Habit2vec: Trajectory Semantic Embedding for Living Pattern Recognition in Population

Hancheng Cao, Fengli Xu, Jagan Sankaranarayanan, Yong Li, *Senior Member, IEEE*, and Hanan Samet, *Fellow, IEEE*

Abstract—Recognizing representative living patterns in population is extremely valuable for urban planning and decision making. Thanks to the growing popularity of location-based applications and check-ins on social networking sites, Point of Interest (POI) of a location is quite often available in the trajectory data, which expresses user living semantics. However, adopting trajectory semantics for living pattern recognition is an open and challenging research problem due to three major technical challenges: effective feature representation, suitable granularity selection for habit unit, and reliable habit distance measurement. In this paper, we propose a representation learning based system named habit2vec to represent user trajectory semantics in vector space, which preserves the original user living habit information. We evaluated our proposed system on a large-scale real-world dataset provided by a popular social network operator including 123,803 users for 1.5 months in Beijing. The results justify the representation ability of our system in preserving user habit pattern, and demonstrate the effectiveness of clustering users with similar living patterns.

Index Terms—Representation learning, trajectory mining, pattern recognition, urban computing

1 Introduction

With the increasing popularity of personal mobile devices and location-based applications, large-scale semantic-rich trajectories of individuals are being recorded and accumulated at a faster rate than ever [1], where Point of Interest (POI) of a location is often available and associated with the trajectories [2]. POI information, as the semantics of location, is a good indicator of the person's behavior at the location [3], [4]. Mining underlying patterns in trajectory semantics through POIs therefore make it possible to recognize typical living patterns in the city. Understanding living patterns in population is of great importance, as it has the potential to reveal people's social and economic status [5], as well as social capital [6], which provides key insights for city planners and decision makers. Despite its great significance, there have been few studies dedicated to living pattern recognition in population via semantic-rich trajectory data.

Recently, emerging research on trajectories focuses on mining frequent patterns. For instance, Lee et al. [7] and Yao et al.[8] propose trajectory clustering methods to cluster users who share similar geographical routes. However, these works on trajectory pattern mining are based on the viewpoint of physical location transition patterns and therefore, constrained to only discovering common mobility pattern of people located in nearby geographical regions. Meanwhile, there have been advancement focusing on mining trajectory semantics similarity. Jiang et al. [9] make use of a PCA-based method to cluster daily patterns of human activities through travel survey data; Furtado et al. [10] propose a multidimensional similarity measures to compare semantic trajectories. However, these works typically measure trajectory semantics solely on static POI type labels. Two trajectories with similar semantics but distinct POI type labels (e.g., supermarkets and shops) will therefore be measured completely different. As a result, these works often involve manually grouping POI type labels, which heavily rely on prior knowledge and result in coarse granularity.

In this paper, we seek to recognize typical living patterns distributed in different geographical locations in population through the semantics embedded in the trajectories. We define the similarity of living patterns as engaging in similar behavior at similar times instead of staying in geographically neighboring location. For instance, the people in the city who follow the weekday routine: sleep at night in residence district, get up at 9 am, go to work in commercial center from 10 am to 6 pm, and get back to arrive at home in the residence district at 8:30 pm, belong to the same living pattern group, though they may be physically far away.

Nevertheless, recognizing typical living patterns in population through POI semantic is challenging. First, there is no ready method to build user habit representation through the varied-length and often biased POI records in trajectory data.

- H. Cao is with the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the Department of Computer Science, Stanford University, Stanford, CA 94305. E-mail: hanchcao@stanford.edu.
- F. Xu and Y. Li are with the Beijing National Research Center for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. E-mail: {xfl15, liyong07}@tsinghua.edu.cn.
- J. Sankaranarayanan is with the Department of Computer Science, University of Maryland, College Park, MD 20742. E-mail: sjagan@gmail.com.
- H. Samet is with the Department of Computer Science, University of Maryland, College Park, MD 20742. E-mail: hjs@umiacs.umd.edu.

Manuscript received 27 Jan. 2018; revised 3 Nov. 2018; accepted 29 Jan. 2019. Date of publication 1 Mar. 2019; date of current version 2 Apr. 2020.
(Corresponding author: Yong Li.)
Digital Object Identifier no. 10.1109/TMC.2019.2902403

Second, it is difficult to select a suitable POI granularity as there are usually multiple levels of coarse-grain granularity. Third, it is also challenging to define accurate metrics to measure the similarity between user habits to take both semantics and time scheduling into account. Classical approach fails to capture the semantic similarity between different POI type as well as temporal variations of POI type semantics.

To address the above three challenges, we propose a representation learning-based system to convert trajectories into living patterns. First, POI type transitions are extracted from raw trajectory data. To solve the feature representation problem, they are put in a preprocessing layer, which outputs a fixed length representation for each individual's unbiased living habit. Next, a representation learning inspired feature extraction layer produces vector representation for each person's living habit preserving both semantic and temporal information in the same space. Therefore, recognizing typical living pattern in population is reduced to a classical clustering problem through our system. Our contributions can be summarized as follows:

- To the best of our knowledge, we are the first to recognize human living patterns in population through trajectory data along with the outsider semantic information that breaks through the geographical constraints.
- We propose a representation learning system named habit2vec to represent the habit of a person as a vector, which upgrades the word2vec model according to special characteristics of trajectories. The system fits the objective of living pattern recognition and solves the feature representation problem.
- Through extensive experiments, we evaluate our proposed habit2vec system on a large-scale real-world dataset that records the trajectory of 123,803 users for 1.5 months. The results justify the ability of our system in preserving user habit information, which discovers 13 typical weekday living habits and 12 typical weekend living habits, coinciding physical meanings. We show that our proposed system achieves significant performance gain over baseline methods.

The rest of the paper is structured as follows. Section 2 reviews related work. Section 3 identifies the problem and discusses the key challenges. Section 4 proposes the framework of our habit2vec system. We evaluate our proposed system in Section 5 and provide concluding remarks in Section 6.

2 Related Work

We summarize the closely related works from three aspects: trajectory mining, activity pattern modelling and representation learning.

Trajectory Pattern Mining. Extensive studies have been dedicated to detecting the prevailing trajectory patterns in large scale spatiotemporal data. However, previous works mostly focus on identifying the popular location sequences shared by different trajectories and grouping trajectories based on their physical closeness [7], [11], [12]. Giannotti et al. [13] designed *T-patterns* framework to address the problem of detecting frequent sub-trajectories in spatiotemporal data. Mamoulis et al. [11] focused on mining frequent periodic mobility patterns, Zheng et al. [14] investigated the problem of detecting frequent traveling paths between fixed locations, and Salidek et al. [15] leveraged principal component analysis to extract mobility pattern for long-term location prediction. As for trajectory clustering, Lee et al. [7] proposed a *partition-and-group* framework to detect popular common sub-trajectories and group similar trajectories based on the shared sub-trajectories. [16] exploited principal component analysis technique to extract latent mobility patterns from raw trajectories and cluster trajectories based on the latent features. However, this line of research is limited in measuring the similarity between trajectories based on their physical closeness, such as distance, overlapping sub-trajectories and co-occurrences, therefore unable to understand semantic patterns behind human mobility. More recently, there have been works on semantic-rich trajectory mining. Zhang et al. [17] proposed a hidden Markov model based approach to discover user groups that share similar mobility patterns taking into consideration mobility semantics. Ying et al. [18], [19] adopted trajectory semantic feature to assist location prediction. Zhang et al. [20] developed Splitter system to mine fine-grained sequential patterns in semantic trajectories. Yuan et al. [21] proposed e a Bayesian non-parametric model to to discover periodic mobility patterns for social media users by modeling the geographical and temporal information. However, none of these works focus on jointly modeling temporal and semantic aspect in human mobility. Different from them, we investigate a novel problem of mining living patterns embedded in trajectories. Instead of building semantic-aware mobility model or mining sequential pattern in semantic trajectory, we design a methodology that captures the semantic features of living patterns in a vector space so as to better understand user social-economic behavior pattern.

Activity Pattern Modelling. Modelling the activity patterns in individuals' daily lives is an increasingly important topic that has been extensively studied in recent years [22]. Some early works study the nature of activity patterns (routine behavior), and compare them with grammar in natural language processing [23], [24]. Other works focus on mining the activities behaviours from survey data. Eagle et al. [25] utilized PCA algorithm to extract the features from semantic annotated trajectories, and then identified clusters of activity patterns. Farrahi et al. [26] adopted distant N-Gram topic model to extract user mobile behavioral patterns. Banovic et al. [27] proposed a decision-theoretic framework to rationalize the casual relationship in human routine behaviour logs. Jiang et al. [9] exploited statistical learning techniques to analyze an activity-based travel survey, where the spatiotemporal points are labeled with activities. There are also works dedicated to understanding user indoor mobility/occupancy pattern for location prediction [28] and smart home heating [29]. While these works aimed to identify the key patterns in individual's daily activities, they heavily relied on human-labeled survey data, which is typically not representative and prevented population scale analysis. In addition, unsupervised methods have also been developed to model individual's activity patterns. Furletti et al. [10] proposed a method to infer the activities behind the GPS records. Furtado et al. [30] developed an unsupervised algorithm to measure similarity between semantic trajectory data, where the spatiotemporal records are associated with POIs, Cao et al. [31] studied user location revisitation patterns in urban space and Xu et al. [32] proposed a clustering method to identify popular temporal modes in population. However, these works fell short of capturing the semantic features of the trajectories, since they did not properly model the correlation between activities or POIs. Different from previous works, in this paper we develop an unsupervised algorithm to model the semantic similarities between unlabeled semantic-rich trajectories (i.e., user POI type transition traces), which shows promising results in living pattern recognition.

