

# SUME: Semantic-enhanced Urban Mobility Network Embedding for User Demographic Inference

FENGLI XU\*, ZONGYU LIN\*, and TONG XIA, Beijing National Research Center for Information Science and Technology, Tsinghua University

DIANSHENG GUO, Tencent Corporation

YONG LI<sup>†</sup>, Beijing National Research Center for Information Science and Technology, Tsinghua University

Recent years have witnessed a rapid proliferation of personalized mobile applications, which poses a pressing need for accurate user demographics inference. Facilitated by the prevalent smart devices, the ubiquitously collected mobility trace presents a promising opportunity to infer user demographics at large-scale. In this paper, we propose a novel Semantic-enhanced Urban Mobility Embedding (SUME) model, which learns dense representation vectors for user demographic inference by jointly modelling the physical mobility patterns and the semantic of urban mobility. Specifically, SUME models urban mobility as a heterogeneous network of users and locations, with various types of edges denoting the physical visitation and semantic similarities. Moreover, SUME optimizes the node representation vectors with two alternating objective functions that preserve the feature in physical and semantic domains, respectively. As a result, it is able to capture the effective signals in the heterogeneous urban mobility network. Empirical experiments on two real-world mobility traces show the proposed model significantly out-performs all state-of-the-art baselines with an accuracy margin of 8.6%~14.3% for occupation, gender, age, education and income inference. In addition, further experiments show SUME is able to reveal meaningful correlations between user demographics and the mobility patterns in spatial, temporal and urban structure domain.

CCS Concepts: • **Information systems** → **Data mining**; • **Human-centered computing** → *Empirical studies in ubiquitous and mobile computing*;

Additional Key Words and Phrases: human mobility, representation learning, user profiling, demographic inference.

## ACM Reference Format:

Fengli Xu, Zongyu Lin, Tong Xia, Diansheng Guo, and Yong Li. 2020. SUME: Semantic-enhanced Urban Mobility Network Embedding for User Demographic Inference. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 98 (September 2020), 25 pages. <https://doi.org/10.1145/3411807>

## 1 INTRODUCTION

Achieving accurate and large-scale user profiles lies at the heart of the increasingly popular personalized applications [1, 2]. The recently available large-scale user behavioral data provides a novel angle for the potential solutions, which draws attention to two key problems: identifying suitable data sources [3] and designing

\*Both authors contributed equally to the paper.

<sup>†</sup>This is the corresponding author.

---

Authors' addresses: Fengli Xu, [fenglixu2020@hotmail.com](mailto:fenglixu2020@hotmail.com); Zongyu Lin; Tong Xia, Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing; Diansheng Guo, Tencent Corporation, Beijing; Yong Li, [liyong07@tsinghu.edu.cn](mailto:liyong07@tsinghu.edu.cn), Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2020 Association for Computing Machinery.

2474-9567/2020/9-ART98 \$15.00

<https://doi.org/10.1145/3411807>

compatible behavior models [4]. On one hand, the correlations between user profiles and their mobility patterns have been widely observed and extensively validated [5, 6]. On the other hand, the localization modules in the prevalent smart devices provide fine-grained mobility data that can potentially scale to population level. Under this context, we aim to design a novel user profiling model that can effectively leverage the mobility data to infer user's demographics, i.e., occupation, gender, age, education and income level.

However, due to the intrinsic complex correlation, it is a non-trivial task to harness the power of mobility data to infer user demographics. Most of the previous work combined mobility data with various other mobile sensory data, such as app usage [3], user-generated content [7] and light sensor data [1], to improve the model performance. However, the additional data sources requirement significantly limits the generalizability of these methods, since they are hard to collect and might raise more privacy concerns. On the other hand, the previous works solely relied on mobility data did not fully exploit the semantic information of urban mobility [6, 8], i.e., modelling urban mobility as transition between physical locations instead of semantic-aware point of interest (POI), such as home and workplace. As a result, it limits the performance of these models.

Inspired by the limitations of previous works, we aim to design a novel semantic-enhanced urban mobility embedding (SUME) model, which is dedicated to learning effective user representations to facilitate accurate user profiling. The designed model consists of three main components. First, we propose to model the urban mobility as a heterogeneous information network, which includes two types of nodes, i.e., user and location. In addition, there are three types of edges among them, where the edge between user and location describes the visitation frequency. The edges among users and among locations denotes their similarity in semantic domain, which measure the similarity in time allocation pattern (i.e. how users dissect their time into smaller chunks) and POI distribution (e.g. number of workplaces and residences), respectively. Such data structure allows us to jointly model the urban mobility patterns and the underlying semantic information. Second, we design a network embedding algorithm to automatically learn a dense representation vector for each user and location node, which is optimized to preserve user's similarity in urban mobility patterns. That is the users with similar visitation patterns across urban regions and the urban regions with similar visiting users will have more similar representation vectors. Therefore, based on the assumption that users with similar profile will have more similar mobility patterns, the learned representation vectors can serve as a important feature for user profiling. Third, simply modelling the similarity in physical mobility patterns might result in an important limitation in leveraging the semantic information. For example, two users with similar profiles might exhibit identical transition patterns among POIs, but have very different representation vectors because they are active in different parts of the city and do not visit similar physical regions. To tackle this limitation, we further propose a heterogeneous network embedding framework to jointly consider their similarity in physical mobility patterns and semantic domain. Specifically, in addition to the similarity in physical visitations, the framework also optimizes the node representation vectors to preserve the similarity in user's time allocation pattern and urban region's POI distribution. By alternating the optimization objective between the physical domain and semantic domain, the proposed framework can effectively propagate relevant signals through the heterogeneous edges in urban mobility network, and learn better representations for user profiling.

To evaluate the effectiveness of the proposed model, we leverage two real-world urban mobility datasets, which are associated with the ground truth of user demographics. By conducting extensive experiments on these real-world datasets, we find out the following important results. First, the proposed model, SUME, significantly outperforms all state-of-the-art baselines by an accuracy margin of 8.6%~14.3% for occupation, gender, age, education and income inference ( $p < 0.01$ ,  $ES > 0.8$ , two-tailed Student's t-test). Second, through a careful ablation study, we show that consistent performance gains are achieved by integrating the semantic information in urban mobility, i.e., the similarity in user's time allocation pattern and urban region's POI distribution. This result justifies the model design that aims to incorporate semantic information, and showcases the proposed model can effectively harness such power. Third, our further experiments analyze our model's robustness across different

user groups. Specifically, there is no significant performance trend across user groups with different mobility characteristics. However, the performance is correlated with other user demographic feature, e.g, the prediction accuracy on female user's occupation is significantly higher than the male's ( $p < 0.001$ , two-tailed Student's t-test). Finally, based on the learned user representation vectors, we show our model can reveal meaningful correlations between user demographic and urban mobility patterns in spatial, temporal and urban structure domain. For example, the signature of users in education category (i.e., students and teachers) is that: they frequently visit university area, but are less likely to visit central business area; they tend to stay in one location during working days, and only visit different locations during daytime in non-working days; they are significantly more likely to visit recreation and education type POIs than general population.

To conclude, the contributions of this work can be summarized as the following three aspects:

- To the best of our knowledge, we are the first to propose a semantic-enhanced urban mobility embedding (SUME) model for user profiling. It models urban mobility behavior as a heterogeneous information network, which is able to capture the signal in physical mobility patterns and the semantic information of urban mobility simultaneously.
- We conduct extensive experiments to evaluate the model performance on a large-scale urban mobility trace along with user demographics. We demonstrate the superiority of the proposed model in terms of achieving significant performance gain against all up-to-date baselines, and having consistent performance across user groups with different mobility characteristics.
- Through an in-depth model analysis, we showcase the proposed model also reveals meaningful patterns in all spatial, temporal and urban structure domain. These results allow us to gain deeper insights on the underlying mechanisms of urban mobility, i.e., how users with different profiles move in the urban space. Such knowledge is of great importance to wide range of applications, including travel survey [9] and urban planning [10].

The rest of this paper is organized as follows. We formally define the problem setting in Section 2, and describe the proposed model in Section 3. After that, we apply our model on a real-world mobility dataset and conduct extensive experiments to evaluate its performance in Section 4. We conduct in-depth analysis to understand captured correlations between user demographic and mobility patterns in Section 5. After systematically reviewing the related works and the limitations of our study in Section 6, we finally conclude our paper in Section 7.

## 2 MOTIVATION AND PROBLEM DEFINITION

### 2.1 Problem Motivation

The recent trend of personalized mobile applications gives rise to the need of large-scale user profiling [1, 2]. Informed by accurate user demographics, application vendors are able to better meet their user's need. For example, location-based services might suggest more suitable entertainment areas to their users by knowing their age groups [11], while mobile e-commerce platforms can improve their recommender system based on the gender information [12]. Previous works in this area mostly focus on leveraging the semantic-aware user behavior data, such as application usages [13], web browsing records [14], and social media profile [15]. For example, a recent survey shows mobile application usage can be exploited to infer user income level and age groups with 64% and 74% accuracy [16], which is beneficial for mobile application development. Due to the prevalence of mobile devices, mobility data has become the most ubiquitously collected personal data, which makes it extremely valuable to large-scale applications compared with other data sources with limited access. However, its application in user demographic inference is woefully inadequate, which is mainly because of its lack of semantics. On the other hand, one recent study showed human experts can accurately identify the important semantic information (e.g., workplace, home and transportation mode) of an individual by solely observing his/her mobility trajectory [17]. Moreover, researchers also showed users with different demographics

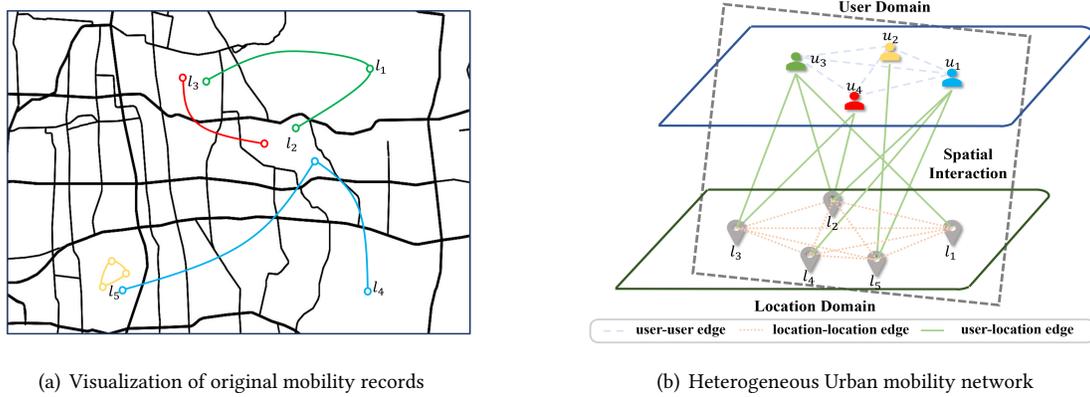


Fig. 1. Visualization of reorganizing mobility records into heterogeneous urban mobility network.

