{\vec{\prime ss}}$.
11:	Update $u_i^{\vec{sr}}$ and $u_j^{\vec{\prime sr}}$ using the objective function O_{sr}.
12:	Overwrite the corressponding part of $u_i^{\vec{ss}}, u_j^{\vec{\prime ss}}$ with $u_i^{\vec{sr}}, u_j^{\vec{\prime sr}}$.
13:	Sample an edge e_{ij}^{ss} from G_{ss}.
14:	Update $u_i^{\vec{ss}}$ and $u_j^{\vec{\prime ss}}$ using the objective function O_{ss}.
15:	END

Table 1. This algorithm needs the geographical constraints network (G_{sc}), the purpose proximity network (G_{sr}), the people flow network (G_{ss}), the number of sampling (T), the initial learning rate (ρ_0), and the number of negative sampling (K) as input. We apply the network embedding model called the "LINE(2nd) model" proposed by Tang et.al [12]. This model approximates second-order proximity between two vertices, optimizing each representation vector. The objective function is as follows:

$$O = - \sum_{(i,j) \in E} w_{ij} \ln p(v_j | v_i) \tag{2}$$

In this equation, w_{ij} indicates the empirical edge weight from the vertex v_i to the vertex v_j. $p(v_j|v_i)$ which is the transition probability from v_i to v_j is estimated using the embedding vector $\vec{u_i}$ of the vertex v_i and the context vector $\vec{u'_j}$ of the vertex v_j as following:

$$p(v_j|v_i) = \frac{\exp(\vec{u'}_j^T \cdot \vec{u}_i)}{\sum_{k' \in |V|} \exp(\vec{u'}_{k'}^T \cdot \vec{u}_i)} \tag{3}$$

We set this objective function for three networks individually and acquire vertex embedding vector sequentially (Line 7,11 and 14) based on the geographical constraint model (Line 8 and 12).

Data Description

In this paper, we conducted experiments with two datasets. One is the getting on and off dataset, and the other one is a dataset of the purpose of use of some main stations. In this section, we explain these two datasets.

The getting on and off dataset

Experiments are carried out using the getting on and off dataset of the IC card user. The six railway companies in Japan's Kansai region provide this dataset. The providers anonymise this dataset. The contents of the dataset mainly consist of six elements: each user's gender, age, getting on and off date and time, and boarding and destination station. The summary of this dataset is shown in Table 2.

Table 2: Overview of the getting on and off dataset.

Starting date	March 01, 2015
Ending date	March 31, 2015
Total number of records	50,925,951
The number of unique users	2,007,507
The number of variety of stations	672
The number of railways companies	6

The purpose of use dataset
This study aims to estimate the role of each station. We took the approach of conducting a questionnaire. The questionnaire was carried out by Lancers[2], a crowdsourcing service in Japan. The respondents were regular users of each railway company. The time span was two weeks from March 23, 2016. The target stations were limited to the top 100 that have a large number of passengers. Respondents were asked to select three stations, from each railway company, that satisfy the following criterion/purpose: the nearest, commuting or attending school, transit, shopping, dining, and entertainment. As a result, we got 2,219 questionnaire results for 96 stations from the total 223 users. In this experiment, we used two results (the nearest, shopping).

Experiment and Result

In this section, we evaluate the effectiveness of our proposed model for geospatial people flow data. For this purpose, we conduct an experiment to collect data from distant areas to which people come for the same purpose. We describe the experimental setting and show the result here.

Input data
We arrange these datasets to input and for experiments. The station–company graph is a graph representing which company a station belongs to. It is an undirected graph, and the weights of all edges are equal. The station–role graph is a graph representing the distribution of purposes for people to go to a station. It is an undirected graph, and the weights of each edge are $P(role|station)$. In other words, the total of all roles for each station is normalized by 1.0. Finally, the getting on and off graph is a graph showing the people getting on and off between two stations. It is a directed graph, and the weight of each edge is $P(destination|boarding)$. That is the sum of the number of users moving from one boarding station at all getting-off stations, which is normalized by 1.0. The station–company garph and the getting on and off graph is made from the getting on and off dataset, and the station–role graph is made from the pupose of use dataset.

[2] http://www.lancers.jp

Experimental procedure and parameter setting
We input three network graphs in the former subsection. We evaluate our proposed method effectiveness compared with the PTE method [11]. The PTE method is the latest network embedding method for a heterogeneous network. This method applies to three different networks which are the word–word, word–document, and word–label networks and can acquire each word, document, and label vector representation. They propose two learning styles, which are the “pre-train” and “joint” learning styles. We select the “joint” learning style, which is slightly better than the pre-train learning style in their report (PTE(joint)). This method can embed all vertices in three network graphs to the same vector spaces. The same vertex in different graphs has the same vector representation among all graphs. This is different from our proposed model, which embed vertices in three network graphs to different vector spaces.