Representation Learning. Representation learning is a category of unsupervised learning method that aims to extract effective and low-dimensional features from the complicated and high-dimensional data [33]. Various algorithms have been designed to capture the features in different data sources. In the area of natural language processing, Mikolov et al. [34], [35] proposed word2vec, a neural network based representation learning model, to extract the features of words' semantic meaning from their sequential orders. Pennington et al. [36] designed a representation learning algorithm, Glove, that captured the semantic meaning of words based on their global co-occurrences features. In addition, Perozzi et al. [37] and Tang et al. [38] introduced the representation learning techniques into complex networks analysis, and they proposed different algorithms to derive representation for nodes' structural roles in the network. The representation learning techniques have also been applied in spatial-temporal data mining. Yao et al. [8] designed a recurrent neural network to capture the physical features of trajectories with a continuous vector, which enabled them to detect trajectories that are similar in speed and acceleration patterns. Zhang et al. [3] modelled the semantic meaning of spatial-temporal points based on their co-occurrence with the texts in social media's checkins. Different from previous works, we develop an algorithm to extract effective semantic representations for individuals' living patterns from their trajectories, i.e., the transition patterns between location semantics. We demonstrate that our derived representations facilitate the task of living pattern recognition on population scale.

3 System Overview

3.1 Motivation and Challenges

In this paper, we aim to cluster population into groups of similar living habits. We consider living habits as people's regular behavior at specific times. As is often the case, a person's behavior at a time is strongly related to his current location. Thus, we are motivated to develop a system to derive the representation of people's living habits from their trajectories.

However, different from previous works on trajectory mining, which aim at clustering people of similar geographical location transition pattern, we seek to group people who share similarity in trajectory semantics, i.e., people who go to similar type of places at similar time. To put it another way, we aim to group individuals sharing similar daily routine but not necessarily in nearby places. For this goal, we filter out other information in trajectory data (e.g., GPS information, user profile) and select the semantic information, i.e., the transition series of different types of POIs, as the principal input. The system outputs clusters of distinct living habits reflected in the trajectory. Recognizing living patterns and clustering people based on POI transition series, however, is challenging for three reasons.

Feature Representation. Raw POI transition series in our daily trajectories are sparse and quite often not uniformly sampled in passively recorded trajectories. For instance, users tend to use their phones more frequently during their leisure time than during working hours. Therefore, there tend to be more POI records at noon or in the evening. If directly using the raw data as a feature, the user's living habit will be represented in a biased way. How to select proper features to represent a user's daily living habit is therefore hard to manage.

Granularity Selection. A person's daily habit is reflected by a trace of POI transitions. However, there are multiple levels of POI types. High-level types fail to capture meaningful types of living patterns since they do not properly distinguish semantics, while finer-grained types capture semantic differences much better. How to select a proper granularity to represent user's living pattern is challenging.

Distance Metric. Clustering people of similar living habits requires a distance metric to measure the similarity between users' living patterns while a good metric should consider both semantic and temporal similarity. In terms of semantics, people with high similarity should go to similar types of locations every day. In terms of a temporal factor, people with great similarity should have analogous time scheduling. How to define an effective distance metric to combine semantics and temporal factors so as to cluster people who appear at similar POI types at close times is of great difficulty.

3.2 System Overview

In order to effectively tackle the above three challenges, we propose a representation learning based system to convert trajectory into living pattern clusters. First, POI type transitions is extracted from raw trajectory data. To solve the feature representation problem, the POI type transitions is then put in a preprocessing layer and the layer outputs a fixed length representation for each individual's unbiased living habit. Next, a feature extraction layer based on representation learning produces vector representation for each person's living habit, which preserves both semantic and temporal information in the same space so that similarity between user habits can be easily determined. Therefore, the granularity selection and distance measuring challenges are resolved. Finally, clustering analysis is made on living habit vectors (along with other user-specified features from trajectory data) to output a living habit group. The framework of our system is shown in Fig. 1.

4 Embedding Semantics in Trajectories

4.1 Preliminary

To better represent semantic differences in people's living habits, we first utilize the lowest level POI type representation. Based on it, we define a habit record Hr as follows.

Definition 1. *A habit record Hr records a person's habit at a specific timestamp, in the format of (p, t), where p represents a POI type, and t represents a specific timestamp, meaning an individual appears in POI type p at timestamp t.*

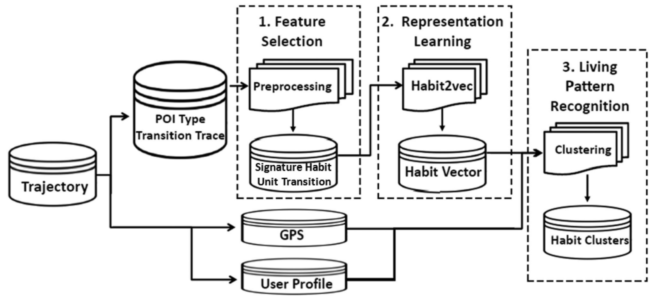


Fig. 1. System architecture.

Based on the above definition for habit record, we further define a person's raw habit trace, denoted by H_{tr}.

Definition 2. *A person's raw habit trace H_{tr} is a relation containing all habit records left by a specific person in the dataset, in the format of $(u_i, Hr_1, Hr_2, \ldots, Hr_n)$, where u_i represents the ID of the individual, Hr_j represents the jth habit record of u_i and n represents the total number of habit records of u_i in the dataset.*

For different users, the total number of habit records n is most likely to be different. As mentioned earlier, trajectory data recorded by mobile devices are often biased in time and of varying length. Furthermore, POI information are usually quite sparse in trajectory data. To effectively represent a person's typical living pattern for later analysis, as well as to filter out redundant information, we therefore carry out a preprocessing step on the raw habit trace H_{tr}, which is motivated by the following two observations.

- Individual mobility follows a strong periodicity pattern and is therefore highly predictable, i.e., people tend to be in similar places at similar time[39].
- Most of us have quite different mobility trajectories on weekdays and weekends.

In the preprocessing, we compress the irregular raw habit trace H_{tr} for a person into two fixed-length POI transition traces: we divide a day into m equal-length time slices (e.g., 30 minutes) for both weekday and weekend and then aggregate the person's raw habit trace H_{tr} on those time slices, which solves the data sparsity problem. To best represent the person's living pattern, we select the most frequent POI type the person visits in each time slices as his typical habit during that time slice, and have the following definition for signature habit trace H_s and habit unit h_j^p.

Definition 3. *Individual's signature habit trace H_s is a feature representing a person's typical POI type he/she visits at a specific time slice, in the format of $(u_i, p_1, p_2, \ldots, p_m)$, where u_i represents the ID of the person, p_j is u_i's most likely POI type to visit during the jth time slice on weekday/weekend and m represents the total number of time slices.*

Definition 4. *Habit unit h_j^p represents a basic unit in POI type - time slice two dimensional space, meaning a visit to POI type p at the jth time slice.*

Definition 5. *A person's weekday/weekend signature habit unit transition H_u is a feature representing a person's typical POI type he/she visits at a specific time slice, in the format of $(u_i, h_1^{p_1}, h_2^{p_2}, \ldots, h_m^{p_m})$, where u_i represents the ID of the person, $h_j^{p_j}$ is u_i's most likely state of habit unit during the jth time slice on weekday/weekend, and m represents the total number of time slices.*

4.2 Representation Learning on Living Habits

Although we have obtained fixed-length signature habit trace H_e to represent the typical living pattern of each user in the dataset on both weekdays and weekends, this feature is not expressive enough for analysis yet. First, it fails to capture the similarity between different types of POIs such as Beijing Style and Shanghai Style restaurant, which is a frequent problem under the finest POI labels. Second, a metric to compare different users' signature habit trace combining semantic and temporal factor, is still hard to define. Therefore, we propose a representation learning method inspired by word2vec to embed semantics and temporal factors of users' signature habit trace in the same space. User habit similarity can therefore be easily determined by classical distance/similarity metrics.

Representation learning, as a growing interest and emphasis on unsupervised learning, aims at transforming complicated, high-dimensional and often redundant real-world data into low-dimensional data while preserving information embedded in the raw data [33].