(e.g., gender and age) tend to exhibit different mobility patterns in various domains [6, 18, 19]. These observations demonstrate the feasibility to extract useful information from mobility data for user demographic inference. However, previous works on mobility data mining focus on capturing the similarity in physical space [20], which overlook the important semantics information in urban context [7, 21]. Inspired by the limitations of previous works, we aim to propose a semantic-enhanced urban mobility model for accurate user demographic inference.

## 2.2 Problem Definition

Now, we formally introduce the parameter notations and problem definition as follows.

**Definition 2.1. (Mobility record)** A mobility record is a triplet  $(u, l, t)$ , which denotes that user  $u$  visits location  $l$  at time  $t$ , where  $l$  denotes a unique area with geographical coordinates (i.e., longitude and latitude) and boundary in the urban space.

**Definition 2.2. (POI)** A POI  $P$  is defined as a uniquely identifiable venue with specific function  $C$ , e.g., residence, workplace and park. In our model, each location has a POI distribution vector which represents the number of different types of POIs within that location.

Figure 1(a) gives an example of the original mobility records, where different colors represent different users. For example, the blue trajectory indicates that a user has visited the locations  $l_5$ ,  $l_2$  and  $l_4$ . It captures how users mobility behavior within the urban space.

**Definition 2.3. (User–Location Mobility Network)** It is denoted as  $G_{ul} = (U \cup L, E_{ul})$ , an undirected bipartite network where  $U$  is the set of user vertices,  $L$  is the set of location vertices, and  $E_{ul}$  is the set of edges between users and locations. If user  $u_i$  visits location  $l_j$ , there will be an edge  $e(u_i, l_j)$  between them. The weight  $w(u_i, l_j)$  is set by the frequency that user  $u_i$  visits location  $l_j$ .

**Definition 2.4. (User–User Similarity Network)** It is denoted as  $G_{uu} = (U, E_{uu})$ , where  $U$  is the set of user vertices, each representing a user, and  $E_{uu}$  is the set of edges between them. The edge weight  $w(u_i, u_j)$  represents the similarity of user  $u_i$  and  $u_j$  in semantic space, where  $w(u_i, u_j) \geq 0, \forall u_i, u_j \in U$ .

**Definition 2.5. (Location–Location Similarity Network)** It is denoted as  $G_{ll} = (L, E_{ll})$ , where  $L$  is the set of location vertices, each representing a location, and  $E_{ll}$  is the set of edges between them. The edge weight  $w(l_i, l_j)$  represents the semantic similarity between location  $l_i$  and  $l_j$ , where  $w(l_i, l_j) \geq 0, \forall l_i, l_j \in L$ .

**Definition 2.6. (Heterogeneous Urban Mobility Network)** It is denoted as  $G = (L \cup U, E_{ul} \cup E_{uu} \cup E_{ll})$ , where  $U$  is the set of user vertices and  $L$  is the set of location vertices. Heterogeneous Urban Mobility Network  $G$  is a combination of  $G_{ul}$ ,  $G_{uu}$  and  $G_{ll}$ , with edge weights  $W(u_i, l_j)$ ,  $W(u_i, u_j)$  and  $W(l_i, l_j)$  corresponding to the edges in these network, respectively.

Specifically, Figure 1(b) showcases the data structure of heterogeneous urban mobility network. The green lines between users and locations depicts the user–location mobility network, which depicts the spatial interaction between user and location. In addition, the blue dashed line and orange dashed line denotes the semantic similarity in user domain and location domain. By organizing the urban mobility as a heterogeneous urban mobility network, we can simultaneously identify the user pairs with similar physical mobility patterns by finding 2–hop neighbors in user–location network as well as the user pairs and location pairs with high semantic similarity by finding the immediate neighbors in user–user network and location–location network respectively. That is we can jointly consider the physical mobility behavior and semantic similarity in urban mobility, which facilitates us to model the complex correlation between urban mobility and user demographic.

Finally, we formally define our problems as follows: given heterogeneous urban mobility networks  $G$ , we aim to learn a projection function  $\Phi$  that projects the user vertices  $u \in U$  into an embedding vector in low-dimensional space  $R^d$ . In this space, users with similar demographic are closer to each other than those who are less similar. The derived embedding vectors should be able to facilitate accurate user demographic inference, and reveal the underlying correlations between user demographic and urban mobility patterns. With the problem formally defined, we are ready to dive into the model design.

### 3 METHOD

Inspired by the phenomenal success of representation learning, we are dedicated to designing a novel network embedding algorithm that learns low-dimensional embedding vectors from the heterogeneous urban mobility network for user demographic inference. However, existing network embedding algorithms mostly focus on homogeneous network [22, 23]. More importantly, they usually are optimized to preserve predefined network metrics, such as first-order proximity [22], second-order proximity [23] and community structure [24], which are not customized for user demographic inference. Therefore, our algorithm needs to address three key obstacles: 1) how to construct the heterogeneous urban mobility network for user demographic inference? 2) how to design optimization metrics to preserve physical mobility pattern and semantic information? 3) how to jointly optimize for these two metrics?

Motivated by these challenges, we first design the user-location network and propose a network embedding algorithm that is able to preserve the physical mobility pattern. Then, we describe how to capture the semantic information into urban mobility. Finally, we propose a unify framework to jointly capture the physical mobility pattern and semantic information in user embedding.

#### 3.1 Urban Mobility Network Embedding

The key intuition behind our model design is the users who share similar urban footprints (frequency distribution among locations) are more likely to have similar demographics [6]. The edge weight  $w(u_i, l_j)$  between user  $u_i$  and location  $l_j$  is set as the frequency that user  $u_i$  visits location  $l_j$ . Therefore, implementing the intuition is equivalent to preserving user’s structural role with respect to urban locations in the embedding vector. That is user vertices that have similar edge weight distribution on location vertices should have similar embeddings. The learning process is illustrated in Figure 2 (a). Specifically, given the embedding vector of user  $u_i$  as  $\vec{u}_i$  and the context vector of location  $l_j$  as  $\vec{l}_j$ , we define the predicted normalized edge weight between them as  $p_m(l_j|u_i)$ , which is computed by:

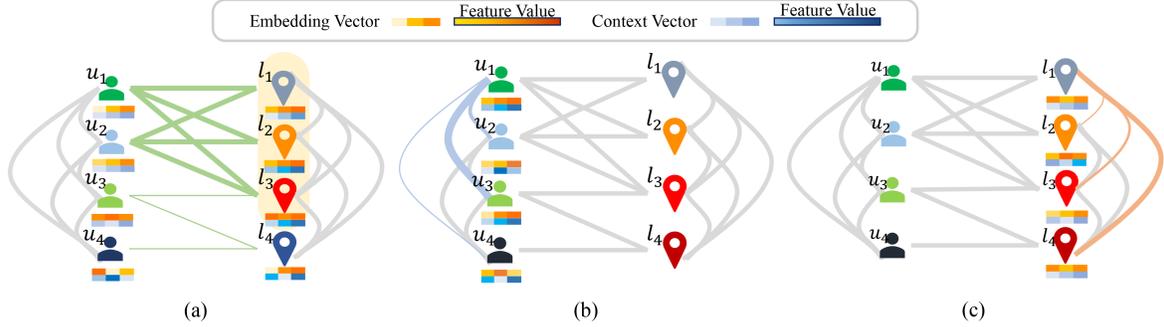


Fig. 2. Illustration of semantic-enhanced urban mobility embedding. Different edge colors denotes different types of edges, with thicker edges represent higher weight. (a) Preserving physical mobility patterns, e.g., vertex  $u_1$  and  $u_2$  should be projected to similar embedding vectors as they share similar neighbors in user-location network. (b) Preserving semantic similarity among users, e.g., vertex  $u_1$  and  $u_3$  should be closely located in the embedding space as they have a strong edge between them. (c) Preserving semantic similarity among locations, e.g., vertex  $l_1$  and  $l_4$  should also be placed closely as they have a strong edge between them.

$$p_m(l_j|u_i) = \frac{\exp(\vec{l}'_j \cdot \vec{u}_i)}{\sum_{l_k \in L} \exp(\vec{l}'_k \cdot \vec{u}_i)}, \quad (1)$$

where  $|L|$  represent the number of location vertices. For each user vertex  $u_i$  in set  $U$ , Eqn.(1) computes the predicted normalized edge weight  $p_s(l_\star|u_i)$  over the entire set of location vertices in  $L$ . To preserve the structural role, we should optimize  $p_s(l_\star|u_i)$  to fit the empirical normalized edge weight distribution  $\hat{w}(l_\star|u_i)$ , which is computed as  $w(l_\star|u_i)/\sum_k w(l_k|u_i)$ . Therefore, we minimize the following objective function:

$$O_m = \sum_{u_i \in U} \alpha_i d(\hat{w}(l_\star|u_i), p_m(l_\star|u_i)), \quad (2)$$

where  $d(\star, \star)$  is defined as the KL-divergence between two distributions, and  $\alpha_i$  is the out degree of user vertex  $u_i$ , which denotes its importance in network. To avoid iterate through user-location pairs that have no edge, we can prove that optimizing Eqn.(2) is equivalent to optimizing the following objective function:

$$O_m = - \sum_{e(u_i, l_j) \in E_{ul}} w(u_i, l_j) \log(p_m(l_j|u_i)). \quad (3)$$

It allows us to only iterate through the user-location pairs with existing edges. By iteratively optimizing the embedding vectors for user vertices, we are able to represent every user vertex  $u_i \in U$  with a  $d$ -dimensional vector  $\vec{u}_i$  that preserve the similarity in physical mobility pattern. For example, user  $u_1$  visits location  $l_1$  three times,  $l_2$  three times and  $l_3$  three times, then his normalized frequency distribution vector among urban locations will be  $[1/3, 1/3, 1/3, 0]$ . Likewise, the normalized frequency distribution vectors for user  $u_2, u_3, u_4$  are  $[1/3, 1/3, 1/3, 0], [0, 0, 1/2, 1/2], [0, 0, 0, 1]$ , respectively. After calculating the KL-divergence between these vectors, we find that the user  $u_2$  is closest to user  $u_1$ , while the user  $u_4$  is farthest to user  $u_1$ . Therefore,  $u_1$  and  $u_2$  will have more similar embedding vectors, which accurately preserves the similarity in physical mobility patterns. The embedding for location vertices can be derived with identical process.

### 3.2 Capturing the Semantic Similarity in Urban Mobility

First, we construct a user-user network to capture the semantic similarity between users. We draw inspiration from a recent study where researchers find urban mobility data can be classified into meaningful clusters based on the time allocation patterns [8]. Specifically, the time allocation patterns are represented as how the users partition their time among different locations. The overall time duration is first segmented into a set of evenly distributed time slices  $C$ . Then, the time slices are partitioned into same subsets if users are in same location during the corresponding time slots, and overall partition scheme is represented as  $P^m$  for user  $m$ . By measuring the similarity in partition scheme, it allows us to model the semantic similarity between user's mobility behavior without the interference of physical proximity, i.e., distant users with similar time allocation patterns can be effectively detected. The underlying assumption is that two users with similar time allocation patterns are more likely to share similar life styles, and hence are more likely to have the similar demographics. Specifically, we use the partition distance to quantify the divergence in time allocation patterns. Given two partition schemes  $P^m$  and  $P^n$  of a collection of time slices  $C$  which represent two users' time allocation patterns, the partition distance between  $P^m$  and  $P^n$  is defined as follows [25],

**Definition 3.1. (Partition Distance  $pd(\star, \star)$ )** The *partition distance*  $pd(P^m, P^n)$  between  $P^m$  and  $P^n$  is computed as the minimum number of time slices that must be removed from  $C$ , so that the two induced partitions are identical.