As described in this paper, we conducted a multi-label classification experiment for station roles. Estimating the role of the station is carried out by a multi-label classification using one-vs-rest SVM[3] method using the learned vector and the tagged role label to a station as the training data. We apply two cross-validations for the station–role data. In other words, G_{ss} and G_{sc} are used all for the sake of training, but we select the half randomly from the station–role data for the G_{sr} training data, and the remaining to the test data. Then, we compare the estimation results between the PTE

[3] We use the “LIBSVM” package [2].

method and the proposed method. We measure the classification performance with the Macro and Micro metrics.

Finally, it is evaluated by the geographical location of the actual station which is near the vector representation of each role label. The evaluation metric is the average value of the standard deviation of the actual geolocation of stations near the role label vector. As the mean of the standard deviation of the nearby stations around the role label is larger, the station group is extracted for the role of the station without the geographical constraints.

Other parameter settings are as follows. We set the total number of sampling size $T = 100,000$, and the SGD learning rate $\rho_t = \rho_0(1 - t/T), \rho_0 = 0.05$. The number of negative sampling K is set as 5. The dimensionality of a vertex vector is set as 100, but in the proposed method, $u_i^{\vec{ss}}, u_i^{\vec{\prime ss}}$ has twice the number of dimension; that is 200. All these parameter settings are the same in the PTE method and in the proposed method.

With the PTE method, we use the "joint training" version. And the word–word network is set as the getting on and off network, the word–document network is set as the station–company network, and the word–label network is set as the station–role network in the experiment.

Table 3: A result of the role estimation (mean value of 2 cross-validations)

method	Macro			Micro		
	Precision(%)	Recall(%)	F1-value(%)	Precision(%)	Recall(%)	F1-value(%)
PTE(joint)	30.556	45.833	36.376	57.190	57.190	57.190
proposed($u_i^{\vec{sc}}$)	31.536	50.000	38.547	63.072	63.072	63.072
proposed($u_i^{\vec{sr}}$)	31.076	47.917	37.500	60.131	60.131	60.131
proposed($u_i^{\vec{ss}}$)	**31.536**	**50.000**	**38.547**	**63.072**	**63.072**	**63.072**

Table 4: the standard deviation of station location around each role vector

method	near@10				near@50			
	role:"nearest"		role:"shopping"		role:"nearest"		role:"shopping"	
	long SD	lat SD	long SD	lat SD	long SD	lat SD	long SD	lat SD
PTE(joint)	0.097	0.023	0.074	0.023	0.129	0.124	0.148	0.118
proposed($u_i^{\vec{sc}}$)	0.444	**0.180**	0.124	0.114	0.346	0.147	0.388	**0.138**
proposed($u_i^{\vec{sr}}$)	0.024	0.031	0.024	0.031	**0.471**	0.142	0.311	0.104
proposed($u_i^{\vec{ss}}$)	**0.513**	0.172	**0.475**	**0.168**	0.459	**0.169**	**0.396**	0.131

Results

We show in Table 3 the results of estimating the role of the station. In Table 3 and 4, we present one PTE result (PTE(joint)) and three proposed method results (proposed($u_i^{\vec{sc}}$), proposed($u_i^{\vec{sr}}$), and proposed($u_i^{\vec{ss}}$)). This is caused by each method output. Though the PTE method embeds all vertices to the same vector space, our proposed method embeds all vertices to three different vector spaces. So each vertex has one vector representation on the PTE method and three vector representations ($u_i^{\vec{sc}}$, $u_i^{\vec{sr}}$, and $u_i^{\vec{ss}}$) on the proposed method. This result compares the performance of station role estimation on the large-scale mobility data between the PTE method and the proposed method. We can see that the proposed ($u_i^{\vec{ss}}$) performs the best of all other methods. The performance of PTE (joint) is inferior to all proposed methods.

We show in Table 4 the results of the standard deviation of station location around each role vector. This result compares the performance of station mapping around each role vector. We can see that the longitude standard deviation (SD) is bigger than latitude SD because this railway network is longer from east to west than from north to south. We can see that the proposed ($u_i^{\vec{ss}}$) method has bigger standard deviation value than other methods on average.

The proposed method embeds distant stations to the place around each role.

Consideration

As described in Section Results, our proposed method achieve better results than the PTE method. These results indicate that, for large-scale movement data that has spatial dependence, the proposed method captures the characteristics of the role of each station better than the PTE method does. People's moving area is usually small, so they live in limited areas. In light of this geographical constraints, our proposed method works better than the PTE method.

