

Revealing Urban Dynamics by Learning Online and Offline Behaviours Together

TONG XIA and YONG LI, Beijing National Research Center for Information Science and Technology, Department of Electrical Engineering, Tsinghua University, China

Urban problems and diseases accompanied by the pace of urbanization have drawn attention to the importance of understanding urban dynamics, while a deep and comprehensive understanding is challenging due to our diversified lifestyles in the modern city. In this paper, we propose an urban dynamics modeling system to characterize the regularity of urban activity dynamics as well as urban functions by learning residents' online and offline behaviours together. Built on a state-sharing hidden Markov model, our system utilizes online activities of App usage and offline activities of mobility in different urban regions and different time slots for learning. The learnt state sequence of each region reveals urban dynamics with the corresponding urban functions. We evaluate our system via a large-scale mobile network accessing dataset, which discovers ten hidden states characterizing different life modes and eight representative dynamic patterns corresponding to different urban functions. These discovered dynamic patterns and inferred functions are validated by social media check-ins and the land-use published by the government with 81% accuracy. Based on our model, we propose two applications, crowd flow prediction and popular App prediction, which outperforms the state-of-the-art approaches by 36.1% and 15.7%, respectively. This study paves the way for extensive city-related applications including urban demand analysis, land-use planning, and activity prediction.

CCS Concepts: • **Information systems** → **Spatial-temporal systems**; **Data mining**; • **Human-centered computing** → *Empirical studies in ubiquitous and mobile computing*; • **Computing methodologies** → Machine learning.

Additional Key Words and Phrases: urban dynamics modeling and revealing, online and offline behaviours, state-sharing HMM, mobile cellular network accessing dataset.

ACM Reference Format:

Tong Xia and Yong Li. 2019. Revealing Urban Dynamics by Learning Online and Offline Behaviours Together. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 1, Article 30 (March 2019), 25 pages. <https://doi.org/10.1145/3314417>

1 INTRODUCTION

Revealing *urban dynamics*, i.e., how the types and intensity of human activities in the city change along with the time, has always been a crucial social-economic task for both researchers and governments [2, 24]. With the ever-increasing urbanization process, human activities in the city are becoming increasingly dynamic and complex, making it more difficult to understand and model [16]. Moreover, nowadays human daily activities include not only the commuting between home and office, meeting friends, shopping, etc. in the physical space, but also the checking-in, liking the online friends, buying and selling goods, etc. in the cyberspace [7, 26],

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, Beijing National Research Center for Information Science and Technology under 20031887521, and research fund of Tsinghua University - Tencent Joint Laboratory for Internet Innovation Technology. Authors' address: Tong Xia, xia-17@mails.tsinghua.edu.cn; Yong Li, liyong07@tsinghua.edu.cn, Beijing National Research Center for Information Science and Technology, Department of Electrical Engineering, Tsinghua University, Beijing, China.

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 the author(s) 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.

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2474-9567/2019/3-ART30 \$15.00

<https://doi.org/10.1145/3314417>

which are more high-dimensional and dynamic. Although uncovering the regularity of such daily activities is challenging, it is of great significance to tackle a series of urban problems, e.g., overcrowding, inadequate infrastructure, traffic congestion, pollution, etc. [50]. What's more, different kinds of dynamics are mainly caused by different land-use of the urban regions [45, 46], which indicates that revealing urban dynamics is directly related with the inference of urban functions to facilitate better urban planning [31].

Until now, most of our understanding about urban dynamics come from traditional surveys conducted by human agents [25]. While this way of collecting data provides detailed information about urban behaviours, it remains hard to update and presents many weaknesses regarding generalization and geographical scope. Luckily, with the ubiquitous mobile devices, massive data recording various human activities is available to reveal urban dynamics. Recently, related works aim to uncover the temporal regularity of human activities, such as detecting the patterns of mobility behaviours in the city via passive mobile positioning data [5], visualizing the level of activities at a given urban location across multiple temporal resolutions [2], etc. However, these existing studies only utilized offline behaviours for urban dynamics understanding by statistics and visualization methods, which cannot realize specific and predictable human activity modeling.

In this paper, our goal is to model urban dynamics with regard to human online and offline behaviours in the city and further to infer urban functions. Despite its practical importance, using both online and offline activities is non-trivial due to three **challenges**: 1) Urban residents' activities are dynamic and complex. Using which kinds of online and offline features to characterize the activity type and intensity is the first challenge. 2) Underlying the various activities, there are several basic city states characterizing different life modes such as busy working or peaceful sleeping, which compose urban dynamics. To model urban dynamics from the city scale, similar activities, whether in the same or different regions, should be detected as the same state. However, on the other hand, though the states are the same, the dynamic patterns of different regions could be different due to their urban functions. Moreover, the data of each region would be limited to learn its own states and dynamics. Therefore, how to build one robust model to learn the difference as well as the similarity is the second challenge. 3) The relation between urban dynamics and urban functions is implicit, and the function of each urban region could be single, compound, or even dynamic. How to infer it according to the learnt activity states is the third challenge.

To overcome these challenges, we propose an urban dynamics modeling system by leveraging the following three key designs. First, we select human online activities of App usage and offline activities of mobility as features. The intuitive motivations of utilizing these two features is that accessing to the Internet through Apps [1] and moving between different places [11] are the most important activities in the cyber and physical urban space, respectively. The App usage, i.e., how frequently the Apps are used, and the mobility, i.e., the volume of crowd flows, can reflect the type and the intensity of activities in a region. To extract these features, we utilize a mobile cellular network accessing dataset, which records when and where a user uses which Apps. Second, to model urban dynamics based on these features, we propose a state-sharing hidden Markov model (HMM). Regarding urban residents' online and offline activities as time series, we aim to detect a common set of hidden states characterizing basic life modes in the city and to reveal urban dynamics by the transition among these states. Particularly, the state set is shared by all regions yet each region has its own state sequence. Through this strategy, similarity and difference among different regions are learnt at the same time, and the problem of data sparsity is also solved. Third, to infer and interpret the urban functions from the learnt state sequences, we design a clustering algorithm to divide these state sequences into several typical dynamic patterns, where each pattern maps to a specific function. By combining the semantics of states in different time slots, the relations between urban dynamics and functions are uncovered. To summarize, the contribution of our work is four-fold:

- (1) We investigate the problem of understanding urban dynamics using human activities in both physical and cyber space recorded by mobile cellular network accessing dataset. To the best of our knowledge, this is the first study to utilize both online and offline behaviours to reveal urban dynamics.
- (2) We propose a novel urban dynamics modeling system based on the state-sharing HMM, where the states characterized by the type and intensity of human activities are shared by all urban regions, but each region has its own state sequence. It achieves qualitative representations of urban dynamics as the transitions between different states.
- (3) We evaluate our method by a real-world large-scale dataset in Shanghai, China. We have learned ten hidden states and discovered eight typical dynamic patterns corresponding to different urban functions. These discovered dynamic patterns and inferred functions are validated by social media check-ins and the land-use published by the government with 81% accuracy.
- (4) We design two applications based on our model, i.e., crowd flow prediction and popular App prediction. Extensive experiments demonstrate that our proposed state-sharing HMM outperforms the state-of-the-art approaches by 36.1% and 15.7%.

2 RELATED WORK

In this section, we introduce the related from three perspectives: urban dynamics modeling, urban functions inferring, and hidden Markov model with its application.

Urban dynamics modeling. Urban dynamics, generally defined as how sociological indicators (e.g., the population, the land use) change over time [9], can be divided into two aspects: long-term urbanization with sustained economic growth [22], and short-term anthropogenic changes and activity rhythms [2, 24, 48]. We focus on the latter one and our goal is reveal urban dynamics in terms of human activities with different types and intensity. Relevant to our work, Zhang *et al.* [48] demonstrated that the activity volume of an area is not uniformly distributed across time, and different areas have different activity volume temporal distributions cross the geo-tagged social data. Also using the geo-tagged data from Twitter, Sofiane *et al.* [2] built activity time series for different cities and found that close neighborhoods tend to share similar rhythms. Louail *et al.* [20] demonstrated that the city shape and hot-spots change with the course of the day. Fabio *et al.* [24] captured the spatio-temporal activity in a city across multiple temporal resolutions, and visualized different activity levels in different time slots. From the perspective of individuals, Clemente *et al.* [8] revealed different urban lifestyles via sequences of purchases and mobility behaviours by coupling credit card data with mobile phone data. In this work, we regard urban activities as time series and aim to reveal the daily regularity hidden in them. Different from the existing works utilizing only offline activities, we model urban dynamics by learning online and offline activities together. Besides, different from the works based on statistical analysis [2] and data visualization [24], we propose a specific model, which achieves urban dynamics understanding and prediction at the same time.