Higher partition distance indicates two users have more different time allocation patterns. Therefore, we use the inverse of partition distance to determine the edge weight in user-user network:  $w_{u_i, u_j \in U} = 1/pd(u_i, u_j)$ , which measure the semantic similarity among users. Note that this similarity metric allows us to capture both temporal and spatial information, since it measures time allocation patterns between users.

On the other hand, we leverage the location's POI distribution to measure the semantic similarity among them. Specifically, each location is associated with a POI distribution vector, which represents the number of different types of POIs within that location. Previous studies demonstrated the urban locations with similar POI distribution usually have similar urban function [10, 21]. Therefore, the location's similarity in POIs distribution can measure their semantic in urban mobility. As a result, we set the location-location edge weight  $w(l_i, l_j)$  as the cosine similarity of two location's POI distributions.

In both user-user network and location-location network, the edge weight denotes the semantic similarity between two vertices, where higher edge weight indicates the connected vertices are more similar in semantics. Therefore, modelling the semantic similarity is equivalent to preserving the proximity in location-location and user-user network. That is the vertices pairs that are connected with high weight edges should have similar embedding vectors. To achieve this goal, we propose an additional objective function. Since optimizing embedding vectors of location vertices is identical with user vertices, we only describes the algorithm of learning user embeddings to avoid redundancy. We first define the predicted edge weight between user vertex  $u_i$  and  $u_j$  as follows:

$$p_s(u_i, u_j) = \frac{1}{1 + \exp(-\vec{u}_i^T \cdot \vec{u}_j)}, \quad (4)$$

where user vertices with similar embedding vectors are predicted to have higher edge weight, which is consistent with our objective. Following this equation, we can compute predicted edge weight distribution  $p_s(u_\star, u_\star)$  over the space  $U \times U$ . The empirical distribution can be computed as  $\hat{w}(u_i, u_j) = w(u_i, u_j) / \sum_{(u_i, u_j) \in E_{uu}} w(u_i, u_j)$ . To optimize the embedding vectors, a natural way is to minimize the KL-divergence of two probability distributions as follow:

$$O_s = d(\hat{w}(u_\star, u_\star), p_s(u_\star, u_\star)), \quad (5)$$

Similarly, the objective function is equivalent to iterative through existing edges with the following equations:

$$O_s = - \sum_{(u_i, u_j) \in E_{uu}} w(u_i, u_j) \log p_s(u_i, u_j). \quad (6)$$

The illustration of node embedding algorithm in user–user network and location–location network is shown in Figure 2 (b) and (c). Take the user-user network as an example, the more similar two users' time allocation patterns are, the more similar embedding vectors they will have, which as a result facilitates the user demographic inference.

### 3.3 Semantic-enhanced Urban Mobility Embedding

Finally, we introduce the unified learning framework, i.e., Semantic-enhanced Urban Mobility Embedding (SUME). The goal is to optimize the vertex's embedding vectors to capture the feature from all types of edge in heterogeneous urban mobility network simultaneously. Specifically, we have described the standalone embedding algorithms on all three sub-networks, i.e., user-user network, location-location network and user-location network. An intuitive approach is to optimize the combination of the corresponding objective functions, which can be formulated as following:

$$O = \lambda_{ul} O_{ul} + \lambda_{uu} O_{uu} + \lambda_{ll} O_{ll}, \quad (7)$$

where

$$\begin{aligned} O_{ul} &= - \sum_{(u_i, l_j) \in E_{ul}} w(u_i, l_j) \log(p_m(l_j | u_i)), \\ O_{uu} &= - \sum_{(u_i, u_j) \in E_{uu}} w(u_i, u_j) \log p_s(u_i, u_j), \\ O_{ll} &= - \sum_{(l_i, l_j) \in E_{ll}} w(l_i, l_j) \log p_s(l_i, l_j), \\ \lambda_{ul} + \lambda_{uu} + \lambda_{ll} &= 1, 0 < \lambda_{ul}, \lambda_{uu}, \lambda_{ll} < 1. \end{aligned} \quad (8)$$

To be specific,  $O_{ul}$  describes the objective to capture physical mobility patterns, while  $O_{uu}$  and  $O_{ll}$  describe the objective to capture semantic similarity in user domain and location domain respectively. In addition,  $\lambda_{ul}$ ,  $\lambda_{ll}$  and  $\lambda_{uu}$  denote the importance of them. However, optimizing  $O$  is computationally expensive, since the evaluation of  $O_{ul}$  requires to summing over the entire vertices set to compute  $p_m(l_\star | u_i)$  for each vertex. To address this problem, we adopt the negative sampling technique to accelerate the optimization [26]. In the optimization of each vertex we sample one positive edge according to its edge weight distribution and multiple negative edges according to noisy distribution. We denote the optimized vertex as  $v_i$ , noisy distribution as  $P_n(v)$  and number of negative edges as  $K$ . Without loss of generality, the  $v_i$  can represent either user vertex or location vertex when optimizing for different vertices. Then, the update gradient for its embedding  $\vec{v}_i$  can be computed as:

$$\log \sigma(\vec{v}_j^T \cdot \vec{v}_i) + \sum_{n=1}^K E_{v_n \sim P_n(v)} [\log \sigma(-\vec{v}_n^T \cdot \vec{v}_i)], \quad (9)$$

where  $\sigma(x) = 1/(1 + \exp(-x))$  is the sigmoid function,  $v_j$  and  $v_n$  are connected vertices on positive edge and negative edge respectively. Following the widely adopted empirical setting [26], we set  $K = 5$  and  $P_n(v) \propto \alpha_v^{3/4}$ , where  $\alpha_v$  is the degree of vertex  $v$ . The key idea of negative sampling is to approximate the gradient of objective function by sampling a small amount of edges, where the first term in Eqn.(9) corresponding to gradient from existing edges while the second term corresponding to absent edges. Following the idea of negative sampling, we can optimize the overall objective function 7 in a edge sampling manner. That is, in each optimization step we first

**ALGORITHM 1:** Semantic-enhanced Urban Mobility Embedding Algorithm

---

**Input:**  $G_{ul}, G_{uu}, G_{ll}$ , number of total samples  $N$ , number of negative samples  $K$ , list of importance coefficient  $\lambda = [\lambda_{ul}, \lambda_{ll}, \lambda_{uu}]$ .

**Output:** user embeddings  $\vec{u}$ , user contextual embeddings  $\vec{u}_c$ , location embeddings  $\vec{l}$ , location contextual embeddings  $\vec{l}_c$ .

Initialize edge type list:  $\mathbf{H} \leftarrow [ul, ll, uu]$ ;

**while**  $iter \leq N$  **do**

- Sample edge type:  $h \leftarrow \mathbf{H} \sim \lambda$ ;
- Draw a positive edge:  $e_p \leftarrow E_h \sim \hat{w}_h$ ;
- Draw  $K$  negative edges:  $\{e_n\} \leftarrow E_h \sim P_n(v)$ ;
- Update vertex embedding according to Eqn.(9);

**end**

---

sample an edge type based on relative importance (i.e.  $\lambda_{ul}, \lambda_{ll}$  and  $\lambda_{uu}$ ), and then optimize the vertices embedding in corresponding sub-network with negative sampling algorithm in Eqn.(9). The complete optimization algorithm is shown in Algorithm 1. Since the context vectors are equivalent to the embedding vectors, we average the embedding vectors and the context vectors as the final output embedding vectors.

By jointly optimizing these three objective functions, we can effectively propagate the important signals through the heterogeneous edges. For example, user  $u_1$  and user  $u_2$  never visits the same locations, but the locations they frequently visit share similar POI distributions. Then, their embedding vectors will still be similar, because the locations they frequently visit have similar embeddings due to the objective in location-location network. As a result,  $u_1$  and  $u_2$  are assigned similar embedding vectors due to the signals propagate through user-user edge and user-location edge. Therefore, our SUME model is able to capture features not limited in physical mobility pattern or semantic similarity, but combining the strength from both of them. It is worth pointing out that the proposed SUME model focus on extracting low-dimensional embedding vectors, which can be readily fed into off-the-shelf classifier models for user demographic inference, such as SVM and random forest. We aim to extract quality representation from urban mobility data for user demographic inference, and thus the downstream classifier models are beyond our scope. In addition, our SUME model can be easily generalized to online mode, where the embedding of newly added users can be computed by only updating the edges connecting them for a few epochs.

## 4 EXPERIMENT

### 4.1 Dataset

We leverage two large-scale real-world mobility trace along with user demographics data to evaluate our models. Their descriptions are as follow.

**4.1.1 Social Media Data.** This dataset is released by a recent study [27, 28]. Specifically, the dataset is collected by the local mobile operator in Shanghai, and contains more than 7 million mobility records covering 602 users and 13,496 base stations for one week, i.e., from 20 Apr. 2016 to 26 Apr. 2016. It records the time stamps and the connected base stations whenever the mobile users access cellular network, e.g., consuming mobile data traffic, sending text and making phone calls, which constitutes a fine-grained mobility trace. On the other hand, it also includes the public profiles users post on the mainstream social media platform in China, i.e., Weibo. The basic information is summarized in Table 1. Specifically, it provides the gender and occupation information for the users. We present the user demographic distribution in Table 2.

Table 1. The basic information of the utilized datasets.

Source	City	Time Duration	Records	Users	Locations
Social media	Shanghai	20th-26th, April, 2016	7,704,708	602	13,496
Social network	Beijing	1st, October-31st, December, 2019	208,795,352	39,999	10,656

Table 2. User demographic distribution in *social media* dataset.

Demographic	Category
Occupation	white collar(61.00%), education(19.74%), sports and health(10.65%), media(5.06%), civil servant(3.56%)
Gender	male(59.97%), female(40.03%)

Table 3. User demographic distribution in *social network* dataset.

Demographic	Category
Occupation	professionals(66.44%), sales workers(12.57%), transport and production(7.41%), administration support(6.85%), healthcare and technicians(4.48%), services(1.30%), managers(0.94%)
Gender	male(58.67%), female(41.32%)
Education	primary school(7.18%), junior high school(6.66%), senior high school(22.13%), undergraduate(51.79%), postgraduate(12.24%)
Income	very low(3.36%), low(14.84%), medium(55.79%), high(24.57%), very high(1.44%)
Age	0~20(5.61%), 20~25(26.76%), 25~30(22.57%), 30~50(42.46%), 50~99(2.60%)

**4.1.2 Social Network Data.** Through close collaborations with Tencent corporation, we were granted access to a mobility trace collected from their social network applications. This dataset is large-scale in terms of covering 39,999 users and 10,656 locations in Beijing. It contains 209 million mobility records from 1 Oct. 2019 to 31 Dec. 2019. Whenever users invoke location-based services, e.g. check-in to a location, the localization module generate a record of user’s physical location. The basic information is summarized in Table 1. Through large-scale user survey, this dataset also collects user demographic information, including gender, age, education level, income level and occupation. The occupations are originally classified based on the standard occupational classification (SOC) system [29]. Since some of the categories are absent in urban space, like farming and military, we merge them into seven categories following the setting in [30]. This occupation taxonomy covers all the main job categories in urban space. Besides, we know most of them have distinct commuting behaviors based on our common sense and previous works [30], which can be compared against our model’s performance. We present the user demographic distribution in Table 3. Note that the age groups are split unevenly to make the user distribution more even, and also to better represent different life stages, such as schooling, early career and retirement.