Word2vec [34], takes advantage of a three-layer neuron network to learn input corpus. It finds a fixed-length low-dimensional representation (often by the hundred assigned by users) for each word. Word representations are learnt in a way such that words sharing common contexts in the corpus are located close to each other in the embedding space, thus word similarity can be easily determined by cosine similarity. Experiments show that word2vec is both effective and efficient in learning word-level semantics.

As in [23], [24], we have discovered a strong similarity between natural language and signature habit uniter transition H_u.

First, natural language and signature habit unit transition can both be viewd as time-dependent series. For each word in the sentence, there can be multiple choices from the dictionary regardless of context. Similarly, there are multiple choices of habit unit for each element in the signature habit unit transition.

Second, both natural language and signature habit unit transition can be approximated by context. In many cases, if given context, we are able to predict nearby words without much trouble. Likewise, a human living pattern has some typical transition modes, which are reflected in POI type transition mode in signature habit unit transition.

Third, large scale of data are available for both natural language and signature habit unit transition to learn their characteristics.

Lastly, the frequency distribution of habit unit is very similar to word frequency distribution in natural language. A typical distribution of habit unit is shown in Fig. 2 (observed in our dataset utilized in experiment), which approximately follows Zipf Law, the governing law in word frequency distribution [40].

Therefore, we draw an analogy between learning representation for signature habit trace and word embedding, as shown in Table 1. Inspired by the idea of word2vec, we are

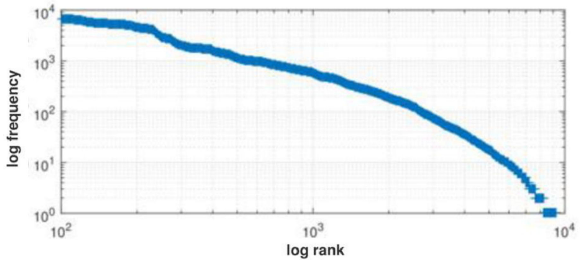


Fig. 2. Statistics of habit unit satisfies Zipf law.

motivated to propose an algorithm to learn an individual's habits from the trajectory, with the name of habit2vec.

4.3 Methodology

The key idea of habit2vec is to embed a user's current visit to a POI type $target$ during a time slice (current habit unit h_j^{pj}) based on its $context$. We first acquire embedding vectors for all habit units. Then, we take mean of all habit units appearing in a user's signature habit unit transition as this user's living habit embedding vector.

The model requires two user-specified hyper-parameters, dim and w. We define hyper parameter dim as the expected length of embedding vectors (same for habit unit embedding vector and user living habit embedding vector). We define another hyper parameter window size w so as to define the context of habit unit appearing in a user's trace. Here we make the assumption that habit units the user appears farther than w time slices won't have direct influence on the user's current habit unit state, and exclude them from the context.

Definition 6. *Suppose person u_i is of habit h_j^p at time slice j, context $C(h_j^p)$ represents all habit units u_i visits in nearby time slices of j. With a user-specified window size w, $C(h_j^p)$ contains habit units person u_i visits from time slice $(j-w)$ to $(j+w)$ (time slice j excluded). The habit unit state h_j^p at the jth time slice is called target.*

Note that we need to pay special attention to the boundary. Instead of treating the habit unit transition as a line, which neglects the dependency between time slices right before and after midnight, we treat each user's habit unit transition as circle. The context of a boundary target is shown as an example in Fig. 3. In this case, window size is assigned 3, and the context of habit unit for this user at 11:00 pm not only includes habit unit at 9:30, 10:00, 10:30 and 11:30 pm, but also takes into account those 'very first' habit units at midnight and 0:30 am.

By sequentially identifying each habit unit in each user's habit unit transitions as a target and sliding the window across the user's habit unit transitions to get the target's context, we get a list of $(target, context)$ training pairs and put them into a three-layer neural network model Continuous Bag-of-Words (CBOW) [35] to learn the habit unit embedding vector. All habit unit embedding vectors are initialized as random dim-dimension vectors. The objective of the neural network is to adjust and find the optimal weights of neuron and habit unit vectors at the output layer such that the possibility of $target$'s appearance is maximized when given $context$.

TABLE 1
Analogy from Habit Trace to Natural Language

Habit unit (a visit to a POI type during a time slice) $\longrightarrow$ word
Habit unit transition mode $\longrightarrow$ grammar rules/fixed
collocation an individual's signature habit unit transition $\longrightarrow$
sentence people's signature habit unit transitions $\longrightarrow$ corpus

The architecture of the CBOW model is shown in Fig. 4, where N is the number of habit units in the dataset. Denote the vector representation of habit unit h_j^p as $w(h_j^p)$. When training $(target_i, context_i)$, the neural network takes all one-hot key representations (the way to represent categorical data where only one label bit is '1' while all other bits are '0') of $context_i$ as input and uses an embedding vector matrix acquired in previous training steps to transform one-hot key representation of $context_i$ to vector representation $\{w(l), l \in C(h_j^p)\}$, as shown in the input layer of Fig. 4. Then a second layer sums up all vector representations of $context_i$ and get the output vector

$$\Phi(C(h_j^p)) = \sum\nolimits_{l \in C(h_j^p)} w(l),$$

as shown in the projection layer of Fig. 4. The third layer transforms vector back into N dimension at the output and then predicts the possibility of $target_i$ given $context_i$ using softmax function(shown in the output layer of Fig. 4). More formally, the posterior probability of $target_i$ given $context_i$ is calculated as follows:

$$p(h_j^p | C(h_j^p)) = e^{w(h_j^p)\cdot\Phi(C(h_j^p))} / \sum\nolimits_{h \in H} e^{w(h)\cdot\Phi(C(h))},$$

where H is the set of all habit units appearing in the dataset. Finally, the training objective of habit2vec is to maximize the average log probability

$$\frac{1}{|H|} \sum\nolimits_{h \in H} \log p(h|C(h)).$$

Using an optimization method such as gradient descent, the weights of the neuron and embedding vector representation are adjusted accordingly. Techniques such as negative sampling [34] can help speed up the training process. The complexity of CBOW training process is log-linear.

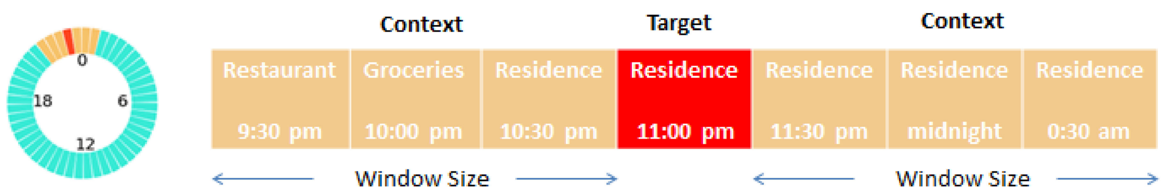


Fig. 3. Learning context through sliding window.

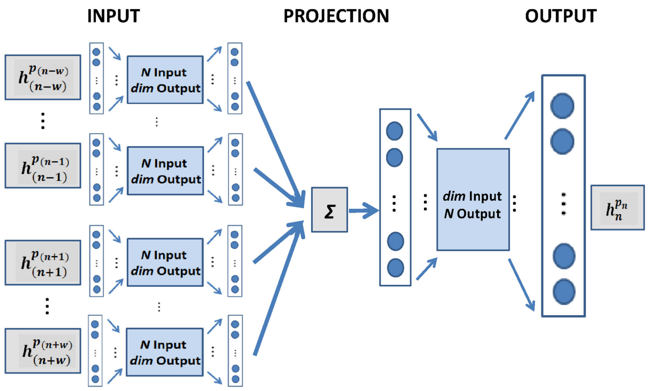


Fig. 4. Architecture of CBOW neural network.

After training with a large enough dataset, the weights of neurons and representation for embedding vectors converge. We therefore obtain embedding vectors for all habit units.

4.4 Clustering Method

After the habit2vec representation learning procedure, we obtain the vector representations for the living habit of each user. Since the living habit has been embedded in a single space, we can use a classical clustering method, such as K-means and density-based method, to cluster similar living habits. People who visit similar POI types at similar time slices will be under the same cluster. Apart from the POI type transition information we use in the habit2vec process, there is usually other trajectory information from the raw dataset, such as GPS information and a user profile. Through user-specified features and a distance/similarity metric, potentially we can get finer living habit clusters with constraints, such as "find clusters of people who have similar POI transition patterns and travel similar distance from home to work".

5 Experiments

Now, based on a large-scale real world spatial-temporal dataset, we implement our proposed habit2vec system to detect human living habits and further cluster them into different groups. We show that habit2vec is capable of capturing semantics in user trajectory, recognizing meaningful living patterns in population, and significantly improve performance over baseline methods.

5.1 Data Description

Our dataset is collected by Tencent, one of the largest social network service providers in China. Its service covers the majority of Chinese citizens with over 0.89 billion monthly active users. Thanks to the wide coverage, users recorded by our dataset can be seen as a good representative of Chinese citizens. We select the data to focus on the largest metropolis Beijing, which is ideal for studying metropolis-level living habits of human beings.