Urban functions inferring. Traditional approaches to infer the actual land-use often rely on costly human surveys, yet it is still coarse-grained and limited in geographical scope. Recently, with the availability of massive data from different sources, efficient methods to infer urban functions have been proposed. Pijiaonowski *et al.* [28] utilized GIS data to model and predict the change of land-use in the city, while Lenormand *et al.* [15] applied a functional network approach to determine land use patterns from mobile phone records. Louail *et al.* [19] utilized origin-destination (OD) matrices to capture the structure of cities and showed that cities essentially differ by their proportion of two types of flows: integrated (between residential and employment hotspots) and random flows. Yuan *et al.* [43] compared urban function discovering results only based on static Points of Interest (POIs) data and based on both POIs and mobility pattern of taxicabs, which proved that mobility reflecting the commute has a strong relationship with urban functions. Besides, Wang *et al.* [34] and Zhang *et al.* [45] showed that the temporal traffic pattern and the activities recorded by check-ins also can be utilized to infer urban functions. Our

work aims to infer the functions by the discovered dynamic patterns via a large-scale mobile cellular network accessing dataset. Moreover, unlike previous works to identify fine-grained yet static urban functions, we pay more attention to the composition of a region's functions with its dynamic changes.

Hidden Markov model and its application. Hidden Markov Model (HMM) is a statistical model in which the time series being modeled is assumed to be a Markov process with unobserved (i.e., hidden) states[29]. In order to improve the performance of HMM, parameter sharing are very helpful to deal with increasingly complex tasks [12]. One well-known example is *shared-distribution HMM*, where clustering is carried out at the distribution level and output distributions are shared with each other if they exhibit acoustic similarity [14]. Another example is *tied-mixture HMM*, which is a kind of semi-continuous HMM [3, 13, 17]. It assumes that each output is generated by large amount of continuous probability density functions (PDFs), while the weight of the PDFs are discrete. By enforcing PDF sharing, it is able to improve the modeling accuracy as its fully using of data. Our proposed *state-sharing HMM* is also a kind of parameters sharing HMM. Unlike *shared-distribution HMM*, our model can be learned end-to-end, which means no following clustering is need to force the parameters shared. Also unlike *tied-mixture HMM*, our model is designed to share a set of states which generates the continuous observations directly instead of only sharing underlying PDFs.

HMM and its variants are also widely used in human activity modeling. One important application is individual mobility prediction [23, 47, 52]. These works assumes that a user's movement is actual the successive transition on several hidden states. These states are the key places he visits frequently (e.g, his home, his office, etc.). The observed spatial points are distributed around the key locations. Therefore, HMM is suitable to first determine the key states under the trajectory, and than predict the next location the user would visit based on the state transition probability. Compared with the existing works, the novelty of our study lies applying HMM in urban dynamics revealing problem. We regard urban dynamics as the transition between hidden states which characterize human activities with different types and intensity, and we also achieve urban dynamics prediction.

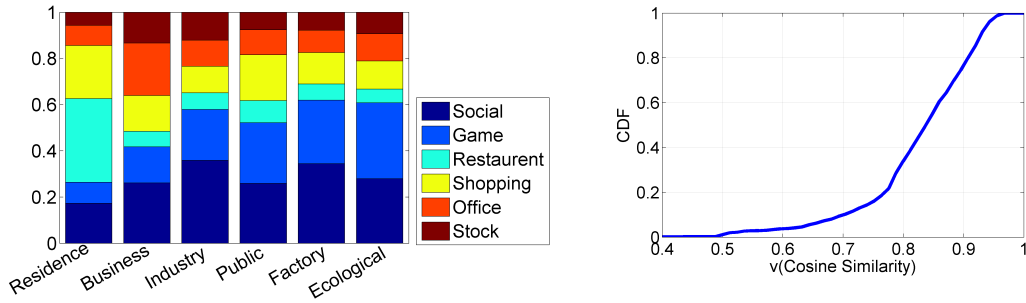
3 PRELIMINARIES

3.1 Motivations

In order to reveal urban dynamics in terms of daily activity rhythms, we investigate the variations of the aggregated online and offline activities along with time. In this section, we discuss the motivations to select mobility as offline activity feature and App usage as online activity feature.

Human mobility, reflecting daily life pattern by its distinct modes, is the most important social-economic activity in the physical space [38]. Consequently, the aggregated mobility behaviour, i.e., how many people leave from, arrive at and stay in each urban region and time slot, reflects urban commuting patterns, indicating that it should be used as offline feature for urban dynamics modeling [42].

While in terms of online activities, it is recently reported that more than 80% of the mobile phone using time in all markets are spent in Apps [1], which demonstrates the purpose of accessing the Internet could be reflected by the App usage. Moreover, the App usage varies significantly with the regions of different land-use, which indicates it has a strong correlation with urban functions [41]. Utilizing the mobile cellular network accessing dataset, we also explore the correlation between App usage and land-use explicitly. As Figure 1(a) shows, App usage in the regions with different land-use are obviously different, which reflects people's different App preferences in different places. We also show the cumulative distribution function of the statistical correlation between land-use components and App usage in Figure 1(b). The results show that for more than 80% of the regions, they are strongly correlated (above 0.76). Therefore, we utilize App usage as online behaviour representation to model urban dynamics and infer urban functions.



(a) For different types of regions (i.e. regions with relatively higher frequency of certain types of land-use), we calculate the percentage of different App categories used. (b) Cumulative distribution function of the statistical correlation between App usage and the land-use.

Fig. 1. Intuitive and statistic correlation between App usage and land-use.

3.2 State-sharing HMM

Regarding the App usage and mobility features in different time slots as time series, we are able to apply HMM to reveal urban dynamics. To deal with the problem of data sparsity and to detect similar as well as different dynamics of urban regions, we propose a state-sharing HMM. Before formally defining the model, we give an exemplified illustration in Figure 2.

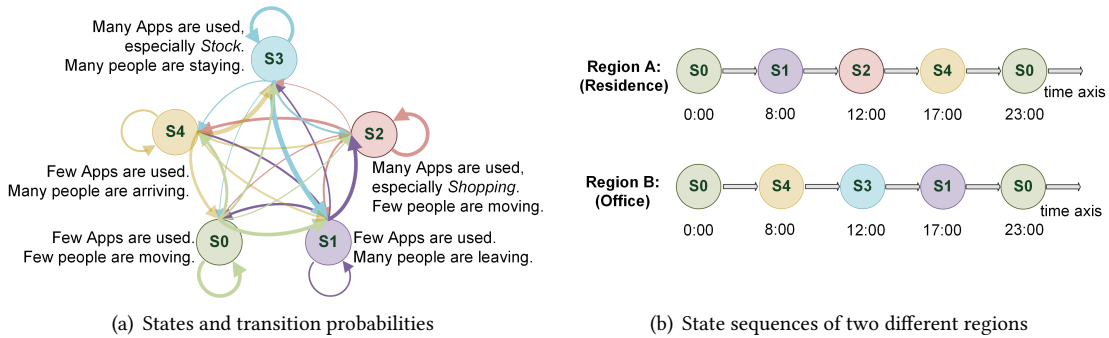


Fig. 2. An exemplified illustration of the state-sharing HMM, where five states with different semantics are shown in five circles of different colors, and their transition probabilities are represented by the colorful arrows. As shown in (a), the thicker the line is, the greater the probability is. Two state sequences of different regions in a working day are shown in (b), where A is a residence region and B is an office region.

In this example, **S0**, **S1**, **S2**, **S3** and **S4** are five hidden states characterizing human online and offline activities with different types and intensity. **S0** denotes the state with low activity intensity both in App usage and mobility. **S2** denotes the state that few people are moving and relaxing Apps are used most frequently, while **S3** denotes the state that few people are moving and official Apps are used most frequently. Both with few Apps used, **S1** denotes a sudden moving-out crowd flow yet **S4** denotes a sudden moving-in crowd flow. Besides, **S0**, **S2** and **S3**

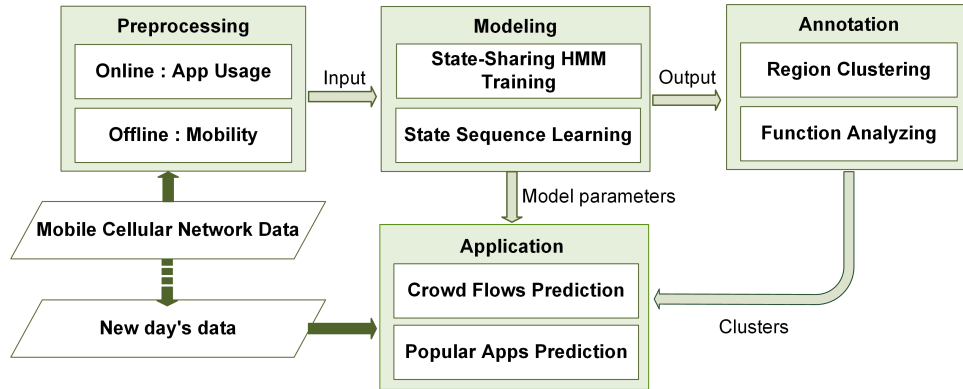


Fig. 3. Framework of the proposed system for urban dynamics modeling.

have larger probability of turning to themselves. Thus, they could continue to appear in a long period before jumping to another state. Compared with **S2** and **S3**, **S1** and **S4** have smaller probability of turning to themselves. Thus, they are more likely to appear in a rather short period. For each urban region, these states appear in turn, which compose its state sequence. There are two different state sequences based on the common set of states and the same transition probabilities as shown in Figure 2(b). The sequence of region A, a residential region, shows the dynamics that few people are active at night, many people leave in the morning, *shopping* Apps are used most frequently during the day, and many people go back in the afternoon. On the contrary, the sequence of region B, which is an official region, shows the dynamics that many people arrive in the morning and leave in the afternoon, and use *stock* Apps instead of relaxing ones frequently during the day. To summary, states are shared but state sequence is unique for each region, which represents the different urban dynamics. The goal of this paper is to discover the shared state set for the city, and own states state sequence for each region.