**4.1.3 Ethical Considerations.** Given the sensitivity of the data, we enforce the following protocols to address the privacy and ethical risks in data sharing and analysis. First, all data is properly anonymized by the data owners before they are shared to us. The real user IDs are never made available to or utilized by the researchers. In addition, our data analysis procedures are reviewed and authorized by the dataset owners to ensure the compliance with privacy protocols in the Term-of-Use statements. Second, all the researchers that have been authorized to access the datasets are bounded by strict non-disclosure agreements, and our research protocol are approved by the local institutional board. Third, we store all the data in a secure off-line server, and only the authorized core researchers can access the data.

**4.1.4 Basic Statistics.** To provide a comprehensive understanding on the utilized datasets, we present the distribution of mobility records, radius of gyration and entropy in Figure 3. Take the social media dataset as an

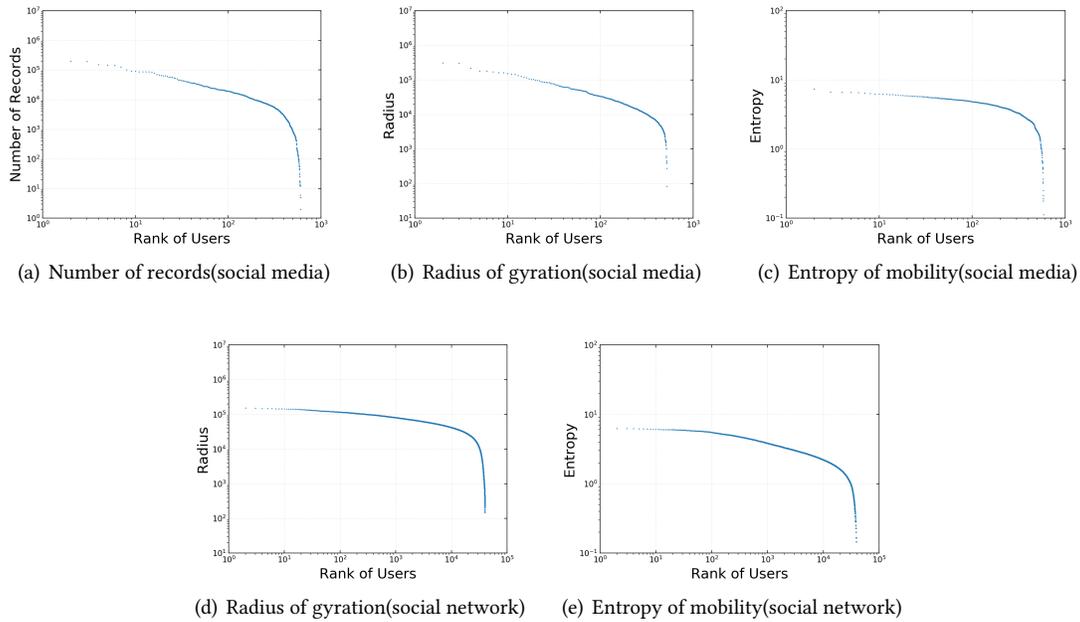


Fig. 3. The distribution of mobility characteristics among user population.

example, we plot the number of mobility records distribution in Figure 3(a), with the users sorted by the number of records. We can observe that the empirical data follows power-law distribution with an exponential cut-off. 81% users (448) have more than 1000 mobility records. In addition, we examine the mobility activity area as the radius of gyration [31], which is computed as the user's root mean square displacement from his average location through all his mobility records. From Figure 3(b), we can observe that the radius of gyration follows similar power-law distribution, with 304 users having less than 13,000 meter radius. On the other hand, we showcase the entropy of mobility behavior in Figure 3(c), which measures the uncertainty of user's mobility across all the locations. We observe more than 400 users have entropy above 2.5. Similar observations are made in social network dataset. Since each user in this dataset has identical number of mobility records, we only present the distribution of radius of gyration and entropy in Figure 3(d) and 3(e).

## 4.2 Experiment Settings

**4.2.1 Baselines.** We select numerous baselines to compare against our method. Specifically, the baselines can be summarized in three categories: classic models, deep learning models and variants of our model.

First, we implement two classic methods to serve as benchmarks: random guess and raw feature.

- **Random guess.** The most basic method that sets a naive benchmark for the prediction task. It randomly chooses an occupation and gender for each user according to the popularity, i.e., frequency distribution.
- **Raw feature.** The classic feature based method. The feature vector of each user is constructed as the frequency distribution over all locations.

We also compare our method with three state-of-the-art network embedding models and two graph neural network models. These models leverage representation learning technique to learn embedding vectors for vertices. Therefore, they are strong competitors to our model.

- DeepWalk [22]. It is a classic network embedding method, which uses skip-gram model [32] to learn embedding vectors for vertices based on the vertices sequence sampled by truncated random walk.
- Node2vec [33]. It generalizes DeepWalk by adopting customized biased random walks, which achieves a better trade-off between exploration and return.
- LINE [23]. It is a network embedding method that learns embedding on a weighted network to represent both first and second order proximity. There are three variants: LINE(1st), LINE(2nd) and LINE(1st + 2nd), which preserve first-order proximity, second-order proximity and both of them, respectively.
- Graph Convolutional Network (GCN) [34]. It is a classic graph neural network model, which learns embedding for each node by iteratively aggregating information from its neighbors. We use a two-layer GCN model (optimal setting), which is trained in an unsupervised manner to ensure fair comparison.
- Graph Attention Network (GAT) [35]. It is a state-of-the-art graph neural network model, which replaces the aggregator function in GCN with attention module. It allows the model to assign different weights to different neighbor nodes. We use a two-layer GAT model (optimal setting), which is trained in an unsupervised manner to ensure fair comparison.

The proposed SUME model has three variants, which stands for the embedding on the partial urban mobility network. The details of the variants are discussed as follows.

- SUME(ul). It is a variant of SUME, which only exploits the user-location network in the unified embedding framework.
- SUME(uu+ul). It is a variant of SUME, which exploits user-location network and user-user network in the unified embedding framework.
- SUME(ll+ul). It is a variant of SUME, which exploits the user-location network and the location-location network in the unified embedding framework.

Except for the random guess method, all the evaluated models generate an embedding vector for each user. Each baseline method follows the optimal implementation released by the authors, and the hyper parameters are tuned to optimal through grid search. For the network embedding and graph neural network baselines, we use the complete mobility network we construct for our model as the input graph. Besides, we use the node id embedding as the node features for the graph neural network baselines [34]. Since all the users in the evaluation datasets have ground truth information, we feed the learned embedding vectors into a supervised classification model for user demographics inference. Without loss of generality, we adopt the widely used support vector machine [36], which can be readily switched to other supervised and semi-supervised models.

**4.2.2 Evaluation Metric.** We adopt the accuracy, precision and recall as performance metrics for model evaluation, which can be computed as follow:

$$Accuracy = \frac{1}{M} \sum_{i=1}^M I(T_i \in S_i(k)), \quad Precision = \frac{1}{N} \sum_{j=1}^N \frac{TP_j}{TP_j + FP_j}, \quad Recall = \frac{1}{N} \sum_{j=1}^N \frac{TP_j}{TP_j + FN_j},$$

where  $N$  is the number of attribute categories, and  $M$  is the number of users. In addition,  $S_i(k)$  is the set of the top- $k$  attributes the model predicts user  $i$  exhibits, and  $T_i$  is the ground truth.  $TP_j, FP_j, FN_j$  are the true positive, false positive and false negative rates in the  $j$ -th category.  $I(T_i \in S_i(k))$  equals to 1 if the ground truth attribute is included in  $S_i(k)$ , otherwise it equals to 0. They are the most adopted metrics in real-world application [37], since they balanced measure model's performance from different aspects. Without loss of generality, we set  $k = 2$  for occupation, education, age and income inference and  $k = 1$  for gender inference respectively. To ensure

the robustness of the experiment results, we randomly split the datasets into 5 subsets, and report the average performance of 5-folds cross-validation. For the inference of each user demographic, we ensure it distributes uniformly in each subsets.

### 4.3 Performance Evaluation

**4.3.1 Performance Comparison with Baselines.** The overall experiment results are summarized in Table 4 and Table 5. For the proposed SUME model, we conduct statistical analysis to examine the significance of performance gain over the best baseline. Specifically, we use Student’s t-test (with Bonferroni correction) to compute the p-value  $p$  and effect size  $ES$  respectively, where  $p < 0.05$  and  $ES > 0.8$  indicate the performance gain is statistically significant [38]. From the experiment results, we have the following observations and conclusions.

Table 4. Performance comparison with baseline models on social media data, where (\*\*) indicates  $p < 0.05$  (with Bonferroni correction) and  $ES > 0.8$  over the best baseline.

Method	Occupation			Gender		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
random guess	0.41	0.37	0.38	0.53	0.48	0.49
raw feature	0.42	0.38	0.40	0.59	0.56	0.58
DeepWalk	0.50	0.43	0.43	0.63	0.58	0.60
Node2vec	0.47	0.42	0.43	0.62	0.57	0.55
LINE(1st)	0.53	0.47	0.48	0.65	0.59	0.61
LINE(2nd)	0.53	0.48	0.47	0.67	0.62	0.60
LINE(1st+2nd)	0.55	0.49	0.52	0.70	0.64	0.65
GCN	0.51	0.46	0.47	0.65	0.62	0.63
GAT	0.52	0.47	0.46	0.65	0.61	0.61
SUME(ul)	0.53	0.48	0.48	0.65	0.62	0.63
SUME(ul+uu)	0.57	0.51	0.52	0.71	0.65	0.68
SUME(ul+ll)	0.56	0.53	0.52	0.70	0.66	0.69
<b>SUME</b>	<b>0.61**</b>	<b>0.56**</b>	<b>0.57**</b>	<b>0.76**</b>	<b>0.71**</b>	<b>0.73**</b>

- 1) In both datasets and across different user demographics, there is no significant difference between random guess and raw feature method. The performance of random guess only degenerates slightly for all user demographic inference. On the other hand, all the representation learning methods, including DeepWalk, Node2vec, LINE, GCN and GAT, consistently achieve notable performance gains over the random guess and feature-based methods. A plausible reason is that the complex correlation between urban mobility and user demographics cannot be captured by the simple frequency distribution vector, which suggests the importance of learning expressive representations. In addition, our SUME model achieves further performance gain over the best representation learning methods. It indicates our model can effectively captures the physical mobility patterns and semantic information with the novel representation learning algorithm.
- 2) The proposed SUME model significantly outperforms all baseline models. In social media Dataset, it provides a relative performance gain of 42.5%~47.4% in occupation inference ( $p < 0.05$ ,  $ES > 0.8$ ) and 25.9%~28.8% in gender inference ( $p < 0.05$ ,  $ES > 0.8$ ), respectively. In addition, it also significantly outperforms the state-of-the-art representation learning method, Line(1st+2nd), by 9.6%~14.3% in occupation inference ( $p < 0.05$ ,  $ES > 0.8$ ) and 8.6%~12.3% in gender inference ( $p < 0.05$ ,  $ES > 0.8$ ). As for the social network

Table 5. Performance comparison with baseline models on social network data, where (\*\*) indicates  $p < 0.05$  (with Bonferroni correction) and  $ES > 0.8$  over the best baseline.