TABLE 3
Data Summary

Coverage	Beijing
Record duration	Sep 17-Oct 31, 2016
Number of users	123,803
Number of unique POIs on weekdays	63,966
Number of unique POIs on weekends	61,827
Number of first-level POI type	17
Number of second-level POI type	189
Number of third-level POI type	405

The obtained spatial-temporal dataset is recorded whenever the users make requests on localization service in different platforms of the same service provider, such as location sharing, location check-ins, location-based social network, etc. GPS information at the timestamp is therefore recorded by localization modules. As the localization is achieved through both GPS and network-based approaches, the obtained location information is of fine-grained spatial granularity. In addition to GPS data, POI information, such as the name of a restaurant, or a specific address, is often recorded at the same time, thus adding semantics to spatial temporal information. Tencent provides a POI dictionary, which maps each POI to POI types of three levels. The number of POI types of the three level is 17, 189 and 405, respectively. For the first-level POI type, the categorization is coarse and covers major categories such as life service, company, real estate, etc, while the second and third-level POI type are much more fine-grained (e.g., distinguish different kinds of restaurants). Some examples of Tencent POI dictionary are listed in Table 2. One key issue with this provided POI dictionary, however, is that many POI categorization are not independent, even for the first level. For instance, office building belongs to the first-level POI type 'real estate', but is closely related to first-level POI type 'company', thus leads to difficulty in living pattern recognition. As we will show later, our proposed habit2vec addresses this challenge with good results.

In the experiment, we implement our system on 123,803 randomly sampled users in Beijing, whose records range from September 17, 2016 to October 31, 2016. All user information in the dataset have been anonymized for privacy concerns. A detailed description of the dataset is summarized in Table 3.

5.2 Data Preprocessing, Habit2vec and Clustering

Based on the POI information from the raw dataset, we use the third-level POI dictionary to convert POI transitions into finest POI type transitions. Then we discretize a day into 48 equal-length time slices (30 minutes every time slice) and aggregate the POI type transitions into a weekday and a weekend day. The parameter is set as 48 as it provides finest

TABLE 2
Example of Tencent POI Dictionary

First-level POI	Restaurant, company, real estate, service, entertainment, school, commercial, infrastructure
Second-level POI	Chinese restaurant, factory, residence, office building, post office, university, bank, shop, transportation
Third-level POI	Beijing restaurant, factory, villa, office building, industrial park, post office, university, bank, supermarket, market, airport

TABLE 4
Similarity Between 5 Major POI Types at 3 am, 10 am, and 8 pm on Weekday

	residence	university	commercial	restaurant	entertainment
residence	1	(-0.03,0.12,0.21)	(-0.08,0.29,-0.25)	(-0.01,-0.06,0.1)	(-0.41,0.11,0.21)
university	(-0.03,0.12,0.21)	1	(0.23,0.32,-0.16)	(0.11,0.02,0.12)	(0.05,0.04,0.12)
commercial	(-0.08,0.29,-0.25)	(0.23,0.32,-0.16)	1	(0.33,-0.08,0.11)	(0.29,0.06,0.10)
restaurant	(-0.01,-0.06,0.1)	(0.11,0.02,0.12)	(0.33,-0.08,0.11)	1	(0.29,-0.1,0.51)
entertainment	(-0.41,0.11,0.21)	(0.05,0.04,0.12)	(0.29,0.06,0.10)	(0.29,-0.1,0.51)	1

granularity and explainability without suffering from data sparsity. We mark time slices missing POI information as "missing" type in the dataset. In the acquired POI type transition trace, 5.2 percent of the time slices in the aggregated weekday and 8.1 percent of the time slices in the aggregated weekend day is marked missing.

As people follow very different living patterns on weekday and weekend, which in turn affects the POI transitions and the context of each habit unit, we separately adopt habit2vec representation learning on weekday and weekend data. After obtaining the vector representation for each habit unit, we take the mean of all habit units of a user as the vector representation for his/her living habit, and then adopt K-means clustering algorithm (with cosine similarity metric) separately on weekday and weekend trace to find groups of similar living habits on weekday/weekend. Following practices in word2vec[34], [35], We set the representation vector length dim as 80, and the window size w as 3 (1.5 hours). As revealed by previous works, a larger dim enables a better preservation of original word semantics, yet the gain is limited when dim is sufficiently large. For window size w, if the value is too small, correlation between habit units will not be properly captured; on the other hand, a too large window size will lead to over estimation of correlation. In the experiment, we carefully tuned these parameters so that optimal performance is achieved in measuring habit unit similarity and identifying living patterns. Finally, we choose the optimal number of cluster K through elbow method [8]. By increasing K from 2 to 30, and calculating the sum of error from user samples to the cluster center, we choose K at elbow point where the sum of error does not drop significantly compared to other points as the number of clusters. In this way, we determine the number of habit clusters for a weekday as 13 and the number of habit clusters for a weekend as 12.

5.3 Results Analysis

5.3.1 Habit Unit Embedding

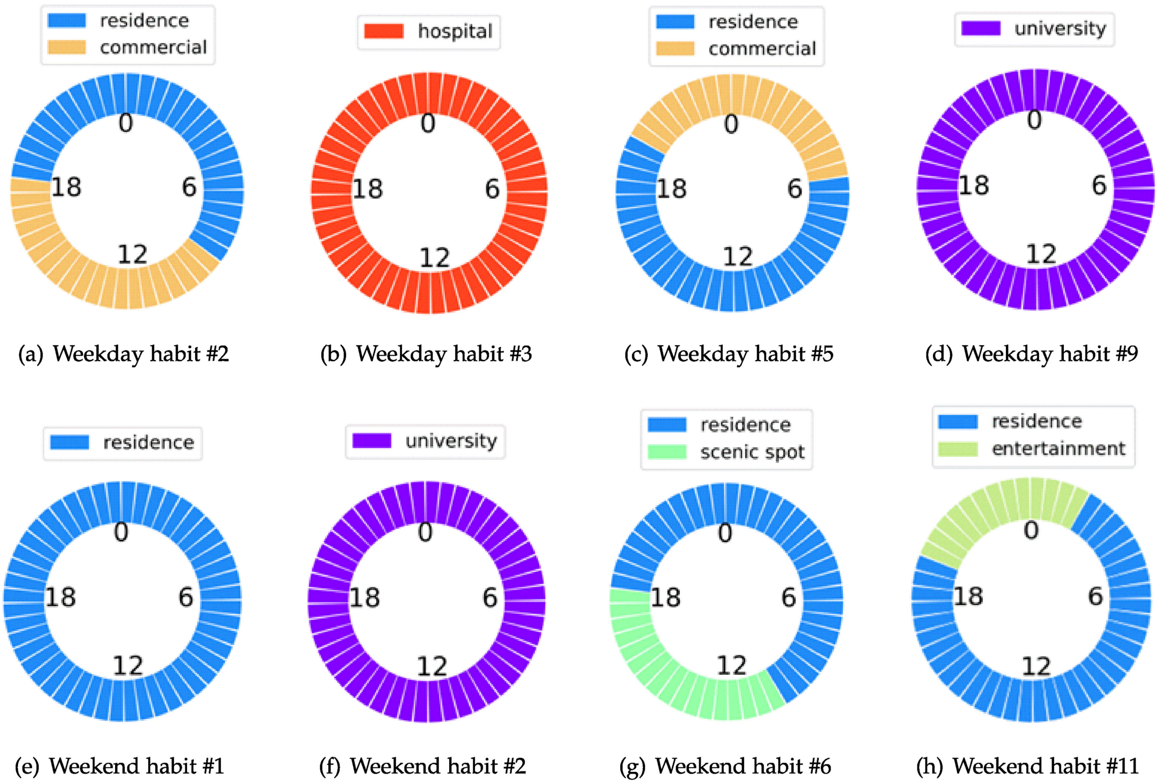
One main objective to adopt habit2vec is to measure the semantic similarity between POI types under the variation of time. We first take a specific POI type as an example to check the effect of habit2vec in distinguishing semantic difference. For instance, first-levl POI type restaurant is subdivisioned into different styles as a Beijing style restaurant, Hunan (a Chinese province) style restaurant, Pizza, etc. We measure the cosine similarity (value between -1 and 1 where the greater the value, the closer is the relationship between the two features) betweeen a Beijing style restaurant and a Hunan style restaurant at 12 noon. The similarity is 0.86, which implies a great similarity. On the other hand, the similarity between a Beijing style restaurant and Pizza (both are restaurants, but they have customers with different purposes, where Pizza is fast food while a Beijing style restaurant is much more formal) at 12 noon is 0.34 while the similarity between a Beijing style restaurant and a factory (they have no relationship) at 12 noon is -0.45. Habit2vec also measures temporal difference. The similarity between a residence at 1 am and a residence at 11:30 pm (both late night) on a weekday is 0.79, while the similarity between a residence at 1 am (late night) and 8 am (morning rush hour) on a weekday is 0.22.