3.3 Problem Statement and Solution Overview

Based on the ideas discussed above, we formally define the urban dynamics modeling problem. Given the online and offline activity observations of different regions in the city, we aim to address the following three aspects regarding urban dynamics: First is to discover the hidden states in the city, which characterizes human activities with different types and intensity. Second is to detect urban dynamic patterns, which is revealed by the hidden state sequence and could be mapped to urban functions. Third is to predict urban dynamics. Specifically, it is to predict the crowd flows and the popular Apps for each region in the next time slot.

We show the framework of our proposed system in Figure 3, which includes four modules named *Preprocessing*, *Modeling*, *Annotation* and *Application*. Specifically, we first extract App usage and mobility as the input features for each region in different time slots from the mobile network accessing records. Then, we fed them into the state-sharing HMM to discover the city states. Based on the HMM, we learn the state sequence of each region, which reveals the urban dynamics and functions. Besides, applications of predicting the crowd flows and popular Apps can be achieved based on the model.

4 METHOD

4.1 Preprocessing

Mobile cellular network accessing data records when and under which base station a user is using which Apps. To detect the dynamics of urban regions, which are segmented by the major road networks, the first step is to

extract aggregated App usage and mobility features from individual records. For convenience, we discretize one day into several time slots. To distinguish them in the working and non-working day, we use index $[1, n]$ for the former and $[n + 1, 2n]$ for the latter. After mapping the base station to the region and change the timestamp to the time slot, the i -th record is denoted by $\langle u_i, \tau_i, g_i, a_i \rangle$, where u_i is the user ID, τ_i is the time slot and g_i is the region ID. In the following parts, we denote the set of users as $U = \{u_1, u_2, \dots, u_M\} (1 \leq m \leq M)$, the set of regions as $G = \{g_1, g_2, \dots, g_R\} (1 \leq r \leq R)$, and the set of time slots as $T = \{\tau_1, \tau_2, \dots, \tau_N\} (1 \leq n \leq N)$, respectively.

4.1.1 Preprocessing for App Usage. Given a urban region, we define the App usage as how many people are using a certain kind of Apps during one time slot. For the whole city, it is denoted as a 3-dimensional tensor $X \in \mathbb{R}^{R \times N \times L_a}$, where R is the number of regions, L_a is number of App categories, and N is the number of time slots. Since different Apps in the same category play similar roles in reflecting human activities, we only use the App category to characterize the type of human online activities. To share the states in the city, we normalize the App usage to eliminate the influence of the population in different regions. We first compute the *TF-IDF* weights for the usage in each time slot. *TF-IDF* [27], short for term frequency-inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus. Here, we use the *TF-IDF* weight to indicate how popular an App category is in the given time slot. The *TF-IDF* weight for the n -th time slot denoted by X_n can be calculated as follows,

$$X_n(r, l) = \frac{X(r, n, l)}{\sum X(r, n, :)} \times \log \frac{\sum X(:, n, :)}{\sum X(:, n, l)}. \quad (1)$$

Then, for each region we conduct the max normalization on the weights over different time slots as follows,

$$\hat{X}(r, n, l) = X_n(r, l) / \max_{1 \leq n \leq N, l \leq L} (X_n(r, l)), \quad (2)$$

where \hat{X} denotes the normalized App usage tensor.

4.1.2 Preprocessing for Mobility. A user's trajectory is defined as the base station sequence he has visited in time order. By traversing the trajectories of all users, we can obtain the mobility features that how many people leave from, arrive at and stay in each region in different time slots [49]. We denote the mobility features as L_r , A_r and S_r , where $L_r = (L_{r,1}, L_{r,2}, \dots, L_{r,N})$, $A_r = (A_{r,1}, A_{r,2}, \dots, A_{r,N})$ and $S_r = (S_{r,1}, S_{r,2}, \dots, S_{r,N}) (1 \leq n \leq N)$ with $L_{r,n}$, $A_{r,n}$ and $S_{r,n}$ denoting the number of people who **leave from**, **arrive at** and **stay in** the r -th region in n -th time slot, respectively. Similar with App usage, considering the difference for land area and population, we normalized these three vectors by their maximum value. The normalized mobility features denoted by \hat{L}_r , \hat{A}_r and \hat{S}_r can be obtained as follows:

$$\hat{L}_{r,n} = L_{r,n} / \max_{1 < n < N} (L_{r,n}), \forall n = 1, 2, \dots, N, \forall r = 1, 2, \dots, R, \quad (3)$$

$$\hat{A}_{r,n} = A_{r,n} / \max_{1 < n < N} (A_{r,n}), \forall n = 1, 2, \dots, N, \forall r = 1, 2, \dots, R, \quad (4)$$

$$\hat{S}_{r,n} = S_{r,n} / \max_{1 < n < N} (S_{r,n}), \forall n = 1, 2, \dots, N, \forall r = 1, 2, \dots, R. \quad (5)$$

4.2 Modeling

4.2.1 State-sharing Hidden Markov Model. For r -th region, we define the observation sequence from the online and offline activities as $O_r = O_{r,1}O_{r,2}\dots O_{r,n}\dots O_{r,N}$, where $O_{r,n}$ denotes the observation in n -th time slot. It is a multi-dimensional feature vector denoted by $O_{r,n} = \langle o_{r,n,1}, o_{r,n,2}, \dots, o_{r,n,l}, \dots, o_{r,n,L} \rangle (L = L_a + 3)$ with $o_{r,n,l}$

presenting observation for App usage from \hat{X} as (1)-(2) and mobility as (3)-(5) as follows:

$$o_{r,n,l} = \begin{cases} \hat{X}(r, n, l), & 1 \leq l \leq L_a, \\ \hat{A}_{r,n}, & l = L - 2, \\ \hat{L}_{r,n}, & l = L - 1, \\ \hat{S}_{r,n}, & l = L. \end{cases} \quad (6)$$

Observation sequence set of the city denoted by $O = \{O_1, O_2, \dots, O_R\}$ is the input of our model, and our goal is to learn the corresponding hidden state sequence under them. We define the common state set including K hidden states as $S = [s_1, s_2, \dots, s_K]$. We assume that all the input observations O are generated by these states. As a general practice for HMM, the observation variables $o_{n,l}$, ($l = 1, 2, \dots, L$) are independent and $o_{n,l}$ is generated from Gaussian, i.e., $p(o_{n,l}|s_n = k) = N(\mu_{k,l}, \sigma_{k,l})$, where $\mu_{k,l}$ and $\sigma_{k,l}$ are the mean and variance for hidden state k . Therefore, we build the state-sharing HMM parameterized by $\theta = \{\pi, A, \mu, \sigma\}$, where:

- (1) $\pi \in \mathbb{R}^{K \times 1}$ denotes the initial distribution over K hidden states, i.e., $\pi_k = p(s_1 = k)$ ($1 \leq k \leq K$);
- (2) $A \in \mathbb{R}^{K \times K}$ denotes the transition probabilities among the K hidden states. If $(n - 1)$ -th state is $s_{n-1} = j$, then the probability for n -th state s_n to be k is given by $A_{j,k}$, i.e., $p(s_n = k | s_{n-1} = j) = A_{j,k}$;
- (3) $\mu, \sigma \in \mathbb{R}^{K \times L}$ denotes the mean and variance of observation probability, i.e., $p(o_{r,n,l} | s_n = k) = \frac{1}{\sqrt{2\pi\sigma_{k,l}}} \exp\left(-\frac{(o_{r,n,l} - \mu_{k,l})^2}{2\sigma_{k,l}}\right)$.

4.2.2 Model Training. The model parameters $\theta = \{\pi, A, \mu, \sigma\}$ can be learned by maximizing the likelihood function, i.e., the probability of the observations represented by human activities for all regions under different hidden states. According to Baum-Welch algorithm [29], the likelihood function can be maximized by optimize the Q-function step by step. Therefore, the key of the parameter inference is to identify the Q-function denoted by:

$$Q(\theta, \theta^t) = \sum_S p(S|O; \theta^t) \ln p(O, S|\theta), \quad (7)$$

which is different from a general HMM. In our model, we need to maximize the likelihood function of all regions at the same time. Thus, Q function can be defined as follows,

$$Q(\theta, \theta^t) = \sum_{r=1}^R \sum_S p(S|O_r; \theta^t) \ln p(O_r, S|\theta). \quad (8)$$

We train the model using Expectation-Maximization (EM) algorithm [4], which is described in detail in Appendix I. Starting with random initialized parameters, EM algorithm alternates between E-step and M-step round by round until the stop condition is satisfied. In the $(t + 1)$ -th round E-step, it computes $Q(\theta, \theta^t) = \sum_{r=1}^R \sum_S p(S|O_r; \theta^t) \ln p(O_r, S|\theta)$. In M-step, it finds a new estimation $\theta^{(t+1)}$ to maximizes Q-function. After several rounds of iteration, the model tends to converge with the output of parameters $\theta = \{\pi, A, \mu, \sigma\}$.