Method	Gender			Education			Income			Age			Occupation		
	Acc.	Prec.	Rec.												
random guess	0.50	0.48	0.50	0.40	0.40	0.38	0.40	0.40	0.39	0.40	0.40	0.40	0.29	0.29	0.28
raw feature	0.53	0.51	0.51	0.46	0.42	0.45	0.45	0.42	0.41	0.46	0.42	0.43	0.33	0.31	0.30
DeepWalk	0.55	0.53	0.53	0.60	0.54	0.54	0.59	0.53	0.52	0.63	0.59	0.58	0.43	0.38	0.38
Node2vec	0.54	0.51	0.52	0.59	0.53	0.51	0.58	0.53	0.54	0.62	0.57	0.57	0.43	0.39	0.38
LINE(1st)	0.61	0.56	0.58	0.71	0.62	0.65	0.71	0.62	0.63	0.67	0.64	0.62	0.49	0.45	0.44
LINE(2nd)	0.62	0.59	0.56	0.72	0.64	0.63	0.73	0.64	0.67	0.69	0.64	0.63	0.51	0.46	0.46
LINE(1st+2nd)	0.63	0.58	0.59	0.72	0.64	0.66	0.73	0.66	0.68	0.71	0.67	0.65	0.53	0.47	0.48
GCN	0.59	0.56	0.56	0.72	0.65	0.64	0.72	0.66	0.66	0.71	0.67	0.66	0.52	0.47	0.47
GAT	0.59	0.57	0.57	0.71	0.65	0.63	0.72	0.67	0.66	0.72	0.68	0.67	0.52	0.48	0.47
SUME(ul)	0.61	0.58	0.56	0.72	0.63	0.65	0.72	0.67	0.69	0.70	0.65	0.66	0.47	0.44	0.45
SUME(ul+ll)	0.65	0.59	0.59	0.73	0.67	0.64	0.73	0.66	0.70	0.73	0.68	0.68	0.53	0.48	0.49
SUME(ul+uu)	0.68	0.65	0.62	0.75	0.66	0.68	0.76	0.69	0.67	0.75	0.68	0.71	0.54	0.50	0.48
<b>SUME</b>	<b>0.72**</b>	<b>0.68**</b>	<b>0.66**</b>	<b>0.79**</b>	<b>0.72**</b>	<b>0.71**</b>	<b>0.80**</b>	<b>0.73**</b>	<b>0.73**</b>	<b>0.79**</b>	<b>0.74**</b>	<b>0.73**</b>	<b>0.59**</b>	<b>0.53**</b>	<b>0.54**</b>

dataset, SUME also significantly outperforms all the baselines across all performance metrics. In terms of accuracy, SUME outperforms the best baselines with a margin of 9.6%~14.3%. Moreover, the accuracy of income level and age inference reaches up to 80% and 79%. These results demonstrate the proposed SUME model consistently achieves preferable performance against state-of-the-art baselines, and is beneficial to real-world application.

- 3) By comparing the performance of the complete SUME model and the degraded versions, we observe consistent performance gains as more components are integrated into the framework. These comparisons indicate that each component of our model play its role in improving the user demographic inference. Also, these results suggest our model can effectively incorporate the semantic information in urban mobility and the physical mobility patterns.

To conclude, our SUME model consistently achieves preferable results compared with classic models and the state-of-the-art network embedding methods. In addition, each component of the proposed model leads to significant performance gain. These results combined to suggest the proposed SUME model's effectiveness in simultaneously capturing the feature of physical mobility patterns and semantic similarity, and therefore it is a powerful model to infer user demographic with mobility data.

**4.3.2 Performance Across Different User Groups.** We also conduct more experiments to evaluate how the model performance varies across user groups with different characteristics. To avoid redundancy, we only show the accuracy of gender and occupation inference on social network dataset. Similar observations are made on other demographics and social media dataset.

Specifically, we split the user population evenly into four groups based on the number of visited locations, radius of gyration and mobility entropy, respectively. The performance of occupation inference across different user groups is displayed in Figure 4(a), while the performance of gender inference is displayed in Figure 4(b). From Figure 4(a), we observe that the user group with the highest number of visited locations ranks first in prediction accuracy, which is probably because the features of user mobility can be better captured with more visited locations. Besides, Figure 4(b) shows similar trend in gender inference. In addition, both Figure 4(a) and Figure 4(b) demonstrate the accuracy of occupation and gender inference slightly decreases as the radius of gyration increases, which does not vary significantly. On the other hand, Figure 4(a) and Figure 4(b) show the inference accuracy of gender and occupation both gradually decrease with the increase of mobility entropy. These combine to suggest the demographic inference are easier in user population with smaller activity area and less

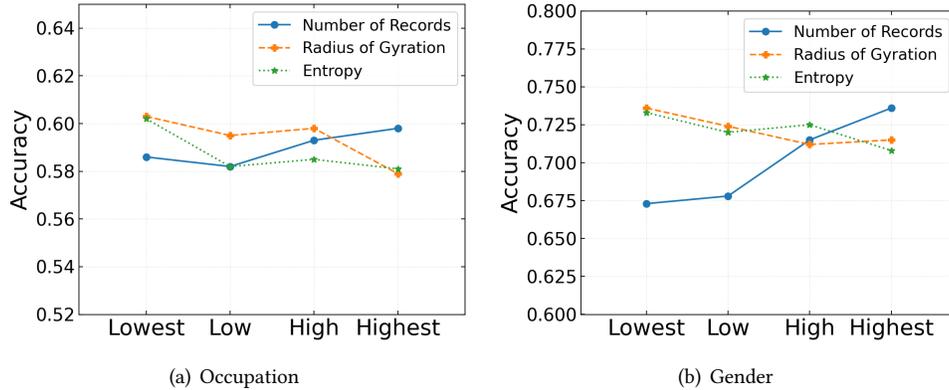


Fig. 4. Performance of gender and occupation inference across users groups with different mobility characteristics.

uncertain mobility behavior. To conclude, the proposed SUME model generally performs better in the user group with more visited locations, smaller radius of gyration and lower mobility entropy. However, the performance does not vary significantly across different user groups, which showcases the generalizability of our model.

## 5 REVEALING THE CAPTURED MOBILITY FEATURES FOR DEMOGRAPHIC INFERENCE

In this section, we aim to understand the mobility feature SUME extracted for user demographic inference. Since the demographic inference is based on the learned embedding vectors for users and locations, we can efficiently extract the most representative users and locations for different demographics by simply computing the similarity between embedding vectors. Specifically, we are dedicated to visualizing the most representative mobility patterns in temporal, spatial and POI visitation domains. Note that we aim to deepen our understandings on how SUME works for user demographics inference. On the other hand, comprehensive analysis on the mobility differences across user groups is beyond our scope. Therefore, we only present the case studies on the gender, occupation and age inference in social network dataset to avoid redundancy. Similar observations are made on other user demographics and social media dataset.

### 5.1 Temporal Domain Patterns

We first pick out the most representative temporal patterns among user population of different occupations, age group and gender, and then examine their characteristics by visualizing them in temporal domain. Specifically, previous studies have established theoretical and empirical analyses on the mobility differences among users with management, clerical and blue collar occupations as well as different genders and age groups [18, 30, 39]. Therefore, we conduct case studies on genders, age groups, and the user occupations of transport and production, manager, and administration support to better connect our findings with literature. The intuition behind the most representative temporal patterns selection is that it should be corresponding to the mobility trace of the “centroid users” among the user population with certain demographics, which is defined as the users with smallest root mean square distance with all other users within the population. Since we aim to examine the temporal patterns, we use the partition distance to measure the distance between users. Such procedure allows us to select the users with minimum overall distance to all users that are classified into certain demographics, i.e., the most representative ones. As for the visualization, we first divide and aggregate the mobility records into a typical working day and a typical non-working day, since they might exhibit different patterns. Then, we evenly segment

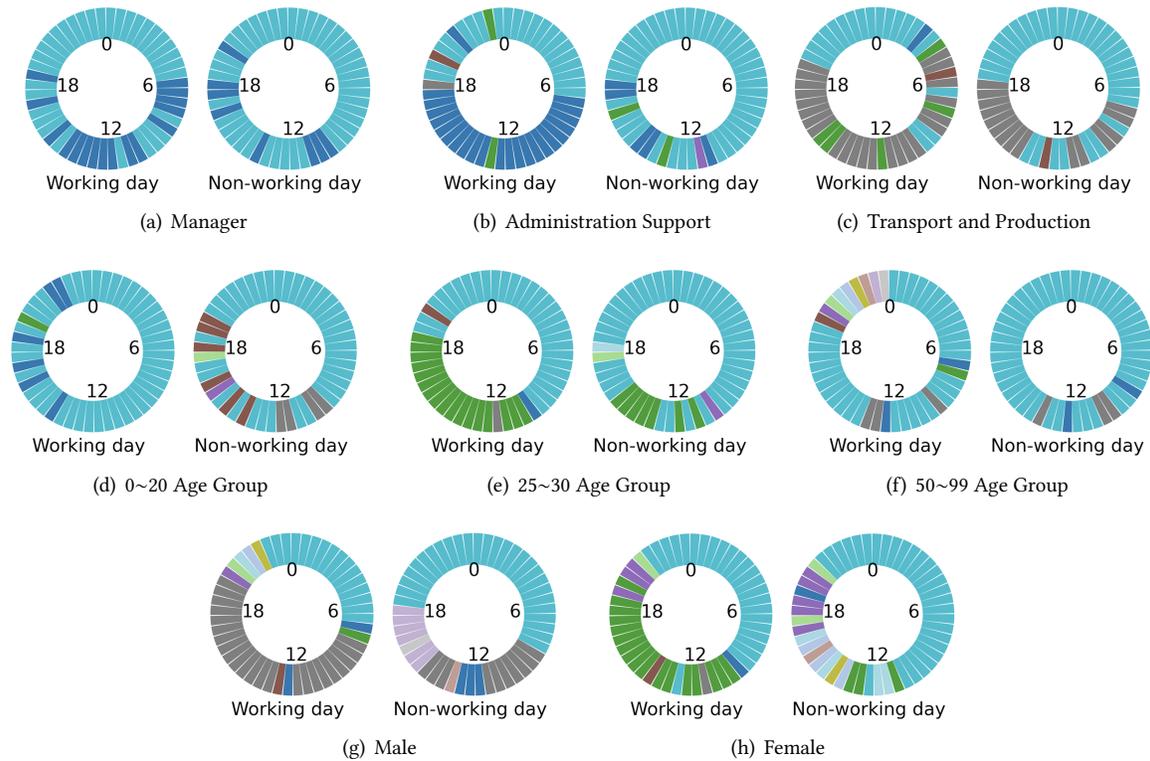


Fig. 5. The visualization of the most representative temporal patterns of different occupations, age groups and genders. Same color within each subgraph indicate users are associated with same locations, while the color scheme is not shared among different subgraphs.

the working day and non-working day into 48 time slices respectively, with each time slice denote a 30-minute period. We link each time slice to the most frequently visited location in that time period, and we visualize them in the form of clock dial as shown in Figure 5. Specifically, for each selected user, the time slices with same colors means they are linked to same locations, while the colors do not contain ordering information. Besides, same color only means same location for an individual user, which does not have a clear meaning among different users. We can observe that the most representative users of different demographics exhibit distinct temporal patterns, we summarize the main observations as follows.