We further test habit unit vectors on a global scale. We select 5 representative major POI types from the first-level and second-level POI dictionary: residence, university, commercial, restaurant and entertainment and check their similarity at different time slices on weekday. For top-level POI type such as restaurant and entertainment (which includes low-level POI types as cinema, club, etc.), we take the mean of all their subdivision POI types at the same time slice as its representation. The cosine similarity between the five POI types at 3 am, 10 am and 8 pm is shown in Table 4. Each non-diagonal cell in the table has three components, referring to the similarity between the two POI type at 3 am, 10 am and 8 pm. For instance, the element at the second row third column (-0.08,0.29,-0.25) means the similarity between residence and commercial district is -0.08 at 3 am, 0.29 at 10 am and -0.25 at 8 pm. From Table 4, we observe that the result of habit2vec is in accordance with our expectation. The five major POI types have quite clear semantic difference, which is reflected in the fact that most elements in the table are much less than 1.

On the other hand, habit2vec has the ability to distinguish semantic variations of POI type at different time. For instance, the similarity between restaurant and entertainment is 0.29 at 3 am, -0.1 at 10 am while 0.51 at 8 pm. This is consistent with our intuition: being in restaurant could simply mean filling the stomach, hanging in out with friends or doing a job (chef) while being in entertainment zone is closely related to entertaining with friends. If a person goes to a restaurant at night, he/she is likely to meet with friend and therefore similar in the purpose of going to entertainment zones. If a person appears in a restaurant 10 am in the morning, either because he/she is hungry or the person works in the restaurant, which is quite different from being in entertainment zones. The semantic difference of POIs at different times, is therefore successfully embedded in our habit2vec approach.

5.3.2 Label User Habit Clusters

We further check the performance of user habit vectors obtained from habit2vec. We implement K-means clustering method separately on weekday and weekend user living habit representation and obtain 13 weekday habit clusters (weekday living patterns) and 12 weekend habit clusters (weekend living patterns). Then, we determine the semantics, or the label of each habit cluster based on 2 criteria.



(a) Weekday habit #2 (b) Weekday habit #3 (c) Weekday habit #5 (d) Weekday habit #9

(e) Weekend habit #1 (f) Weekend habit #2 (g) Weekend habit #6 (h) Weekend habit #11

Fig. 5. Four weekday and four weekend living patterns detected by habit2vec.

- The statistics of POI types in the cluster;
- The living habit closest to the cluster center.

For instance, we determine the label of weekday cluster #2 and cluster #5 through the following way. We first make a statistics of the POI type in cluster #2 and #5. We find that POI type commercial building accounts for 36.7 percent, while residence accounts for 52.8 percent of all POI types appearing in cluster #2. In the meantime, POI type commercial building accounts for 25.3 percent, while residence accounts for 45.3 percent of all POI types appearing in cluster #5. As we are not able to tell the difference between semantics of the two clusters, we move on using the cluster center temporal information. We find out that the living habit of user whose habit vector is closest to the vector of the cluster #2 center stays in residence at night and goes to commercial building in the day (Fig. 5a), while user closest to cluster #5 is on the contrary (Fig. 5c). Thus, we label them differently. This example shows that habit2vec clustering is good at measuring difference in time scheduling.

In this way, we label semantics of all weekday and weekend cluster. The semantics, possible identity as well as the population proportion of the 13 weekday living habit clusters are summarized in Table 5 while the semantics and the population proportion of the 12 weekend living habit clusters are summarized in Table 6. The cluster results show that habit2vec not only captures distinctions in users' visits to different POI types, but also distinguishes schedule difference.

Note that we are using cluster center to represent typical living habit in the city, which turn out to be simple enough to be represented by one or two POI types. However, not all users in the a living habit cluster are similar to cluster center in terms of POI type labels. For instance, in weekday habit cluster #1, we observe people who spend most of their day time at auto repair, auto service, local market, pharmacies, etc. (different from habit center POI type "shops"), but they don't show up in other habit clusters. They belong to cluster #1 as the semantics of working in these POI types during the day are much more similar to working in shops than working elsewhere (other habit clusters). In addition, users in a living habit cluster generally have variations from cluster

TABLE 5
Semantics, Possible Identity and Population Proportion of 13 Weekday Living Habit Clusters

ID	Semantics	Possible Identity	%
#1	stay in residence at night, stay in shops for the day	shop owner, shop assistant	5.7%
#2	stay in residence at night, stay in commercial building for the day	white collar	7.1%
#3	stay in hospital the whole day	doctor, nurse, patient	1.0%
#4	stay in industrial zone whole day	engineer, laborer	4.1%
#5	stay in commercial building at night, in residence for the day	white collar	7.3%
#6	stay in residence the whole day	retired, freelance	28.0%
#7	POI type missing	-	5.9%
#8	stay in residence most of the day, go to commercial buildings briefly	senior white collar	4.7%
#9	stay in university the whole day	college student	9.8%
#10	stay in residence most of the day, go to shops briefly in the day	retired, freelance	14.1%
#11	irregular life, skip from one POI type to another	people leading irregular life	2.0%
#12	stay in suburb residence most of the day, go to market in the day	local business owner	2.9%
#13	stay in residence at night, go to schools in the day	teacher, student	7.4%

TABLE 6
Semantics and Population Proportion of 12 Weekend Living Habit Clusters

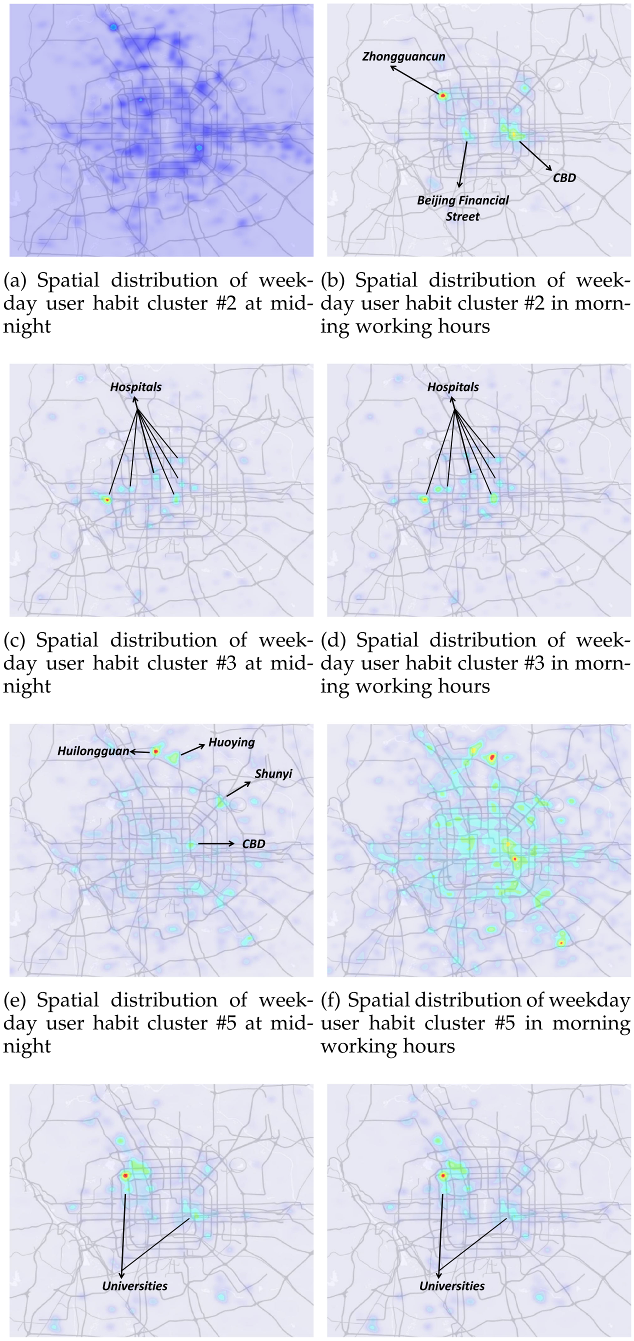
ID	Semantics	%
#1	stay in residence the whole day	34.0%
#2	stay in university the whole day	7.1%
#3	stay in residence at night, go shopping in the day	6.2%
#4	stay in residence at night, go to gym in the day	1.8%
#5	stay in university most of the day, go shopping briefly in the day	6.8%
#6	stay in residence at night, go to scenic spot in the day	3.9%
#7	stay in residence at night, go to university, shops in the day	8.8%
#8	stay in industrial zone the whole day	3.2%
#9	stay in residence most of the day, go shopping in the evening	8.5%
#10	POI type missing	7.8%
#11	stay in residence most of the day, stay in entertainment zones in the evening	7.3%
#12	stay in residence at night, go shopping and gym in the day	4.6%

center, e.g., different users may go to different type of restaurant at different time slices. Despite the variations, their general living habit fit the pattern defined by cluster center. Therefore, habit2vec is capable of capturing fine-grained semantic similarity dynamically in user habit without prior knowledge, compared with previous works which directly consider static POI type labels (often involve manually grouping POI labels).