4.2.3 State Sequence Learning and Clustering. After model training, we obtain K hidden states parameterized by $\{\mu, \sigma\}$ and the transition probability parameterized by $\theta = \{\pi, A\}$. Therefore, we can learn the hidden state sequence under each observation sequence O_r by Viterbi algorithm [33]. As discussed before, the common states set characterizes the fundamental life modes in the city, which reveals the similarity among different regions, while the hidden state sequence of each region indicates the uniqueness of its own dynamics.

In order to detect typical dynamic patterns, we further cluster urban regions according to their state sequences. To divide the regions with similar dynamics into the same cluster, we measure their distance by Hamming distance of their state sequences [10]. Specifically, we define the distance of each two state sequences as the number of different states they have in the corresponding time slots. For example, as shown in Figure 2, states of these two region in 2-th, 3-th and 4-th time slot are different, thus the distance is 3. Because states are discrete

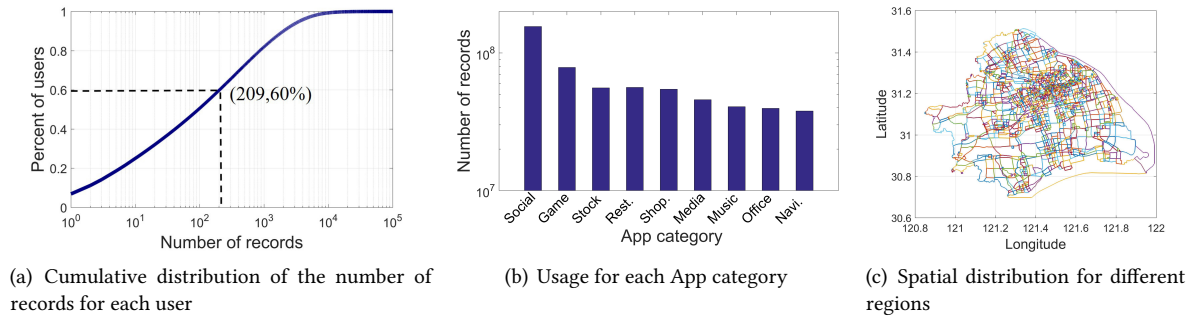


Fig. 4. Data overview.

and cannot be averaged, we apply K-medoids algorithm for clustering [38]. K-medoids defines the intra-cluster distance as the average of other regions to the center. Unlike K-means, it selects one region from all with the minimized intra-cluster distance as the cluster center in each round. To determine how many patterns are suitable, we adopt *Davies-Bouldin index* (DBI) [6] to determine the number of clusters, which reflects the ratio between inter-cluster distance and intra-cluster distance. A smaller DBI usually indicates a more effective clustering. Since people attend different activities in the places with specific urban functions, the discovered states and dynamic patterns are closely related to urban functions. Therefore, we can map dynamic patterns into urban functions.

5 EXPERIMENTS

In this section, we set up a series of experiments to evaluate our proposed model. We first introduce the used datasets and the parameters. Then we discuss the learned dynamics and inferred functions in details. After that, we evaluate the the performance of our proposed application.

5.1 Datasets

1) Mobile cellular network accessing data. We use this dataset to extract human online and offline activity features. This dataset containing mobile users' accessing logs to the cellular network is collected by collaborating with a major mobile network operator in Shanghai, China from April 20th to 26th, 2016. Through deep packet inspection, each access record is characterized by an anonymized user ID, timestamp, cellular base station with GPS location and the metadata of the networking communication. With the geographical information of cellular base stations, we can recover users' trajectories and further to extract offline mobility features. Beside, we can also identify Apps from the networking metadata by adopting SAMPLES [40] to extract online App usage.

Overall, the dataset contains over 1,700,000 unique devices, 2000 unique Apps, and 9800 base stations in Shanghai. We divide these Apps into 9 categories by referring to the App Store (iOS Apps) and Google Play (Android Apps), which includes *social*, *music*, *reading*, *game*, *shopping*, *restaurant*, *navigation*, *office*, and *stock*. The cumulative distribution of the number of records for each user is shown in Figure 4(a). It shows that more than 40% of the users are highly active with than more 200 records. The frequency of mobile users accessing different App categories is shown in Figure 4(b), where App category is arranged in descending order of using frequency. This figure indicates that all kinds of App are used for more than 10 million times in one week and the most popular category is *social*. This large-scale and fine-grained dataset enables the representative of our study.

2) Road network. The road network data is used to obtain the boundaries of urban regions. Urban regions segmented by roads are the basic geographic units of residents' daily life and we use them to reveal urban

dynamics. We crawled the road network from the online map service. Using the major roads of high way and ring roads, Shanghai is divided into 1595 non-overlapping regions spontaneously. Their distribution is shown in Figure 4(c). From this figure we can observe that the regions in downtown are fine-grained and those around suburb are coarse-grained.

3) Check-in dataset. We utilize social-media check-ins in different time slots to validate the uncovered urban dynamics. Check-in refers that a user records his activity with the location information using smartphones [51]. We collected this dataset from one of the most popular location-based service provider in China. Each check-in record consists of anonymized user ID, check-in time, check-in location and point of interest (POI). Totally, it contains over 130,000 records appearing in Shanghai and covers the POI categories including *entertainment, hotel, shopping, leisure, fitness, school, tourism, transportation, finance, company, business, factory, industry, technology park, economic development zone, high-tech development zone, residence, life service, township* and *village*.

4) Land-use map. We use this dataset to validate the inferred urban functions. We collected the latest urban land-use planning map from the government, which contains six land-use types including *residence, business, industry, public infrastructure, farming and forestry*, and *ecological restoration*.

Ethics. We have carefully considered the ethical issues of the data and taken effective measures to protect users' privacy by referring to previous related works which also utilized massive individual data [32, 41]. **First**, when the users use the network service provided by the cellular network operator or social-media platform, they have authorized that their data can be collected and analyzed by the provider. Moreover, all the personal identification information have been stripped (or replaced by a random string) by the provider and we never had the direct access to the actual user ID. **Second**, the interpretation of urban dynamics and functions is of great significance to urban planning. These datasets are used for the public problem solving benefiting the city management, not for the commercial purposes or personal purposes. **Third**, in our study, all the data has been aggregated into different regions, which does not contain any user's preference, indicating that individual privacy is well protected. **Last but not least**, the raw individual data is stored in the provider's servers, which is accessed only by the employees. We only utilize aggregated pre-processed results to carry out this study. Our research has been reviewed and approved by both the provider and local university institutional board.

5.2 Experiments Setting

In order to determine the time granularity of our model, we compare the activity observations when using different length of time slot (i.e., 10 minutes, 20 minutes, 30 minutes, 1 hour, 2 hours, 3 hours, 4 hours). When the length of time slot is shorter than 1 minutes, the observations of neighbor time slots are very similar. But when the length of time slot is longer than 3 hours, the dynamics that the observation sequence reflects are too coarse-grained. Thus, we show the parts of the observations of 30 minutes and 2 hours in Figure 5(a), (b), (d) and (e). Compare Figure 5(a) and (c), we can observe that the changing trend with time of App usage for different time granularity is consistent. Compare Figure 5(b) and (d), we can observe that peak and valley of mobility for different time granularity also appear in the same time. Since the short one time slot is, the larger the number of time slots is. Therefore, we set the length of the time slot as 2 hours as a trade-off between time granularity and model complexity.

Then, to decide how many hidden states are suitable to present the characters of whole city, we analyze the principal components of the mean vectors $\{\mu_1, \mu_2, \dots, \mu_k, \dots, \mu_{20}\}$ of the Gaussian distribution and the probability of observation $P(O|\theta) = \prod_{r=1}^R p(O_r|\theta)$ for different number of hidden states. The cumulative variance ratio (CVR) and $P(O|\theta)$ with different number of states are shown in Figure 5(c) in red and blue, respectively. From the results we can observe that when the number of hidden states is more 10, both the two curves converge to their maximum. Thus, we set $K = 10$ to learn states as different as possible at the lowest training cost.