- Occupation:** As is shown in Figure 5(a)~5(c), the most representative users with the occupations of manager, administration support, and transport and production share a similar pattern of mostly staying in one location during working day's daytime and in another location during the nighttime, which is consistent with the common nine-to-five working schedule. However, the users with manager occupation seem to have a more flexible working schedule. They sometimes might stay at home during the typical working hours, and occasionally work overtime in the non-working days. On the other hand, the users with administration support and transport and production occupations tend to have a fixed working schedule during the typical working hours. These findings echo previous study in this area that high

pay management jobs have more access to flexible working schedule compared with other relatively low pay occupations [40]. Besides, the users with transport and production occupation sometimes stay at the working sites even during midnight, which is probably because of the shift work that is common in this occupation [41].

- **Age group:** The temporal patterns of three age groups, i.e. 0~20, 25~30 and 50~99 are presented in Figure 5(d)~5(f). Typically, these age groups mainly correspond to the life stages of going to school, starting a career and retirement. We observe that the users in 0~20 age group mainly stay at one location, and occasionally go out during the afternoon in working days and daytime in non-working days. This probably correspond to the schedules of boarding schools and universities, where users only go out when they finish the classes. Besides, the 25~30 age group exhibits a stable nine-to-five schedule in working days, and occasionally works overtime or visit other locations in non-working days. On the other hand, the 50~99 age group tend to mainly stays at one location, and occasional goes out in the morning and evening of working days and the daytime of non-working days.
- **Gender:** Finally, we present the temporal patterns of male and female users in Figure 5(g) and 5(h), respectively. We observe that both male and female users exhibit a similar nine-to-five schedule in working days. On the other hand, male users seem more likely to work overtime in non-working days, while female users tend to visit multiple different locations. These observations probably could be explained by the findings in previous studies that females engage more in multi-purpose and multi-stop trips in order to do household errands and take up gender differentiated roles [18, 42].

In summary, through the temporal patterns visualization and analysis, we find the proposed SUME model can effectively capture the distinct and semantic-aware temporal mobility pattern for the user demographic inference. Although users with different occupations, genders and age groups exhibit similar sleeping cycle, but they differ significantly in working schedule and non-working day activities. These observations echo previous studies on the mobility differences across different user demographics, and also help us to understand how SUME works.

## 5.2 Spatial Domain Patterns

Here, we aim to pick out the most representative locations in the classification of each user demographic, and visualize them in spatial domain. Since the proposed SUME model embeds users and locations into same representation space, we can efficiently look up the most representative locations for certain user demographic by computing the similarity between embedding vectors. Based on the Algorithm 1, we find that the dot product between user's embedding vectors and location's contextual embedding vector measures the importance of this location in inferring that user's demographic, where the higher dot product value indicates higher importance. Guided by this mechanism, we first compute the average embedding vectors among the users that are classified into certain demographics, and then select the top 30 locations for each average embedding vector as the most representative locations. We visualize the spatial distribution of the most representative locations for different user demographics in Figure 6. We adopt the heatmap format for visualization, where red color denotes high density of representative locations and blue color denotes low density. Besides, we also present their differences in mobility characteristics in Table 6.

- **Occupation:** Figure 6(a)~6(c) demonstrate the spatial distribution of the most representative locations of the user population with manager, administration support, and transport and production occupations. We observe distinct differences between them. Specifically, the key features of users with manager occupation is mainly distributing in the downtown area, with three peaks around the famous CBD areas – Zhongguancun (IT center), trade center and financial street. On the contrary, users with administration support and transport and production occupation distribute more evenly in the urban space. On the other hand, Table 6 shows the users with manager occupation has the least radius of gyration and the most mobility entropy

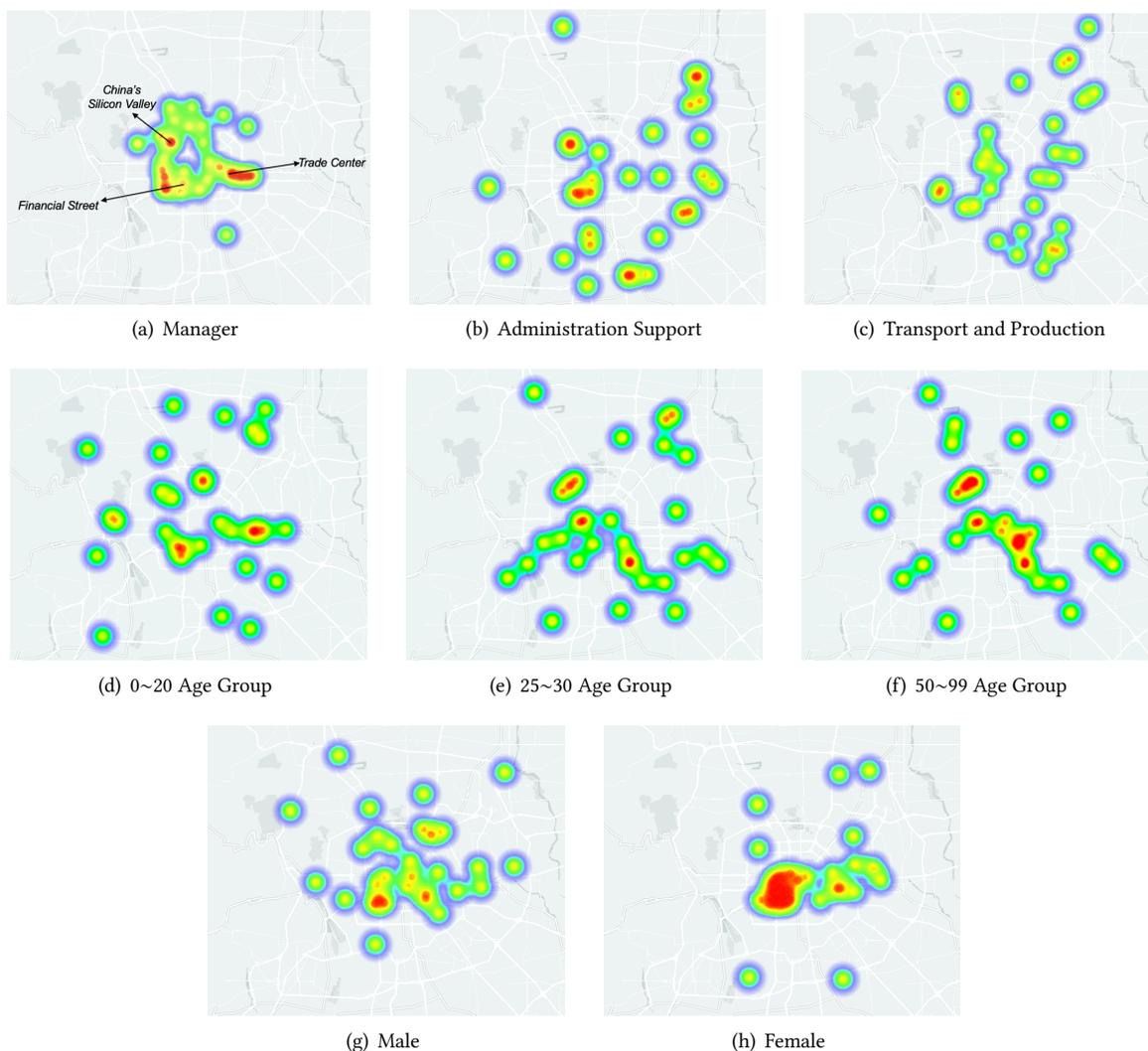


Fig. 6. Spatial distribution of the most representative locations for the inference of different user demographics.

compared with administration support and transport and production. It indicates users with manager occupation has smaller activity areas but allocate their time more evenly among locations, which is consistent with previous observations on spatial and temporal patterns. In addition, users with transport and production occupation have the largest radius of gyration, which is likely due to their occupation requirement for transportation.

- **Age group:** Figure 6(d)~6(f) show the spatial distribution of the most representative locations of different age groups. We observe that they correspond to different locations but all distribute rather evenly in urban space. On the other hand, Table 6 shows the users in 25~30 age group have the largest radius of gyration

Table 6. Differences in mobility characteristics across different genders, occupations and age groups. (-) and (+) represents the groups with least and most value in comparison respectively, which all pass the Student’s t-test (with Bonferroni correction) with  $p < 0.05$  and  $ES > 0.8$ .

Mobility feature	occupation			age			gender	
	manager	admin. support	transport and production	0~20	25~30	50~99	male	female
Radius of gyration(m)	27130.29 (-)	27947.30	29335.14 (+)	23121.90 (-)	30002.30 (+)	27991.72	30715.18 (+)	27681.20 (-)
Mobility entropy	1.69 (+)	1.60 (-)	1.64	1.36 (-)	1.74 (+)	1.53	1.73 (+)	1.63 (-)

Table 7. The results of the deviation of POIs visiting behaviours among different occupations. \*represents  $p\text{-value} \leq 0.05$  and \*\*represents  $p\text{-value} \leq 0.01$  (Student’s t-test with Bonferroni correction). Blocks with red colors are the distinct features.

occupation	business	life service	residence	sports and health	education	recreation	transportation	travel
manager	0.0523*	-0.0744	0.0227	0.0161	0.0739**	0.0476**	-0.0170	-0.0029
administration support	0.0266	0.0607**	0.0021	0.0075	-0.0353	-0.0452	-0.0300	-0.0042
transport and production	-0.0789	0.0136	-0.0248	-0.0235	-0.0386	-0.0024	0.0470**	0.0071*

and mobility entropy, while the users in 25~30 age group have the least. Similar observations have been made in wide range of countries, which is mainly attributed to lack of transportation means (e.g. private cars) in the young and old population[19, 43, 44]

- **Gender:** From Figure 6(g) and 6(h), we observe distinct features on the most representative locations of different genders. Specifically, the representative locations of female users mainly concentrate in the downtown area, while male users distribute more evenly in the urban space. Previous study find evidences that female users tend to visit less unique locations and mainly spend their time on the most frequently visited locations. On the other hand, Table 6 also shows female users generally are less active in mobility. Researchers find evidences that holding less driver licenses, lower socioeconomic status and the fear of personal safety all contribute to the gender gap in urban mobility worldwide [18, 45].