It is worth mentioning, however, due to lack of information and the unsupervised nature of habit2vec, some recognized living patterns may represent people of different identities. For instance, weekday habit cluster #3 represents people who stay in hospital the whole day, and they could be doctors, nurses or patients. Also, habit2vec is not able to distinguish fine-grained location semantics in some cases, e.g., dormitories and lecture halls in universities, since they have the same POI label ‘university’. Nevertheless, we argue that habit2vec, as many other successful data mining techniques, nicely completes ‘search and filter’, ‘read and extract’ and ‘schematize’, and helps establish reasonable hypotheses in the sensemaking process [41] with minimal manual effort. To further support or disconfirm hypotheses generated by habit2vec on user identities, more information and domain knowledge is needed.

5.3.3 *Spatial Analysis on User Habit Cluster*

We visualize the spatial distribution of four interesting habit clusters for weekday (Fig. 6), and four habit clusters for weekend (Fig. 7) at midnight (2 am) and in the morning (10:30 am). We further evaluate the clustering performance of habit2vec, and obtain the following featured clusters.

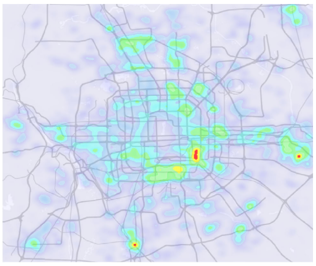


(a) Spatial distribution of weekday user habit cluster #2 at midnight
(b) Spatial distribution of weekday user habit cluster #2 in morning working hours
(c) Spatial distribution of weekday user habit cluster #3 at midnight
(d) Spatial distribution of weekday user habit cluster #3 in morning working hours
(e) Spatial distribution of weekday user habit cluster #5 at midnight
(f) Spatial distribution of weekday user habit cluster #5 in morning working hours
(g) Spatial distribution of weekday user habit cluster #9 at midnight
(h) Spatial distribution of weekday user habit cluster #9 in morning working hours

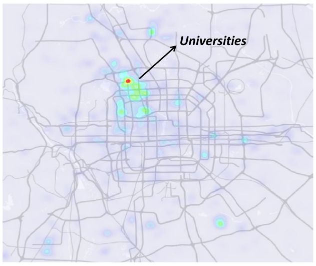
Fig. 6. Spatial distribution of four weekday living habit user clusters at midnight and in morning working hours.

- Weekday cluster #2 represents people who stay in residence at night and stay in commercial building for the day. The spatial distribution of the cluster at midnight spreads across the city. While in the morning working hours, users of this cluster aggregates in commercial center as CBD and Zhongguancun.
- Weekday cluster #3 represents people who stay in hospital for the entire day. The spatial distribution of the cluster at midnight looks almost the same as the cluster in the morning, and the distribution resembles the distribution of Beijing’s hospitals.
- Weekday cluster #5 represents people staying in commercial building at night and going back to residence in the day. The spatial distribution of the cluster is the opposite of Weekday cluster #2, despite this group concentrates more in suburban commercial center as Huoyin.
- Weekday cluster #9 represents people staying in university for the entire day. The spatial distribution of

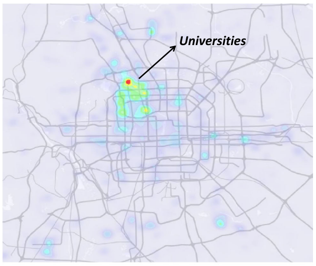
(a) Spatial distribution of weekend user habit cluster #1 at midnight



(b) Spatial distribution of weekend user habit cluster #1 in the morning

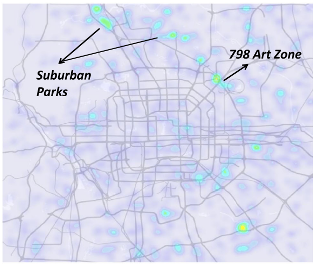


(c) Spatial distribution of weekend user habit cluster #2 at midnight

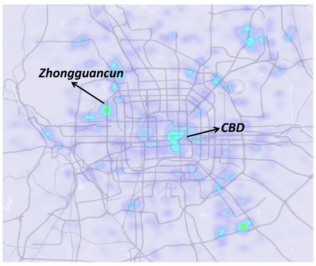


(d) Spatial distribution of weekend user habit cluster #2 in the morning

(e) Spatial distribution of weekend user habit cluster #6 at midnight



(f) Spatial distribution of weekend user habit cluster #6 in the morning



(g) Spatial distribution of weekend user habit cluster #11 at midnight

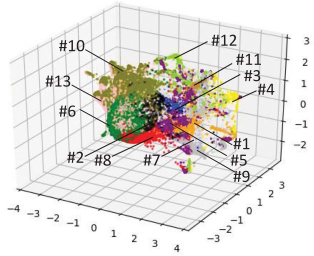
(h) Spatial distribution of weekend user habit cluster #11 in the morning

Fig. 7. Spatial distribution of four weekend living habit user clusters at midnight and in the morning.

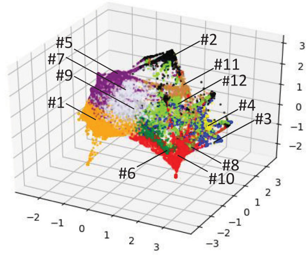
the cluster is the same as university distribution in Beijing.

- Weekend cluster #1 represents people staying in residence for the entire day. This habit group covers the whole city.
- Weekend cluster #2 represents people staying in university for the entire day. The spatial distribution of the cluster is quite similar to weekday cluster #9.
- Weekend cluster #6 represents people staying in residence at night who go to scenic spot in the day. The spatial distribution of the cluster in the day highlights the scenic spots in suburb Beijing.
- Weekend cluster #11 represents people staying in residence in the day and going out to entertainment district at night. The spatial distribution of the cluster is similar to the spatial distribution of Beijing's entertainment district.

From the above analysis, we come to the conclusion that the spatial distribution of habit clusters is in conformity with the habit labels we assigned previously.



(a) Weekday



(b) Weekend

Fig. 8. Visualization of 3D structure of weekday and weekend habit clusters.

5.3.4 Visualization of Habit Cluster Structure

We also visualize the low-dimensional structure of weekday and weekend habit clusters. We adopt the widely used high-dimensional data visualization technique, t-SNE[42], to project user habit vector obtained from habit2vec on 3D space. Figs. 8a and 8b show the distribution of weekday user habit vector and weekend user habit vector respectively, where each point in the 3D space represents the weekday/weekend living habit of a user and points with the same color stand for users of the same detected living pattern. As demonstrated by Fig. 8, users of the same weekday/weekend living pattern aggregate while users of different living patterns are disperse. Therefore, each habit cluster can be represented by its centroid, which proves the effectiveness of habit2vec in representing user living habit and finding distinct living patterns in population.

5.3.5 Relationship between Weekday and Weekend

We further evaluate the performance of habit2vec by studying the relationship between weekday and weekend user habit clusters. As mentioned earlier, we notice the difference between weekday and weekend POI transition mode and therefore separately train weekday and weekend habit unit vectors. We notice that there is a strong correlation between weekday cluster #9 (university whole day) and weekend cluster #2 (university whole day) and cluster #5 (university +briefly shopping in the day), where 64 percent of people in weekday cluster #9 appears in weekend cluster #2 while 15 percent of people in weekday cluster #9 appears in weekend cluster #5. There is also a correlation between weekday cluster #6 and weekend cluster #1 (both means staying in residence for the whole day), where 63 percent of people in cluster #6 end up in weekend cluster #1. On the other hand, there is no simple one-one or one-two matching between other weekday and weekend habit clusters, indicating that other groups have more flexibility in living style. The result of

TABLE 7
Relationship between Weekday and Weekend Living Habit Clusters

weekday \ weekend	#1	#2	#5	#9
#2	31.2%	1.0%	0.5%	10.2%
#6	62.6%	0.1%	0.3%	7.3%
#8	28.9%	0.2%	1.1%	6.7%
#9	4.2%	64.2%	14.8%	1.2%

weekday and weekend living pattern clusters is shown in Table 7. Each element in the table refers to the percentage of people in a weekday habit cluster who belongs to a weekend habit cluster. For instance, element 0.1 percent in the third row third column in the table means that 0.1 percent people in weekday cluster #6 belongs to weekday cluster #2. In summary, the results of weekday-weekend habit cluster relationship lives up to our expectation and highlights the complex composition of users' weekday and weekend living patterns.