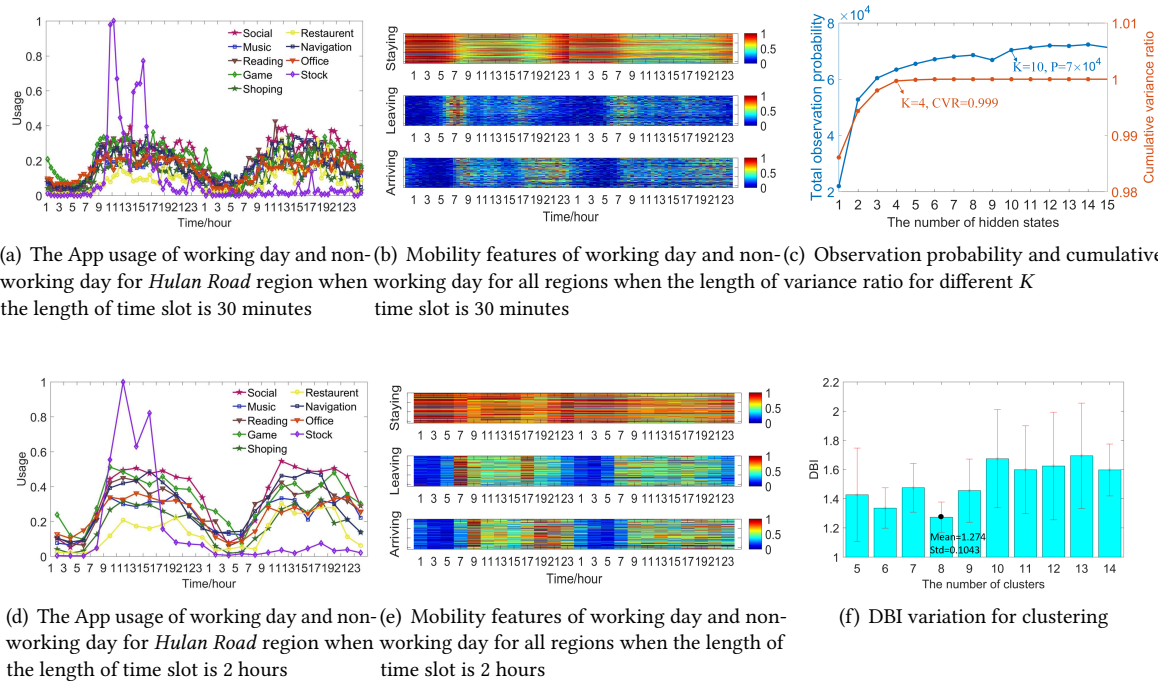


Fig. 5. Parameters setting.

Finally, we use DBI to determine the number of dynamic patterns. For different cluster numbers, we conduct the experiments for 10 times and show the mean and standard deviation of DBI for cluster number in Figure 5(f). When the number of clusters is 8, both the minimum mean and standard deviation are achieved, which means 8 patterns are mostly stable and suitable.

To conclude, we use six days' mobile cellular network accessing data for model learning and leave the rest day for prediction. There are 1596 regions, each with an activity observation sequence of length 24 and dimension 12. Based on these observations, we learn 10 hidden states and detect 8 typical dynamic patterns.

5.3 Results.

5.3.1 States Interpretation. We have learned 10 hidden states as described in Figure 6, where the depth of the color indicates the activity intensity, and the frequency of different App used reflect the activity type. From the results we can observe that **S4** with the lightest color is the state indicating that people all fall asleep. Similarly, we can infer that **S7** is the state people go to work from home and **S2** is the state that people go home from work. **S0**, **S1** and **S3** correspond to working state with *social*, *reading* and *stock* Apps frequently used, while **S5**, **S6** and **S8** correspond to leisure state with *social*, *game* and *navigation* Apps frequently used. The remainders **S9** is a relatively inactive state with the lightest color for all Apps. In this state, few people are moving and use Apps, which indicates people are having a rest at home.

5.3.2 Dynamic Patterns. We obtain 8 typical dynamic patterns by clustering the learnt state sequences. The state sequences of each pattern are shown in Figure 7(a) and (c), 8(a), (c), and (e), and 9(a),(c) and (e), where each row

S0	S1	S2	S3	S4	S5	S6	S7	S8	S9	Transition probabilities
<i>Social</i>	<i>Stock</i>	<i>Social</i>	<i>Social</i>	<i>Social</i>	<i>Social</i>	<i>Social</i>	<i>Social</i>	<i>Social</i>	<i>Social</i>	
<i>Reading</i>	<i>Social</i>	<i>Navigation</i>	<i>Stock</i>	<i>Music</i>	<i>Stock</i>	<i>Reading</i>	<i>Music</i>	<i>Navigation</i>	<i>Music</i>	
<i>Stock</i>	<i>Reading</i>	<i>Music</i>	<i>Navigation</i>	<i>Reading</i>	<i>Music</i>	<i>Game</i>	<i>Reading</i>	<i>Music</i>	<i>Reading</i>	
<i>Game</i>	<i>Game</i>	<i>Reading</i>	<i>Reading</i>	<i>Game</i>	<i>Reading</i>	<i>Shopping</i>	<i>Game</i>	<i>Reading</i>	<i>Game</i>	
<i>Shopping</i>	<i>Shopping</i>	<i>Game</i>	<i>Game</i>	<i>Shopping</i>	<i>Game</i>	<i>Restaurant</i>	<i>Shopping</i>	<i>Game</i>	<i>Shopping</i>	
<i>Navigation</i>	<i>Restaurant</i>	<i>Shopping</i>	<i>Shopping</i>	<i>Restaurant</i>	<i>Shopping</i>	<i>Navigation</i>	<i>Restaurant</i>	<i>Shopping</i>	<i>Restaurant</i>	
<i>Office</i>	<i>Navigation</i>	<i>Restaurant</i>	<i>Restaurant</i>	<i>Navigation</i>	<i>Restaurant</i>	<i>Office</i>	<i>Navigation</i>	<i>Restaurant</i>	<i>Navigation</i>	
<i>Music</i>	<i>Office</i>	<i>Office</i>	<i>Stock</i>	<i>Office</i>	<i>Navigation</i>	<i>Music</i>	<i>Office</i>	<i>Office</i>	<i>Office</i>	
<i>Restaurant</i>	<i>Music</i>	<i>Stock</i>	<i>Office</i>	<i>Stock</i>	<i>Office</i>	<i>Stock</i>	<i>Stock</i>	<i>Stock</i>	<i>Stock</i>	
<i>Staying</i>	<i>Staying</i>	<i>Staying</i>	<i>Staying</i>	<i>Staying</i>	<i>Staying</i>	<i>Staying</i>	<i>Staying</i>	<i>Staying</i>	<i>Staying</i>	
<i>Leaving</i>	<i>Arriving</i>	<i>Arriving</i>	<i>Arriving</i>	<i>Arriving</i>	<i>Arriving</i>	<i>Arriving</i>	<i>Leaving</i>	<i>Leaving</i>	<i>Leaving</i>	
<i>Arriving</i>	<i>Leaving</i>	<i>Leaving</i>	<i>Leaving</i>	<i>Leaving</i>	<i>Leaving</i>	<i>Leaving</i>	<i>Arriving</i>	<i>Arriving</i>	<i>Arriving</i>	

Fig. 6. The Gaussian mean of emission probability for each hidden state and the transition probabilities among these hidden states we obtain. Each state is shown in a column, 9 categories of Apps and 3 kinds of mobility are ranked in descending order by the mean value respectively, where the *Bold* denotes a value higher than 0.65, the *Normal* denotes a value higher than 0.45 but no more than 0.65, and the *Light* denotes a value lower than 0.45.

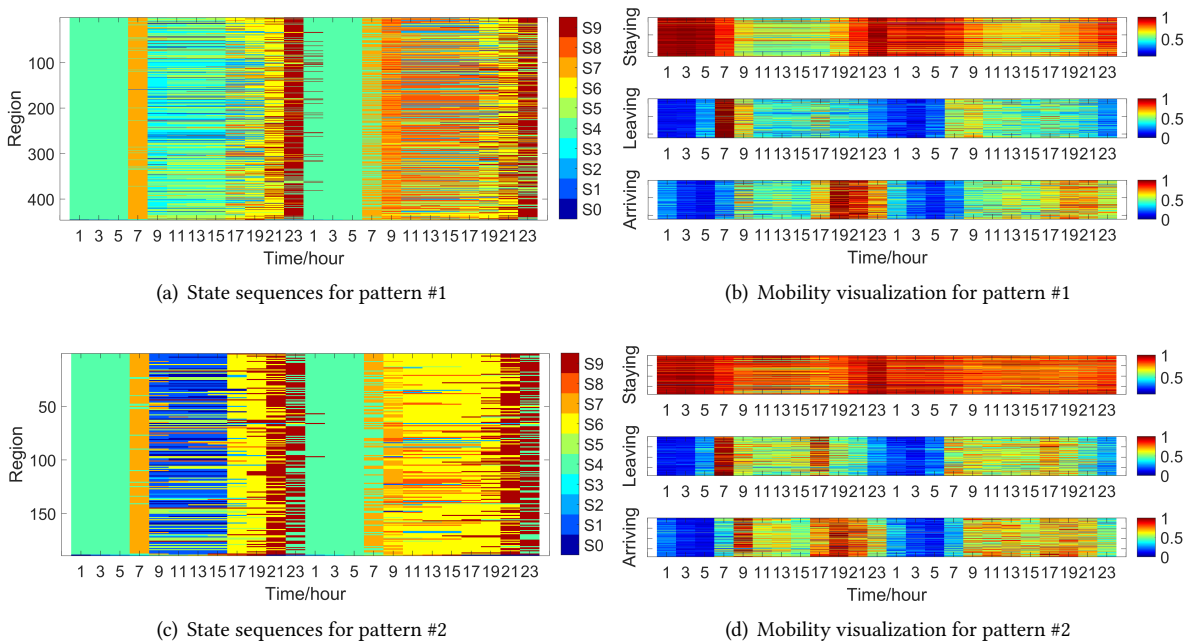


Fig. 7. Visualization of state sequences and mobility for the **residential**, where the state sequence of each region is shown in a row on the left and the normalized mobility of each region is shown in a row on the right.