Through visualizing the most representative locations our model captures, we showcase the signature spatial patterns it utilizes for user demographic inferences. Specifically, we find users with different occupations, ages and genders indeed have distinct spatial distribution and mobility characteristics in urban space. These observations provide insights into how SUME learns effective representations from the spatial domain features of urban mobility.

### 5.3 Understanding the Patterns in POI Visitation

In order to further examine the distinct patterns in POI visitation, we analyze the POI distribution in the most representative locations selected by previous spatial analysis. To remove the inherent difference in the number of different types POIs, we preprocess the POI distribution vectors with z-score normalization. Therefore, a positive value indicates users visit this type of POIs more frequently than the general population, while a negative value means otherwise. Then, we perform statistical analysis on the average POI distribution across the most representative locations of different user demographics, which is shown in Table 7~9. The distinct features of each user demographic is highlighted with red colors. We make the following observations.

- **Occupation:** Table 7 shows POI visitation frequency differences among users with manager, administration support, and transport and production occupations. We observe that users with manager occupation visit business, education and recreation types of POIs significantly more frequently than other two occupations. On the other hand, users with transport and production occupations visit transportation and travel types

Table 8. The results of the deviation of POIs visiting behaviours among different age groups. \*represents  $p\text{-value} \leq 0.05$  and \*\*represents  $p\text{-value} \leq 0.01$  (Student's t-test with Bonferroni correction). Blocks with red colors are the distinct features.

age	business	life service	residence	sports and health	education	recreation	transportation	travel
0~20	-0.0780	-0.1057	-0.0016	-0.0115	<b>0.0403 *</b>	-0.0576	-0.0315	-0.0058
25~30	<b>0.0470 *</b>	0.0534	-0.0202	0.0040	-0.0681	0.0353	0.0008	0.0058
(50,99]	0.0311	0.0523	<b>0.0218 *</b>	0.0075	0.0278	0.0223	<b>0.0307 **</b>	0.0000

Table 9. The results of the deviation of POIs visiting behaviours among different genders. \*represents  $p\text{-value} \leq 0.05$  and \*\*represents  $p\text{-value} \leq 0.01$  (Student's t-test with Bonferroni correction). Blocks with red colors are the distinct features.

gender	business	life service	residence	sports and health	education	recreation	transportation	travel
male	-0.0185	0.0434	0.0325	0.0286	0.0133	0.0164	0.0054	-0.0039
female	<b>0.0185 **</b>	-0.0434	-0.0325	-0.0286	-0.0133	-0.0164	-0.0054	0.0039

POIs more frequently. These observations are probably due to their differences in work locations, and users with manager occupation generally has higher salary to spend on education and recreation.

- **Age group:** Table 8 demonstrates POI visitation differences among different age groups. We observe users of 0~20 age group visit education type POIs significantly more frequently, and 25~30 age group visits business type POIs more frequently. These observations fit with our common understandings on the life stages of going to schools and starting a career. On the other hand, the distinct feature of users older than 50 is visiting residence and transportation type POIs more frequently. One plausible explanation is that these users spend more time in home because of retirement, and use public transportation more frequently due to ticket free policy for old citizens.
- **Gender:** Finally, we demonstrate the gender differences on POI visitation in Table 9. We can observe that the only significant difference is female users visit business type POIs more frequently, which is a bit counter-intuitive considering the widely observed gender gap in workplace. One plausible reason is China has a rather high women employment rate, which reaches up to 89% [46], and the workplace gender inequality mainly exists in the salary. In addition, female users tend to spend time on less locations, and hence lead to more frequent visitation.

To conclude, we find distinct features for each type of user demographic in terms of POI visitation frequency, which provides us with deeper insights into the underlying correlations between POI visiting behaviours and user demographics.

## 6 RELATED WORKS AND LIMITATIONS

With the proliferation of mobile devices, the increasingly available mobility data is gaining popularity in numerous applications, including urban planning [10], business sites selection [47], recommender system [48], and user profiling [8]. In this paper, we focus on leveraging the mobility data to infer user's occupations, which is a crucial task to dissect urban mobility and has wide-range application scenarios at the same time. Now, we review the most relevant related works, and summarize them into the following three aspects.

### 6.1 User Profiling with Mobility Data

The prevalent personalized mobile applications give rise to the need of high-quality user profiling [1, 2, 14]. It has been demonstrated as an important link to improving the user experience in recommender system [12], mobile application scheduling [1], location based services [11] and so on. Various data sources, such as web browsing records [14], user-generated content [49] and social network feature [50], have been exploited to construct user

profiles. Although mobility data has become a ubiquitously collected data source, its applications in inferring user profile are relatively more limited. Previous works either focus on leveraging the mobility data to infer social connections since they are closely correlated [51, 52], or combined the mobility data with other mobile sensory data, such as app usage and light sensor, to infer user attributes like age and gender [3, 4]. One key obstacle for inferring the complex user attributes solely based on mobility data is the difficulty of understanding the motivations behind mobility transitions, which prevents the models from capturing the underlying correlation between mobility behavior and user attributes. On the other hand, one recent study demonstrated that human experts can accurately identify the important functional locations, e.g., workplace, home and transport, by solely observing user's trajectories [17]. It showcased the feasibility of inferring the user occupations based on mobility data. In addition, another important related work proposed a temporal pattern based trajectory clustering algorithm, which effectively captured the similarity in time allocation patterns [8]. It showed promising results in inferring user occupations by predicting them as the most frequent ones within user clusters.

Different from most previous works, we aim to infer the user demographics based on the ubiquitously collected mobility data. Instead of straightforward clustering algorithm, we are dedicated to developing a sophisticated representation learning algorithm, which is able to extract the complex correlation between user demographics and mobility behavior as well as capture the semantics of urban mobility at the same time.

## 6.2 Modelling the Context of Urban Regions

Urban mobility tends to be closely correlated with urban structure, because it is often driven by concrete motivations and consists of regular transitions between different contextual locations [53, 54], such as home, workplace and shopping malls. Therefore, modelling the underlying context of urban regions plays an important role in understanding urban mobility patterns [9], which also holds the key to achieve accurate user profiling from mobility data. Numerous attempts have been made to dissect the urban area into smaller and context-aware functional regions with various data sources [10, 55, 56], such as entertainment area and office area. On the other hand, researchers also proposed novel algorithms to detect the dynamic activities in urban area with large scale mobility data [7, 21]. Instead of classifying urban regions into semantic explicit categories, recent studies also looked at learning implicit feature vectors for urban regions with up-to-date representation learning techniques [57, 58]. The derived region representations have been shown to be effective in predicting the crime rate and house price of different urban regions.

Although we are dedicated to harnessing the power of urban structure in user profiling, we are not set to derive an explicit representation of the urban region context. On the other hand, we formulate the urban mobility as a heterogeneous network of user and location, and learn representation of urban regions implicitly with the goal to optimize user profile. Experiments show these designs allow us to effectively make up the shortcoming of semantically unclear mobility data, and significantly improve the performance of user profiling.

## 6.3 Network Embedding Algorithms

Embedding algorithms, also known as representation learning, are a major branch of deep learning techniques, which aims to derive low-dimensional and dense vector that captures the feature in certain aspect [59–61]. An important seminal work is the Word2Vec model that is originally designed to model the word semantics in natural language processing [59]. Inspired by the connection between word co-occurrence and network structure, the embedding techniques are later adapted to learn the structural role of nodes in network, which are referred to as network embedding algorithms [22, 23, 33]. These algorithms have been proved effective in the applications of modelling user-generated content [62], inferring social relationships [63], profiling urban regions [64, 65] and detecting urban activities [7]. Previous works have designed network embedding algorithms to preserve the

node's similarity in terms of link strength and common neighbours, which optimizes the node embeddings to represent homogeneity and structural role respectively.

In this paper, we model urban mobility as a heterogeneous network, with edges between user and location to capture user's location visitation patterns and the edges among locations and users themselves to capture the semantic similarity. Such settings allow us to learn semantic-enhanced urban mobility representation to better capture its complex correlation with user occupations. On the other hand, it also poses new challenges to properly measure the semantic similarity and design semantic-aware urban mobility embedding algorithms.

#### 6.4 Limitations

We note several limitations of our works. First, the results presented in our study are mainly derived from the population of Chinese users. Although we collect two datasets to cover different platforms, it might still suffer from cultural bias. We are actively searching for datasets with different cultural background. However, it is non-trivial to attain additional large-scale mobility data, especially those with user demographic information. We encourage researchers to reproduce our results on their datasets. Second, the evaluation settings in our paper are correlated with the experiment platforms. For example, the taxonomy of occupation, income level and age group are defined by the platforms of data collection. However, the consistent performance boost in the experiments of age, gender, occupation, income and education demonstrate our model's generalizability to other application scenario. Finally, the proposed SUME model is not customized for a specific task. Instead, we aim to extract embedding vectors that cover a wide spectrum of user demographic. Additional performance gain might be achieved by fine-tuning the embedding vector for specific tasks. We leave this as an important future work.

## 7 CONCLUSION

In this paper, we investigate the problem of leveraging large-scale urban mobility data to infer user demographics. To properly capture the important semantic feature in urban mobility, we model it as a heterogeneous information network, with the heterogeneous edges denoting the user physical mobility patterns as well as the semantic similarity. Furthermore, we propose a semantic-enhanced urban mobility embedding (SUME) algorithm to learn high quality representation for each user, which is effective in accurately inferring user's demographics. Experiments show that our model achieves significant performance gain over the state-of-the-art baselines in terms of improving the prediction accuracy with a margin of 8.6%~14.3% for occupation, gender, age, education and income inference. In addition, the learned user representations also reveal important correlations between user demographics and their mobility behavior patterns in spatial, temporal and urban structure domain, which might provide important insights on interpreting urban mobility patterns.

## ACKNOWLEDGMENTS

This work was supported in part by The National Key Research and Development Program of China under grant 2018YFB1800804, the National Nature Science Foundation of China under U1936217, 61971267, 61972223, 61941117, 61861136003, Beijing Natural Science Foundation under L182038, Beijing National Research Center for Information Science and Technology under 20031887521, and research fund of Tsinghua University - Tencent Joint Laboratory for Internet Innovation Technology.

## REFERENCES

- [1] Joohyun Lee, Kyunghan Lee, Euijin Jeong, Jaemin Jo, and Ness B Shroff. Context-aware application scheduling in mobile systems: What will users do and not do next? In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1235–1246. ACM, 2016.