5.3.6 Evaluation via Ground Truth

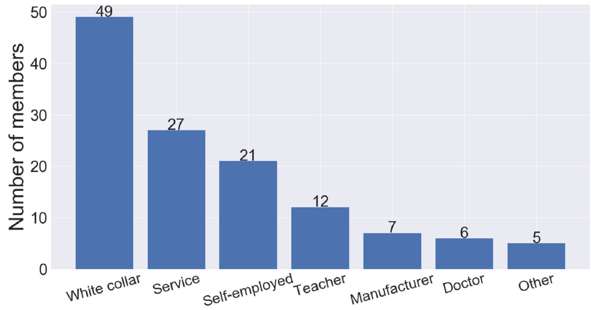
Finally, we evaluate the living patterns recognized through habit2vec via ground truth data.

First, we randomly select 200 users from our dataset and ask 20 volunteers (8 female and 12 male) to label the most likely living patterns of those users by viewing their weekday signature habit unit transition (the 48-length POI type series, whose components represent the most likely POI type the user visits during a specific 30-minute time slice on weekdays). Of the 200 users, 189 users' identities labelled by volunteers match the results recognized through habit2vec, with an accuracy of 94.5 percent. This result further verifies that habit2vec is capable of recognizing meaningful living patterns.

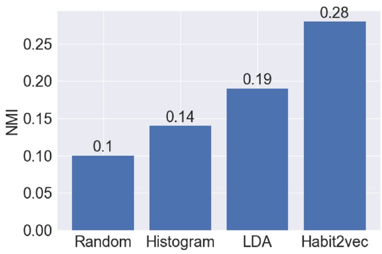
Then, we compare the performance of habit2vec over baseline algorithms. To the best of our knowledge, there is one existing work [25] aiming at approaching a similar living pattern recognition problem, which proposes a PCA-based method to extract behavior patterns for 100 people with 5 location semantics ('work', 'home', 'no signal', 'else' and 'off'). However, the method requires constructing a user-feature matrix, which is of high space complexity. In fact, running the method on our dataset is far beyond the capacity of normal machine's memory. We therefore leverage the following two baseline methods as comparison instead:

Histogram. For each user, we count the frequency the user visits each type of POI in his/her signature habit trace, and form user POI type visit histogram feature to represent his/her habit. Then we perform K-means clustering on this histogram feature. We set the clustering number k as 13 to be the same as the number of weekday living patterns recognized by habit2vec.

LDA. Latent Dirchlet Allocation (*LDA*) [43] is a classical unsupervised method in natural language processing to extract semantic features for documents, which has been recently widely adopted in spatial temporal data mining and user profiling [2], [26]. We first use *LDA* on user POI transition trace, extract latent feature of each user and perform K-means clustering on the latent feature. We set the clustering number k as 13 to be the same as the number of



(a) Occupation distribution of self-reported ground truth data



(b) habit2vec vs. baselines

Fig. 9. Performance comparison of habit2vec over baseline methods.

weekday living patterns recognized by habit2vec, and set the number of latent feature as 12, which has been carefully tuned to obtain optimal performance.

As there are no available data on real user living pattern categorization, we adopt user self-reported occupation as ground truth so as to approximate his/her weekday living pattern. We acquired self-reported occupations of 127 anonymized users recorded in our dataset from Tencent, with 7 different types of occupation. We randomly selected users of different characteristics so as to ensure that the ground truth data is representative to the greatest extent possible. The distribution of different types of occupations of the ground truth is demonstrated in Fig. 9a.

We use normalized mutual information (*NMI*), a popular performance analysis metric in clustering analysis [44], to measure the performance of weekday living pattern recognition. The range of *NMI* is between 0 and 1, and a greater *NMI* value indicates a better match between identified living patterns and the ground truth user occupation types, e.g., if *NMI* reaches 1, it means that users within each identified living pattern cluster has the same type of occupation, and that all users with the same type of occupation fall into the same living pattern cluster. Denote Y_k as the set of users with living pattern k, Z_j as the set of users whose occupation type is j, and M as the total number of users, NMI is defined as follows:

$$I(Y,Z) = \sum_k \sum_j \frac{|Y_k \cap Z_j|}{M} \log \frac{M|Y_k \cap Z_j|}{|Y_k| \cap |Z_j|},$$

$$H(*) = -\sum_{X_i \in *} \frac{X_i}{M} \log \frac{X_i}{M},$$

$$NMI(Y,Z) = \frac{2 \times I(Y,Z)}{H(Y)+H(Z)},$$

where $I(Y, Z)$ is mutual information between identified patterns and ground truth occupation categories while $H(*)$ is the entropy.

We illustrate the performance of random assignment, histogram feature baseline, LDA feature baseline and habit2vec in Fig. 9. We observe that habit2vec outperforms histogram and LDA feature by a large margin, with 100 and 47.4 percent performance gain, respectively, which indicates that habit2vec can identify user living pattern much better than baseline approaches. We attribute this performance gain to the fact that habit2vec can capture the temporal and semantic correlation between different type of POIs. Note that in the current evaluation process, we use the plug-in estimator for probability distribution estimation when calculating entropy, which may subject to bias and errors given the size of ground truth data. We anticipate more evaluation work on habit2vec when better ground truth data becomes available.

6 Conclusions

In this paper, we used semantic information embedded in trajectories to identify typical living patterns in a population. We proposed a representation learning method called habit2vec to mine the users' signature living habit to embed semantics and time in the same space. We evaluated the effectiveness of our proposed system based on a real-world dataset with 123,803 users, and successfully discovered 13 and 12 meaningful weekday and weekedn living patterns respectively. The experiment showed that habit2vec is capable of preserving both semantics and time information in users' living habit. In the future, we plan to predict the career and social-economic status based on living habits recognized by our system.

Acknowledgments

This work was supported in part by the National Key Research and Development Program of China under grant 2017YFE0112300, the National Nature Science Foundation of China under 61861136003, 61621091, and 61673237, the Beijing National Research Center for Information Science and Technology under 20031887521, and research fund of Tsinghua University - Tencent Joint Laboratory for Internet Innovation Technology, and the US National Science 924 Foundation under grants IIS-13-20791 and IIS-1816889.

References

[1] N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A. T. Campbell, "A survey of mobile phone sensing," *IEEE Commun. Mag.*, vol. 48, no. 9, pp. 140–150, Sep. 2010.

[2] J. Yuan, Y. Zheng, and X. Xie, "Discovering regions of different functions in a city using human mobility and POIs," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2012, pp. 186–194.

[3] C. Zhang, K. Zhang, Q. Yuan, H. Peng, Y. Zheng, T. Hanratty, S. Wang, and J. Han, "Regions, periods, activities: Uncovering urban dynamics via cross-modal representation learning," in *Proc. 26th Int. Conf. World Wide Web*, 2017, pp. 361–370.

[4] H. Cao, J. Feng, Y. Li, and V. Kostakos, "Uniqueness in the city: Urban morphology and location privacy," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 2, 2018, Art. no. 62.

[5] B. Wegener, "Job mobility and social ties: Social resources, prior job, and status attainment," *Amer. Sociol. Rev.*, vol. 56, no. 1, pp. 60–71, 1991.

[6] P. S. Adler, "Social capital: Prospects for a new concept," *Academy Manage. Rev.*, vol. 27, no. 1, pp. 17–40, 2002.

[7] J. G. Lee, J. Han, and K. Y. Whang, "Trajectory clustering: A partition-and-group framework," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2007, pp. 593–604.

[8] Y. Di, Z. Chao, Z. Zhihua, H. Jianhui, and B. Jingping, "Trajectory clustering via deep representation learning," in *Proc. Int. Joint Conf. Neural Netw.*, 2017, pp. 3880–3887.

[9] S. Jiang, "Clustering daily patterns of human activities in the city," *Data Mining Knowl. Discovery*, vol. 25, no. 3, pp. 478–510, 2012.

[10] B. Furletti, P. Cintia, C. Renso, and L. Spinsanti, "Inferring human activities from GPS tracks," in *Proc. ACM SIGKDD Int. Workshop Urban Comput.*, 2013, Art. no. 5.

[11] N. Mamoulis, H. Cao, G. Kollios, M. Hadjieleftheriou, Y. Tao, and D. W. Cheung, "Mining, indexing, and querying historical spatio-temporal data," in *Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2004, pp. 236–245.

[12] S. Nutanong, E. H. Jacox, and H. Samet, "An incremental Hausdorff distance calculation algorithm," *Proc. VLDB Endowment*, vol. 4, no. 8, pp. 506–517, 2011.

[13] H. Jeung, L. Y. Man, and C. S. Jensen, "Trajectory pattern mining," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2007, pp. 330–339.

[14] Y. Zheng, L. Zhang, X. Xie, and W. Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in *Proc. Int. Conf. World Wide Web*, 2009, pp. 791–800.

[15] A. Sadilek and J. Krumm, "Far out: Predicting long-term human mobility," in *Proc. AAAI Conf. Artif. Intell.*, 2012, pp. 814–820.

[16] K. Zhang, Y. R. Lin, and K. Pelechrinis, "EigenTransitions with hypothesis testing: The anatomy of urban mobility," in *Proc. 10th Int. AAAI Conf. Web Social Media*, 2016, pp. 486–495.