presents one region’s dynamics from the working day to the non-working day. The corresponding normalized mobility \hat{S} (Staying), \hat{L} (Leaving) and \hat{A} (Arriving) are shown in Figure 7(b) and (d), 8(b), (d) and (f), and 9(b), (d) and (f), respectively. We divide these 8 dynamic patterns into three classes to discuss their functions as follows:

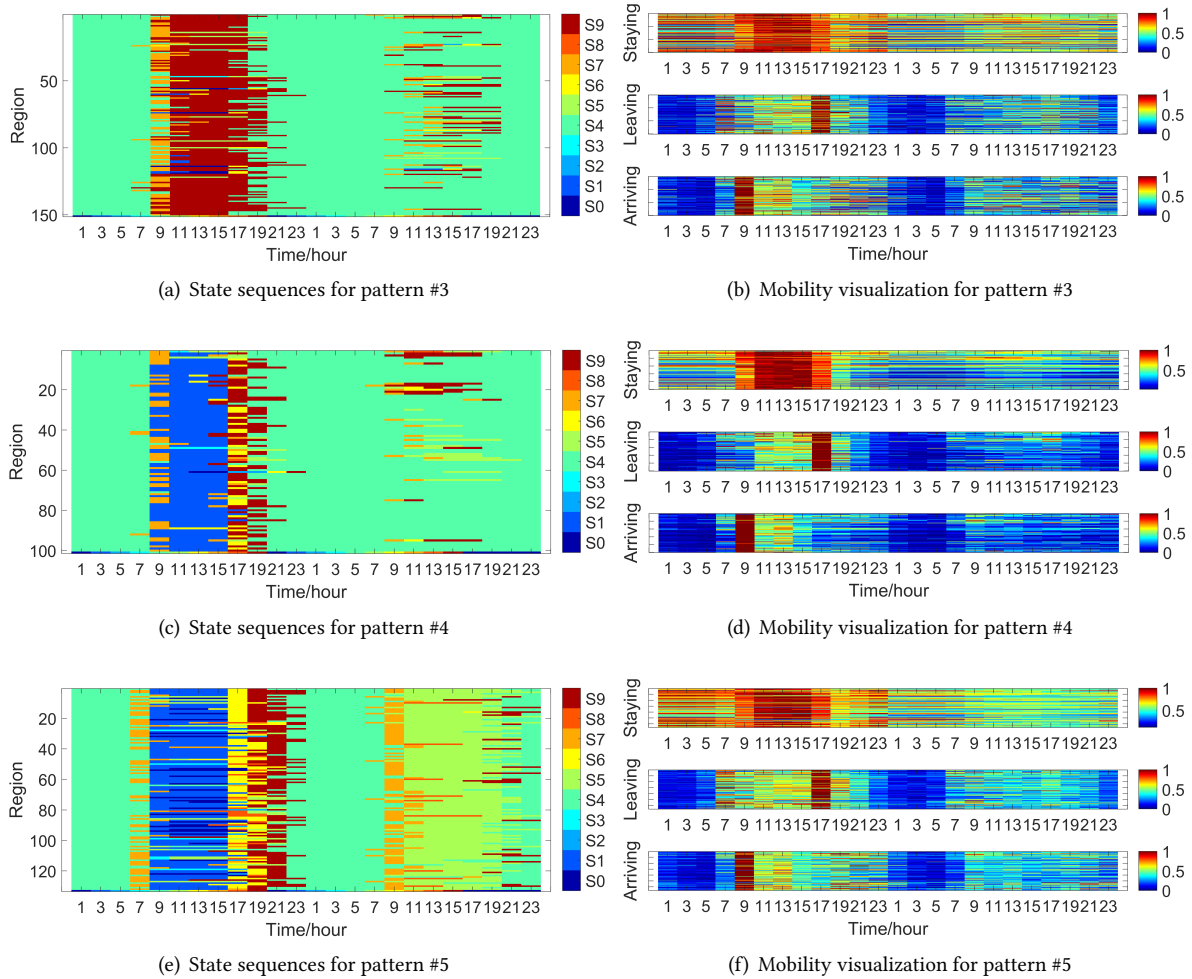


Fig. 8. Visualization of state sequences and mobility for the **official**, where the state sequence of each region is shown in a row on the left and the normalized mobility of each region is shown in a row on the right.

Residential regions: the first class shown in Figure 7 includes pattern #1 and #2. The common feature is that **S7** (in orange color) consistently appears at around 7am of the working day, which means people go to work from home at that time. From the three mobility features visualized in Figure 7(b) and (d), we can observe that lots of people stay here before 7am, after 7pm of the working day, and in the whole non-working day. Besides, there are both a leaving peak (in dark-red color) at around 7am and an arriving peak at around 7pm of the working day when people go to work and back to home, respectively. These characters give more indicators for the residential region. The difference between these two patterns is also obvious. **S0** (in dark-blue color) and **S1** (in navy-blue color) appear more frequently in pattern #2 than #1, which represents that a small group of people come to regions in pattern #2 and use *stock* Apps in the working day. From the mobility of pattern #2, we can also observe that there are an arriving peak at around 7am and a leaving peak at around 7pm. Thus, compared with pattern

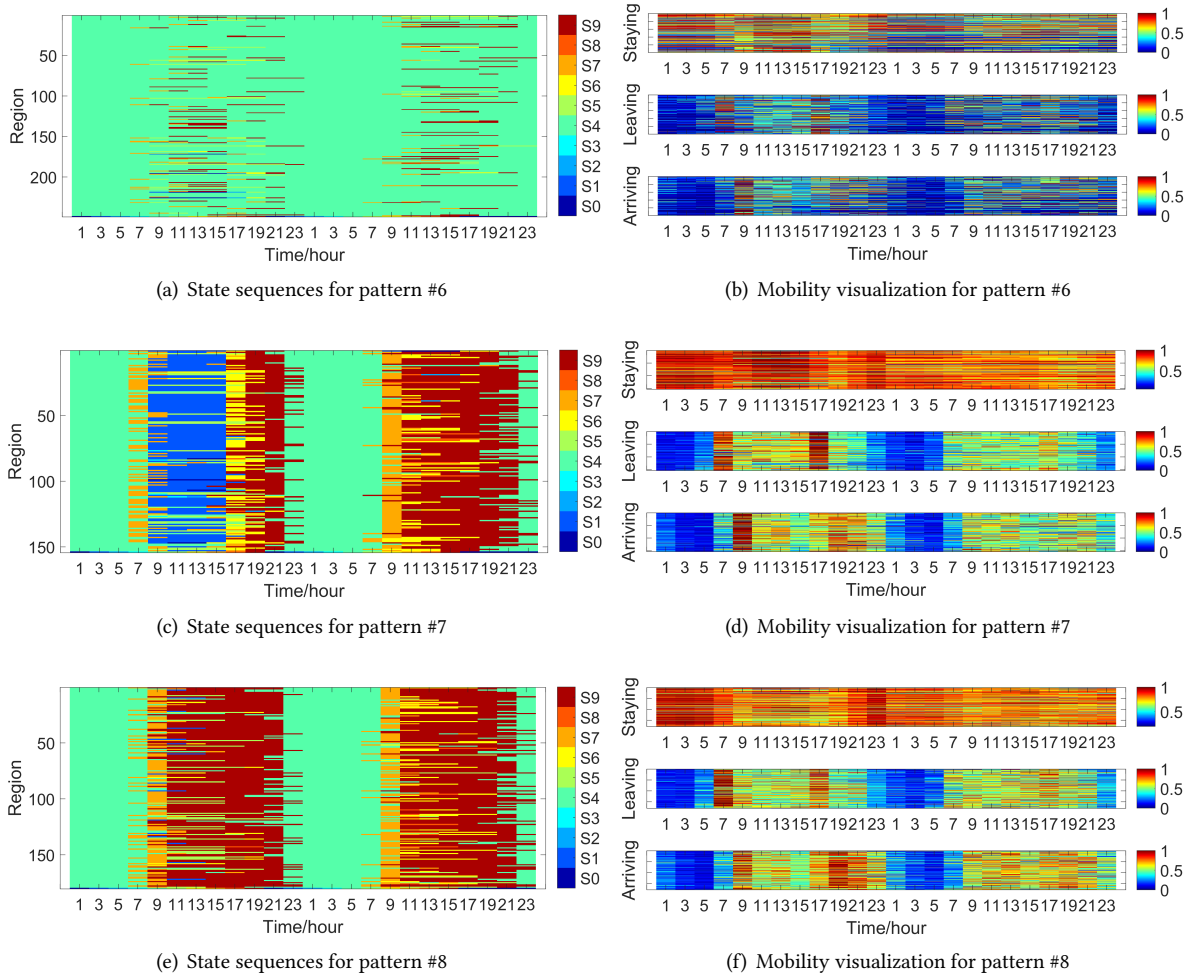


Fig. 9. Visualization of state sequences and mobility for the **compound**, where the state sequence of each region is shown in a row on the left and the normalized mobility of each region is shown in a row on the right.

#1, regions in pattern #2 mainly contain houses but also contain some office areas. **In summary**, pattern #1 corresponds to single residence function, while pattern #2 corresponds mainly to residence but partly to office function.