- [2] Pengyang Wang, Yanjie Fu, Hui Xiong, and Xiaolin Li. Adversarial substructured representation learning for mobile user profiling. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 130–138. ACM, 2019.
- [3] Bjarke Mønsted, Anders Mollgaard, and Joachim Mathiesen. Phone-based metric as a predictor for basic personality traits. *Journal of Research in Personality*, 74:16–22, 2018.
- [4] Zhiwen Yu, En Xu, He Du, Bin Guo, and Lina Yao. Inferring user profile attributes from multi-dimensional mobile phone sensory data. *IEEE Internet of Things Journal*, 2019.
- [5] Feixiong Luo, Guofeng Cao, Kevin Mulligan, and Xiang Li. Explore spatiotemporal and demographic characteristics of human mobility via twitter: A case study of chicago. *Applied Geography*, 70:11–25, 2016.
- [6] Ke Zhang, Yu-Ru Lin, and Konstantinos Pelechris. Eigentransitions with hypothesis testing: The anatomy of urban mobility. In *Tenth International AAAI Conference on Web and Social Media*, 2016.
- [7] Chao Zhang, Keyang Zhang, Quan Yuan, Haoruo Peng, Yu Zheng, Tim Hanratty, Shaowen Wang, and Jiawei Han. Regions, periods, activities: Uncovering urban dynamics via cross-modal representation learning. In *Proceedings of the 26th International Conference on World Wide Web*, pages 361–370. International World Wide Web Conferences Steering Committee, 2017.
- [8] Fengli Xu, Tong Xia, Hancheng Cao, Yong Li, Funing Sun, and Fanchao Meng. Detecting popular temporal modes in population-scale unlabelled trajectory data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1):46, 2018.
- [9] Shan Jiang, Yingxiang Yang, Siddharth Gupta, Daniele Veneziano, Shounak Athavale, and Marta C González. The timegeo modeling framework for urban mobility without travel surveys. *Proceedings of the National Academy of Sciences*, 113(37):E5370–E5378, 2016.
- [10] Fengli Xu, Pengyu Zhang, and Yong Li. Context-aware real-time population estimation for metropolis. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1064–1075. ACM, 2016.
- [11] Giuseppe Sansonetti. Point of interest recommendation based on social and linked open data. *Personal and Ubiquitous Computing*, 23(2):199–214, 2019.
- [12] FO Isinkaye, YO Folajimi, and BA Ojokoh. Recommendation systems: Principles, methods and evaluation. *Egyptian Informatics Journal*, 16(3):261–273, 2015.
- [13] Sha Zhao, Zhiling Luo, Ziwen Jiang, Haiyan Wang, Feng Xu, Shijian Li, Jianwei Yin, and Gang Pan. Appusage2vec: Modeling smartphone app usage for prediction. In *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, pages 1322–1333. IEEE, 2019.
- [14] Ram Keralapura, Antonio Nucci, Zhi-Li Zhang, and Lixin Gao. Profiling users in a 3g network using hourglass co-clustering. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking*, pages 341–352. ACM, 2010.
- [15] Leqi Liu, Daniel Preotiuc-Pietro, Zahra Riahi Samani, Mohsen E Moghaddam, and Lyle Ungar. Analyzing personality through social media profile picture choice. In *Tenth international AAAI conference on web and social media*, 2016.
- [16] Sha Zhao, Shijian Li, Julian Ramos, Zhiling Luo, Ziwen Jiang, Anind K Dey, and Gang Pan. User profiling from their use of smartphone applications: A survey. *Pervasive and Mobile Computing*, page 101052, 2019.
- [17] Ilaria Liccardi, Alfie Abdul-Rahman, and Min Chen. I know where you live: Inferring details of people’s lives by visualizing publicly shared location data. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 1–12. ACM, 2016.
- [18] Laetitia Gauvin, Michele Tizzoni, Simone Piaggese, Andrew Young, Natalia Adler, Stefaan Verhulst, Leo Ferres, and Ciro Cattuto. Gender gaps in urban mobility. *arXiv preprint arXiv:1906.09092*, 2019.
- [19] Anu Siren and Liisa Hakamies-Blomqvist. Private car as the grand equaliser? demographic factors and mobility in finnish men and women aged 65+. *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(2):107–118, 2004.
- [20] Chao Zhang, Keyang Zhang, Quan Yuan, Luming Zhang, Tim Hanratty, and Jiawei Han. Gmove: Group-level mobility modeling using geo-tagged social media. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1305–1314, 2016.
- [21] Shan Jiang, Joseph Ferreira Jr, and Marta C Gonzalez. Discovering urban spatial-temporal structure from human activity patterns. In *Proceedings of the ACM SIGKDD international workshop on urban computing*, pages 95–102. ACM, 2012.
- [22] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 701–710. ACM, 2014.
- [23] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. Line: Large-scale information network embedding. In *Proceedings of the 24th international conference on world wide web*, pages 1067–1077. International World Wide Web Conferences Steering Committee, 2015.
- [24] Mohammad Mehdi Keikha, Maseud Rahgozar, and Masoud Asadpour. Community aware random walk for network embedding. *Knowledge-Based Systems*, 148:47–54, 2018.

- [25] Dmitry A Konovalov, Bruce Litow, and Nigel Bajema. Partition-distance via the assignment problem. *Bioinformatics*, 21(10):2463–2468, 2005.
- [26] K. Chen G. S. Corrado T. Mikolov, I. Sutskever and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, pages 3111–3119, 2013.
- [27] Huandong Wang, Chen Gao, Yong Li, Gang Wang, Depeng Jin, and Jingbo Sun. De-anonymization of mobility trajectories: Dissecting the gaps between theory and practice. In *The 25th Annual Network & Distributed System Security Symposium (NDSS'18)*, 2018.
- [28] Yi Ren, Weimin Mai, Yong Li, and Xiang Chen. Predicting socio-economic levels of individuals via app usage records. In *International Conference on Machine Learning and Intelligent Communications*, pages 199–210. Springer, 2019.
- [29] Theresa Cosca and Alissa Emmel. Revising the standard occupational classification system for 2010. *Monthly Labor Review*, 133(8):32–41, 2010.
- [30] Sunhee Sang, Morton OKelly, and Mei-Po Kwan. Examining commuting patterns: results from a journey-to-work model disaggregated by gender and occupation. *Urban studies*, 48(5):891–909, 2011.
- [31] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. Understanding individual human mobility patterns. *nature*, 453(7196):779, 2008.
- [32] David Guthrie, Ben Allison, Wei Liu, Louise Guthrie, and Yorick Wilks. A closer look at skip-gram modelling. In *LREC*, pages 1222–1225, 2006.
- [33] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864. ACM, 2016.
- [34] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
- [35] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. *arXiv preprint arXiv:1710.10903*, 2017.
- [36] Johan AK Suykens and Joos Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, 1999.
- [37] Andrew I Schein, Alexandrin Popescul, Lyle H Ungar, and David M Pennock. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 253–260. ACM, 2002.
- [38] Paul D Ellis. The essential guide to effect sizes: An introduction to statistical power, meta-analysis and the interpretation of research results. 2010.
- [39] Adil Cubukgil and Eric J Miller. Occupational status and the journey-to-work. *Transportation*, 11(3):251–276, 1982.
- [40] Rebecca Glauber. Limited access: Gender, occupational composition, and flexible work scheduling. *The Sociological Quarterly*, 52(3):472–494, 2011.
- [41] Terence M McMenamin. A time to work: recent trends in shift work and flexible schedules. *Monthly Lab. Rev.*, 130:3, 2007.
- [42] D Brown, G McGranahan, and J Dodman. Urban informality and building a more inclusive. *London: Resilient and Green Economy*, 2014.
- [43] Lisa I Iezzoni, Ellen P McCarthy, Roger B Davis, and Hilary Siebens. Mobility difficulties are not only a problem of old age. *Journal of general internal medicine*, 16(4):235–243, 2001.
- [44] Antonio Paez, Darren Scott, Dimitris Potoglou, Pavlos Kanaroglou, and K Bruce Newbold. Elderly mobility: demographic and spatial analysis of trip making in the hamilton cma, canada. *Urban Studies*, 44(1):123–146, 2007.
- [45] Ismael Peña-López et al. Big data, big impact? towards gender-sensitive data systems. 2019.
- [46] Yong Jing Teow, Saloni Goel, Tara Shrestha Carney, Alex Cooper, Jon Terry, and Katy Bennett. Women in work index 2019, 2019.
- [47] Bin Guo, Jing Li, Vincent W Zheng, Zhu Wang, and Zhiwen Yu. Citytransfer: Transferring inter-and intra-city knowledge for chain store site recommendation based on multi-source urban data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(4):135, 2018.
- [48] Chen Gao, Chao Huang, Yue Yu, Huandong Wang, Yong Li, and Depeng Jin. Privacy-preserving cross-domain location recommendation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(1):1–21, 2019.
- [49] Shangsong Liang, Xiangliang Zhang, Zhaochun Ren, and Evangelos Kanoulas. Dynamic embeddings for user profiling in twitter. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 1764–1773. ACM, 2018.
- [50] R Logesh, V Subramaniaswamy, and V Vijayakumar. A personalised travel recommender system utilising social network profile and accurate gps data. *Electronic Government, an International Journal*, 14(1):90–113, 2018.
- [51] Dingqi Yang, Bingqing Qu, Jie Yang, and Philippe Cudre-Mauroux. Revisiting user mobility and social relationships in lbsns: a hypergraph embedding approach. In *The World Wide Web Conference*, pages 2147–2157. ACM, 2019.

- [52] Ran Cheng, Jun Pang, and Yang Zhang. Inferring friendship from check-in data of location-based social networks. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, pages 1284–1291. ACM, 2015.
- [53] Zhixian Yan, Dipanjan Chakraborty, Christine Parent, Stefano Spaccapietra, and Karl Aberer. Semantic trajectories: Mobility data computation and annotation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4(3):49, 2013.
- [54] Fei Wu, Zhenhui Li, Wang-Chien Lee, Hongjian Wang, and Zhuojie Huang. Semantic annotation of mobility data using social media. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1253–1263. International World Wide Web Conferences Steering Committee, 2015.
- [55] Jameson L Toole, Michael Ulm, Marta C González, and Dietmar Bauer. Inferring land use from mobile phone activity. In *Proceedings of the ACM SIGKDD international workshop on urban computing*, pages 1–8. ACM, 2012.
- [56] Jing Yuan, Yu Zheng, and Xing Xie. Discovering regions of different functions in a city using human mobility and pois. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 186–194. ACM, 2012.
- [57] Zijun Yao, Yanjie Fu, Bin Liu, Wangsu Hu, and Hui Xiong. Representing urban functions through zone embedding with human mobility patterns. In *IJCAI*, pages 3919–3925, 2018.
- [58] Hongjian Wang and Zhenhui Li. Region representation learning via mobility flow. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, pages 237–246. ACM, 2017.
- [59] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.
- [60] Jeffrey Pennington, Richard Socher, and Christopher Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014.
- [61] Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE transactions on pattern analysis and machine intelligence*, 35(8):1798–1828, 2013.
- [62] Jian Tang, Meng Qu, and Qiaozhu Mei. 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*, pages 1165–1174. ACM, 2015.
- [63] Yanwei Yu, Hongjian Wang, and Zhenhui Li. Inferring mobility relationship via graph embedding. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3):147, 2018.
- [64] Shanshan Feng, Gao Cong, Bo An, and Yeow Meng Chee. Poi2vec: Geographical latent representation for predicting future visitors. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [65] Bo Yan, Krzysztof Janowicz, Gengchen Mai, and Song Gao. From itdl to place2vec: Reasoning about place type similarity and relatedness by learning embeddings from augmented spatial contexts. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, page 35. ACM, 2017.