[17] C. Zhang, K. Zhang, Q. Yuan, L. Zhang, T. Hanratty, and J. Han, "GMove: Group-level mobility modeling using Geo-tagged social media," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, pp. 1305–1314.

[18] J. C. Ying, W. C. Lee, T. C. Weng, and V. S. Tseng, "Semantic trajectory mining for location prediction," in *Proc. ACM SIGSPATIAL Int. Conf. Advances Geographic Inf. Syst.*, 2011, pp. 34–43.

[19] J. J.-C. Ying, W.-C. Lee, and V. S. Tseng, "Mining geographic-temporal-semantic patterns in trajectories for location prediction," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 1, 2013, Art. no. 2.

[20] C. Zhang, J. Han, L. Shou, J. Lu, and T. L. Porta, "Splitter: Mining fine-grained sequential patterns in semantic trajectories," *Proc. VLDB Endowment*, vol. 7, no. 9, pp. 769–780, 2014.

[21] Q. Yuan, W. Zhang, C. Zhang, X. Geng, G. Cong, and J. Han, "PRED: Periodic region detection for mobility modeling of social media users," in *Proc. 10th ACM Int. Conf. Web Search Data Mining*, 2017, pp. 263–272.

[22] C. Parent, S. Spaccapietra, C. Renso, G. Andrienko, N. Andrienko, V. Bogorny, M. L. Damiani, A. Gkoulalas-Divanis, J. Macedo, and N. Pelekis, "Semantic trajectories modeling and analysis," *ACM Comput. Surveys*, vol. 45, no. 4, 2013, Art. no. 42.

[23] B. T. Pentland and H. H. Rueter, "Organizational routines as grammars of action," *Administ. Sci. Quart.*, vol. 39, no. 3, pp. 484–510, 1994.

[24] N. Li, S. Kambhampati, and S. Yoon, "Learning probabilistic hierarchical task networks to capture user preferences," in *Proc. Int. Jont Conf. Artif. Intell.*, 2009, pp. 1754–1759.

[25] N. Eagle and A. S. Pentland, "Eigenbehaviors: Identifying structure in routine," *Behavioral Ecology Sociobiology*, vol. 63, no. 7, pp. 1057–1066, 2009.

[26] K. Farrahi and D. Gaticaperez, "Extracting mobile behavioral patterns with the distant N-Gram topic model," in *Proc. Int. Symp. Wearable Comput.*, 2012, pp. 1–8.

[27] N. Banovic, T. Buzali, F. Chevalier, J. Mankoff, and A. K. Dey, "Modeling and understanding human routine behavior," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2016, pp. 248–260.

[28] C. Koehler, N. Banovic, I. Oakley, J. Mankoff, and A. K. Dey, "Indoor-ALPS: An adaptive indoor location prediction system," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 171–181.

[29] J. Scott, A. J. B. Brush, J. Krumm, B. Meyers, M. Hazas, S. Hodges, and N. Villar, "PreHeat: Controlling home heating using occupancy prediction," in *Proc. Int. Conf. Ubiquitous Comput.*, 2011, pp. 281–290.

[30] A. S. Furtado, D. Kopanaki, L. O. Alvares, and V. Bogorny, "Multidimensional similarity measuring for semantic trajectories," *Trans. GIS*, vol. 20, no. 2, pp. 280–298, 2016.

[31] H. Cao, Z. Chen, F. Xu, Y. Li, and V. Kostakos, "Revisitation in urban space versus online: A comparison across POIs, websites, and smartphone apps," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 4, 2018, Art. no. 156.

[32] F. Xu, T. Xia, H. Cao, Y. Li, F. Sun, and F. Meng, "Detecting popular temporal modes in population-scale unlabelled trajectory data," *Proc. ACM Interactive Mobile Wearable Ubiquitous Technol.*, vol. 2, no. 1, 2018, Art. no. 46.

[33] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.

[34] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, pp. 3111–3119, 2013.

[35] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *arXiv preprint arXiv:1301.3781*, 2013.

[36] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Language Process.*, 2014, pp. 1532–1543.

[37] B. Perozzi, R. Alrfou, and S. Skiena, "DeepWalk: Online learning of social representations," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2014, pp. 701–710.

[38] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei, "LINE: Large-scale information network embedding," in *Proc. Int. Conf. World Wide Web*, 2015, pp. 1067–1077.

[39] C. Song, Z. Qu, and N. Blumm, "Limits of predictability in human mobility," *Sci.*, vol. 327, no. 5968, pp. 1018–1021, 2010.

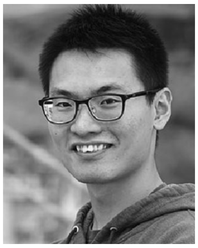
[40] S. T. Piantadosi, "Zipf's word frequency law in natural language: A critical review and future directions," *Psychonomic Bulletin Rev.*, vol. 21, no. 5, pp. 1112–1130, 2014.

[41] P. Pirolli and S. Card, "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis," in *Proc. Int. Conf. Intell. Anal.*, 2005, pp. 2–4.

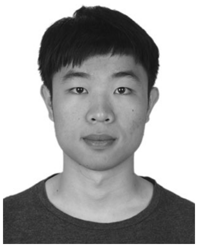
[42] L. Maaten and G. Hinton, "Visualizing data using t-SNE," *J. Mach. Learn. Res.*, vol. 9, no. 2605, pp. 2579–2605, 2008.

[43] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, no. Jan., pp. 993–1022, 2003.

[44] N. X. Vinh, J. Epps and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *Journal of Machine Learning Research*, vol. 11, pp. 2837–2854, Oct 2010.



Hancheng Cao received the BS degree in electronic information science and technology from Tsinghua University, Beijing, China, in 2018, and he is currently working toward the PhD degree in Computer Science Department, Stanford University, Stanford, California. His research interests include data mining, ubiquitous computing, and computational social science.



Fengli Xu received the BS degree in electronics and information engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2015, and he is currently working toward the PhD degree in the Electronic Engineering Department, Tsinghua University, Beijing, China. His research interests include human mobility, mobile big data mining, and user behavior modelling.

Jagan Sankaranarayanan received the PhD degree in computer science from the University of Maryland, in 2008. He currently works at Google on their data infrastructure team. His research interests include database management problems related to GIS, big data, analytics, and mobility. He received the Best Paper Award for ACM SIGMOD, SIGSPATIAL, *Computers & Graphics Journal*, the 10 year Impact Award from SIGSPATIAL, and a Best Invention of Year Award from NEC Labs.

Yong Li (M'09-SM'16) received the BS degree in electronics and information engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2007 and the PhD degree in electronic engineering from Tsinghua University, Beijing, China, in 2012. He is currently a faculty member of the Department of Electronic Engineering, Tsinghua University. Dr. Li has served as general chair, TPC chair, and TPC member for several international workshops and conferences, and he is on the editorial board of two IEEE journals. His papers have total citations of more than 6,000. Among them, 10 are ESI Highly Cited Papers in Computer Science, and four received conference Best Paper (run-up) Awards. He received the IEEE 2016 ComSoc Asia-Pacific Outstanding Young Researchers and Young Talent Program of China Association for Science and Technology.

Hanan Samet received the BS degree in engineering from UCLA, and the MS degree in operations research and the MS and PhD degrees in computer science from Stanford University. He is a distinguished university professor of computer science with the University of Maryland, College Park. His doctoral dissertation dealt with proving the correctness of translations of LISP programs which was the first work in translation validation and the related concept of proof-carrying code. He is the author of the recent book *Foundations of Multidimensional and Metric Data Structures* published by Morgan-Kaufmann, an imprint of Elsevier, in 2006, an award winner in the 2006 best book in computer and information science competition of the Professional and Scholarly Publishers (PSP) Group of the American Publishers Association (AAP), and of the first two books on spatial data structures *Design and Analysis of Spatial Data Structures*, and *Applications of Spatial Data Structures: Computer Graphics, Image Processing, and GIS*, both published by Addison-Wesley, in 1990. He is the founding editor-in-chief of the *ACM Transactions on Spatial Algorithms and Systems* (*TSAS*), the founding chair of ACM SIGSPATIAL, a recipient of the 2009 UCGIS Research Award, the 2011 ACM Paris Kanellakis Theory and Practice Award, the IEEE Computer Society's 2014 Wallace McDowell Award, and a fellow of the ACM, IEEE, AAAS, International Association for Pattern Recognition (IAPR), and UCGIS. He received best paper awards in the 2007 *Computers & Graphics Journal*, the 2008 ACM SIGMOD and SIGSPATIAL Conferences, the 2012 SIGSPATIAL MobiGIS Workshop, the 2013 SIGSPATIAL GIR Workshop, as well as best demo paper awards at the 2011 and 2016 ACM SIGSPATIAL Conferences, and one of the best papers in the 2009 IEEE ICDE Conference selected for publication in the *IEEE Transactions on Knowledge and Data Engineering*. He was elected to the ACM Council as the Capitol Region Representative for the term 1989-1991, and is an ACM distinguished speaker.