Official regions: the second class shown in Figure 8 including pattern #3, #4 and #5. The common feature is that **S4** and **S5** (in green-series color) appear frequently during daytime in non-working day, which means few people comes here in the weekend. The arriving peak at around 7am and the leaving peak at 7pm of working day also support the interpretation of official areas. During daytime in weekday, **S9** (in dark-red color) appears frequently in pattern #3, while **S0** (in dark-blue color) and **S1** (in navy-blue color) appear frequently in pattern #4 and #5, representing that people work in the regions of pattern #3 use less Apps especially the *stock* Apps than in the regions of pattern #4 and #5. Comparing pattern #4 and #5, we can observe that **S7** (in orange color) also

appears at around 7 am when people leaving from home. Thus regions in pattern #5 mainly contain office but also contain some residential areas. **In summary**, pattern #3 corresponds to single office function with few Apps used, pattern #4 corresponds to single office function with App used frequently, while pattern #5 corresponds mainly to office but partly to residence function.

Compound regions: the third class shown in Figure 9 including pattern #6, #7 and #8. The transition of states is more complex than the first two classes. For pattern #6, **S7** (in orange color) appears at anytime of the two days. The frequency of using Apps is rather stationary and people do not move intensively in a certain time slot. Therefore, regions in pattern #6 are more likely to be far suburb with low population density. For pattern #7 and #8, people are very active both in weekday and weekend. The ration of staying is consistently high both in day and night, and people move frequently in the weekend. Thus these regions can be compound zones performing different functions in different time slots. Compared with pattern #8, **S0** (in dark-blue color) and **S1** (in navy-blue color) in pattern #7 appear more frequently in working day. Besides, as shown in Figure 9(f), the number of people who move in and move out the regions in pattern #8 is always large at all time slot during the day. Therefore, pattern #7 covers the areas mainly playing the rule as office but also as houses and entertainment sometime. Pattern #8 contains houses, entertainment areas, office as well as transportation hubs. **In summary**, pattern #6 corresponds to suburb function, pattern #7 corresponds mainly to office but partly to residence and entertainment function, while pattern #8 corresponds to highly dynamic and complex functions.

In summary, these state sequences of the two typical days, reveal the regularity of human activities both in terms of type and intensity. Different state sequences in the working time and non-working time also can indicate the functions of a region. Consequently, we discover Pattern #1, #3, #4 and #6 with fundamental functions of living and working, and we also discover Pattern #2, #5, #7 and #8 with compound and dynamic functions.

5.4 Validations

5.4.1 Validation with Check-ins. Check-ins reveal the purpose of user's activity in the city intuitively. Although POIs are static, their popularity in different time slots would be various (e.g. *restaurant* POIs are most popular at noon), and this dynamic popularity can be reflected by checked-in frequency. Therefore, we use check-ins to validate the detected dynamic patterns. Specifically, we explore the correlations between the state sequences and check-in POIs of each pattern. First, we merge the consecutive time slots under the same hidden state according to the center of each pattern. Then, we calculate the percentage of all kinds of POIs checked-in during the merged time slot. The results for three classes as mentioned in Section 5.3.2 are shown in Table 1, 2 and 3, respectively. In these tables, the **Time** row shows the time slot segmentation results, where *W* denotes working day and *N* denotes non-working day, and the number after *W* or *N* represents the hour of the day (e.g., *W0-W5* means from working day 0:00 to working day 5:59). The rest rows show the hidden state, the most frequently checked-in POI category, and the checked-in percentage in the corresponding time slots, respectively. The observations from these tables are as follows:

(1) From Table 1, we can observe that *residence* POIs are more popular in the sleeping hour (*W0-W5, N0-N5*) with the percentage of 37.8% in pattern #1 and 29.9% in pattenr #2, while *transportation* POIs are more popular during rush hour (*W6-W7, N6-N9*) with the percentage exceeding 25% when people go to work. This indicates the dynamics that people rest in the night and go out in the morning, which is consist with the transition from **S4** to **S7**. Besides, during the daytime, people in pattern #1 are more likely to visit *leisure* POIs, while in pattern #2 *company* and *tourism* POIs are more popular, which further proves that the function of pattern #2 changes from residence to office when the time changes from night to day. This also supports the existence of **S0** in pattern #2.

(2) For the official regions as shown in Table 2, few people are active at night (*W0-W7, N0-N9*), which supports why **S4** always appears at night. *business* and *factory* POIs are distinct in the working hour (*W10-W17, W8-W15*) with the percentage exceeding 20%, which means many people work here during the day as **S1** indicates. It is

Table 1. Check-in POI components in different time slots for the **residential**, where POIs corresponding to their main function are highlighted.

Time	Pattern #1				Pattern #2			
	W0-W5 N0-N5	W6-W7 N6-N9	W8-W15	N10-N17	W0-W5 N0-N5	W6-W7 N6-N9	W8-W15	N8-N19
State	S4	S7	S5	S8	S4	S7	S0	S8
Check-in POI	Residence	Transportation	Leisure	Leisure	Residence	Transportation	Company	Tourism
Percentage	37.8%	25.3%	28.7%	45.5%	29.9%	25.5%	18.4%	41.2%

Table 2. Check-in POI components in different time slots for the **official**, where POIs corresponding to their main function are highlighted, and – denotes there are very few check-in records.

Time	Pattern #3			Pattern #4			Pattern #5		
	W0-W9 N0-N9	W10-W17	N10-N17	W0-W7 N0-N9	W8-W15	N10-N17	W0-W5 N0-N9	W8-W15	N10-N17
State	S4	S9	S4	S4	S1	S4	S4	S1	S5
Check-in POI	-	Factory	Shopping	-	Business	Tourism	Residence	Business	Transportation
Percentage	-	21.3%	33.5%	-	42.7%	25.6%	18.3%	20.1%	23.4%

Table 3. Check-in POI components in different time slots for the **compound**, where POIs corresponding to their main function are highlighted, and – denotes there are very few check-in records.

Time	Pattern #6			Pattern #7			Pattern #8		
	W0-W9 N0-N9	W10-W15 W20-W21	N10-N19	W0-W5 N0-N7	W8-W15	N10-N19	W0-W5 N0-N7	W8-W21	N10-N19
State	S4	S9	S4	S4	S1	S9	S4	S9	S9
Check-in POI	-	Leisure	Transportation	Residence	Business	Tourism	Transportation	Transportation	Tourism
Percentage	-	19.1%	22.7%	15.6%	21.2%	20.1%	52.4%	15.6%	20.1%

worth noting that in pattern #5, the most popular POIs in the night is *residence* while in the day is *office*, which shows that the function changes from residence to office. This is also consistent with our previous analysis for pattern #5.

(3) For the compound regions as shown in Table 3, the most popular POIs are significantly different in different time slots, especially for pattern #7, which presents dynamic functions. Moreover, the percentage of *transportation* POIs is generally high (i.e., more than 15%) in different time slots in pattern #8, which indicates its high volume of crowd flow. Interestingly, *Disney Park* and *Hongqiao Airport* are in this pattern, which all highly support the existence of **S9**.

In summary, we analyze the POIs checked-in in different time slots for different patterns. All these results are consistent with the semantic of learnt states and revealed dynamics.

5.4.2 Validation with Urban Land-use. In order to further verify the inferred functions, we compare them with the urban land-use published by the local government. Figure 10(a) shows the land-use map including 6 types of *residence*, *business*, *industry*, *public infrastructure*, *framing and forestry*, and *ecological restoration*. From the result we can observe that *Business* areas (in red color) such as office buildings and residential areas (in creamy-white color) are concentrated in the city centre, while outside are other functional areas. We also show the spatial distribution of the eight dynamic patterns identified by our model in Figure 10(b), where each pattern is presented

in a unique color. Besides, the corresponding land-use types to our patterns as well as their functions are illustrated in Table 4. From Figure 10 and Table 4, we can observe that the most central working areas labeled by blue dash dot and the major living areas labeled by red dash-dot are well consistent between the land-use and our patterns. Besides, some regions with the large area are found to have the compound function, especially the *Pudong* and *Jinshan* region, which is reasonable since such the land-use of such large regions cannot be single. Functions of each region are not available directly from the land-use plotting published by the government. To quantitatively evaluate the dynamic patterns, we randomly select 100 regions of different patterns and manually label their functions by checking their locations on the map. Finally, the functions of 81% of the selected regions are consistent with our results, which further demonstrates the effectiveness of our revealed dynamics from human activities.