These results indicate that, in large-scale movement data that has the spatial dependency, the proposed method captures well the characteristics of the role of each station, compared to the PTE method.

On the other hand, our proposed method ($u_i^{\vec{s}r}$) results are not the best results on the station role estimation and the station geolocation SD. Our research aims to decompose the effect of geographical constraints and purpose proximity completely. So, we think these results show that our research has considerable room for improvement, which leave future work.

Summary

In this paper, we proposed a novel embedding method to obtain a vector representation of a geospatial area using transition patterns of people from large-scale data of smart cards. We showed that the proposed method estimated the role of the station better than the heterogeneous network embedding method, PTE. In the future, there will be a need to carry out a multi-faceted evaluation or apply the proposed method to other datasets. Also, the proposed method can be regarded as simultaneous learning of vector representations of multiple graphs.

REFERENCES

1. Y. Bengio, A. Courville, and P. Vincent. 2013. Representation Learning: A Review and New Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35, 8 (2013), 1798–1828.

2. Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A Library for Support Vector Machines. *ACM Trans. Intell. Syst. Technol.* 2, 3, Article 27 (May 2011), 27 pages.

3. Takahiko Kusakabe and Yasuo Asakura. 2014. Behavioural data mining of transit smart card data: A data fusion approach. *Transportation Research Part C: Emerging Technologies* 46 (2014), 179 – 191.

4. Ying Long and Zhenjiang Shen. 2015. *Geospatial Analysis to Support Urban Planning in Beijing*. Springer International Publishing, Cham, Chapter Discovering Functional Zones Using Bus Smart Card Data and Points of Interest in Beijing, 193–217.

5. Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. In *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 3111–3119.

6. Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. DeepWalk: Online Learning of Social Representations. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '14)*. ACM, New York, NY, USA, 701–710.

7. Santi Phithakkitnukoon, Teerayut Horanont, Giusy Di Lorenzo, Ryosuke Shibasaki, and Carlo Ratti. 2010. Activity-aware Map: Identifying Human Daily Activity Pattern Using Mobile Phone Data. In *Proceedings of the First International Conference on Human Behavior Understanding (HBU'10)*. Springer-Verlag, Berlin, Heidelberg, 14–25.

8. Masamichi Shimosaka, Keisuke Maeda, Takeshi Tsukiji, and Kota Tsubouchi. 2015. Forecasting Urban Dynamics with Mobility Logs by Bilinear Poisson Regression. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 535–546.

9. L. Sun and J. G. Jin. 2015. Modeling Temporal Flow Assignment in Metro Networks Using Smart Card Data. In *Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on*. 836–841.

10. Lijun Sun, Der-Horng Lee, Alex Erath, and Xianfeng Huang. 2012. Using Smart Card Data to Extract Passenger's Spatio-temporal Density and Train's Trajectory of MRT System. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing (UrbComp '12)*. ACM, New York, NY, USA, 142–148.

11. Jian Tang, Meng Qu, and Qiaozhu Mei. 2015a. PTE: Predictive Text Embedding Through Large-scale Heterogeneous Text Networks. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '15)*. ACM, New York, NY, USA, 1165–1174.

12. Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. 2015b. LINE: Large-scale Information Network Embedding. In *Proceedings of the 24th International Conference on World Wide Web (WWW '15)*. ACM, New York, NY, USA, 1067–1077.

13. Daisaku Yokoyama, Masahiko Itoh, Masashi Toyoda, Yoshimitsu Tomita, Satoshi Kawamura, and Masaru Kitsuregawa. 2014. A framework for large-scale train trip record analysis and its application to passengers ' flo prediction after train accidents. In *Advances in Knowledge Discovery and Data Mining*. Springer, 533–544.

14. Fuzheng Zhang, Nicholas Jing Yuan, Yingzi Wang, and Xing Xie. 2015. Reconstructing Individual Mobility from Smart Card Transactions: A Collaborative Space Alignment Approach. *Knowl. Inf. Syst.* 44, 2 (Aug. 2015), 299–323.

15. F. Zhang, J. Zhao, C. Tian, C. Xu, X. Liu, and L. Rao. 2016. Spatiotemporal Segmentation of Metro Trips Using Smart Card Data. *IEEE Transactions on Vehicular Technology* 65, 3 (March 2016), 1137–1149.

16. Juanjuan Zhao, Chen Tian, Fan Zhang, Chengzhong Xu, and Shengzhong Feng. 2014. Understanding temporal and spatial travel patterns of individual passengers by mining smart card data. In *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*. 2991–2997.

17. Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, and Wei-Ying Ma. 2008. Understanding Mobility Based on GPS Data. In *Proceedings of the 10th International Conference on Ubiquitous Computing (UbiComp '08)*. ACM, New York, NY, USA, 312–321.