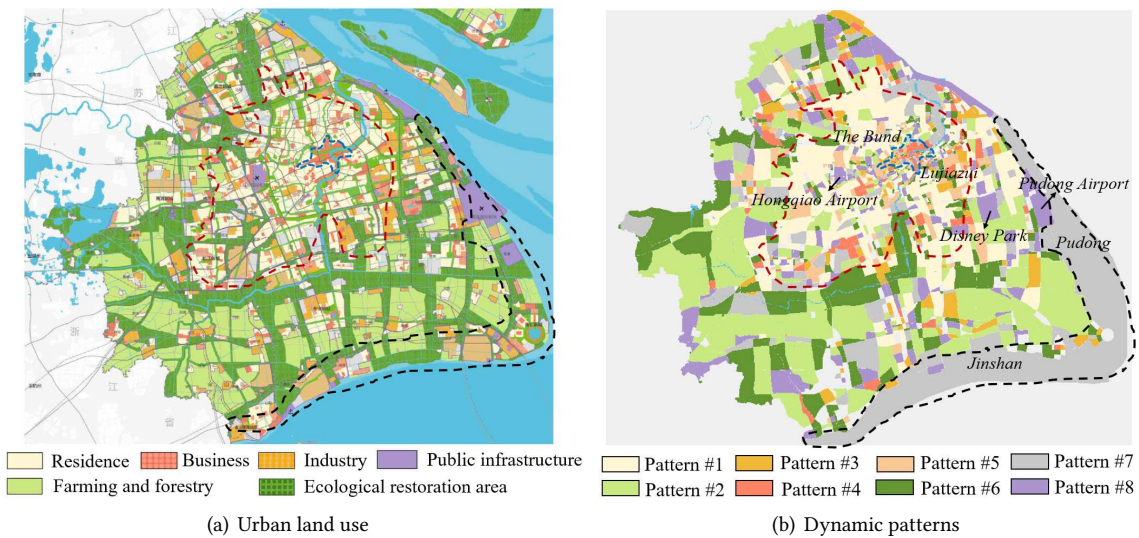


Fig. 10. The spatial distribution for different regions in Shanghai, where within the polygon marked by blue dash dot is downtown *business* area, outside the blue polygon but within the red polygon is downtown *residence* area, and the polygon labeled by black dash dot denotes the most two coarse-grained regions *Pudong* and *Jinshan*.

Table 4. Relationship between land-use and our patterns.

Dynamic Patterns	Pattern #1	Pattern #2	Pattern #3	Pattern #4&5	Pattern #6	Pattern #7&8
Land-use / Region	Residence	Farming	Industry	Business	Ecological	<i>Pudong, Jinshan, etc</i>
Functions	living	living, working	working	working, living	compound	compound

5.5 Application of Prediction

5.5.1 Tasks Design. Our proposed model can learn the patterns of human dynamic activities in different regions in the city, which enable the prediction of people's activities in next time slot. We design two prediction tasks as follows:

Crowd flow prediction: Given the mobility observation of a new day with the length of n time slots, we predict how many people will move in, move out, or stay in $(n + 1)$ -th time slot for each region.

Popular App prediction: Given the App usage observation of a new day with the length of n time slots, we predict which Apps are used frequently in $(n + 1)$ -th time slot for each region.

As mentioned before, six days' data are used for model training and the rest day's data is used for prediction evaluation. It is worth noting that testing data is pre-processed with the same normalization factors and same parameters of the typical working as mentioned in section 3.1. Specially, for r -th region, we define the given observations as $T = T_1, T_2, \dots, T_N$, where $T_r = T_{r,1}T_{r,2}\dots T_{r,n}$ ($R = 1959, 1 \leq n \leq 11, L = 12$). We also define the candidate pool as all the observations in the training data. Then, our goal is to predict $T_{r,n+1}$ for region r at $(n + 1)$ -th, which is determined by:

$$s_{r,n} = \arg \max_{1 \leq j \leq K} \gamma(s_{r,n}^j), \quad (9)$$

$$s_{r,n+1} = \arg \max_{1 \leq k \leq K} A_{j,k}, \quad (10)$$

$$T_{r,n+1} = \arg \max_{r', 1 \leq n' \leq N} p(O_{r',n'} | s_{n+1} = k), \quad (11)$$

where region r' and region r are in the same pattern [30]. We first use the current observations to determine the optimal state sequence of these n time slots, and then to determine the next state by (9) and (10). The history observation with the maximum probability in $(n + 1)$ -th state from the candidate pool is selected as the next observation by (11) [21]. Through anti-normalization, we obtain the number of crowd flows. By sorting the observation of Apps, we obtain the most frequently used Apps.

5.5.2 Metrics and Baselines. To fully evaluate our system, we adopt three evaluation metrics: *TopN-hitrate*, *TopN-accuracy* and *error ratio*, which are defined in Appendix II. The first two metrics are used for popular apps prediction, among which *TopN-hitrate* is the percentage of regions whose TopN Apps are successfully predicted (i.e., correct for at least one), while *TopN-accuracy* reflects how many of the TopN Apps are predicted correctly. The last one is for crowd flows prediction, which is defined as the mean ratio of the difference between the prediction and the ground truth. As a comparison, we also conduct the prediction on our dataset by the following three baselines:

HV : In this method, the prediction of $(n + 1)$ -th time slot is the **Historical Value** (HV) in the training data at the same time slot (i.e., $(n + 1)$ -th time slot in the typical working day) of the same region, directly [49].

ARIMA: **A**uto-**R**egression **I**ntegrated **M**oving **A**verage (ARIMA) is a well-known model for time series analysis and prediction. We conduct the ARIMA prediction with the most suitable orders (i.e., suitable p , d and q) for each region [49].

HMM: In order to show the ability of our model to deal with data sparsity, we train an independent HMM for each region and conduct the prediction based on its own states instead of the shared ones [30].

5.5.3 Results and Analysis. We evaluate the performance of different regions with different length of given observations. First, we partly show the results for pattern #2 (a residential pattern) and #4 (an official pattern) in Figure 11. Over all, the prediction error is always lower than 50% and the accuracy is always higher than 50%. Specifically, from Figure 11(a) and (c), we can observe that the prediction for the number of people staying here is more reliable than that of people moving in or moving out. The error of leaving and arriving prediction are maximum at 11pm for pattern #2, while the staying prediction error are minimum at 11pm for pattern #4. Because there are usually few people in the official zones at night, thus the accuracy of the prediction would decline. Besides, about 5pm there is a rush hour in Friday, thus the prediction of mobility is not as correctly as usual. From Figure 11(b) and (d), we can observe that *Top3-accuracy* and *Top3-hitrate* is stable, where *Top3-accuracy* is always more than 50% and the *Top3-hitrate* is always more than 90% for all patterns in all time slots, which demonstrates our overall prediction accuracy is credible in different time slots.

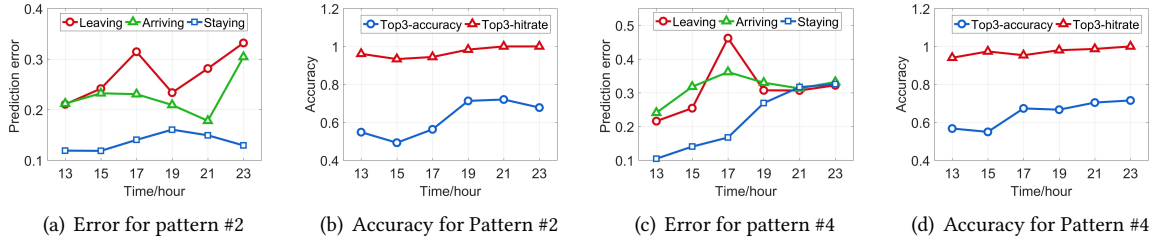


Fig. 11. Crowd flows prediction error and popular Apps prediction accuracy in different time slots.

Table 5. Prediction performance compared with baselines.

Metric	Error for leaving	Error for arriving	Error for staying	Top3-accuracy	Top3-hitrate
HV	85.7%	99.4%	70.4%	51.9%	90.9%
ARIMV	54.0%	47.0%	20.0%	43.9%	79.8%
HMM	54.2%	44.9%	34.7%	42.9%	75.9%
OUR	29.4%	28.3%	19.8%	60.0%	94.6%
Outperform	45.6% (ARIMV)	39.8% (ARIMV)	1.0% (ARIMV)	15.6% (HV)	4.4% (HV)

The complete results are shown in Table 5. Among all the methods, our model always achieves the best performance across different metrics. The total average crowd flow prediction error for leaving, arriving and staying of different patterns are 29.4%, 28.3% and 19.8%, obviously lower than the best baseline for 46.6%, 39.8% and 1.0%, respectively. For popular App prediction, the total average *Top3-accuracy* and *Top3-hitrate* is 60.0% and 94.6%, also obviously higher than the best baseline method for 15.6% and 4.4%. Over all, the prediction error for crowd flows (i.e., the average error of leaving, arriving and staying) is 25.8% and the *Top3* prediction accuracy for App usage is 60.0%, which outperforms the state-of-the-art baseline (i.e., ARIMA for crowd flows prediction and HV for popular Apps prediction) by about 36.1% and 15.7%, respectively. These results do not only demonstrate we have realized an accurate modeling for online App usage and offline mobility, but also indicate that we have achieved reliable urban dynamics prediction.

6 DISCUSSION

Traditional urban dynamics understanding from human surveys is time-consuming, expensive yet coarse-grained [25]. In this paper, we propose a model to efficiently reveal urban dynamics via large-scale mobile accessing dataset. Regarding human activities as time series, we are able to detect typical urban states and different dynamic patterns. We discuss the implications and applications of our work for both researchers and city planners as follows.

6.1 Implications

First of all, we have discovered ten typical states characterizing different life modes in the city. Since we are the first to utilize App usage as online behaviour features to model urban dynamics, the states we detected (e.g., working with App used frequently, working with few App used, having fun through Apps, sleeping with low crowd flow volume, rush hour with high crowd flow volume, etc.) are more semantic-rich and can provide comprehensive understanding for urban dynamics. By comparing these states, we can observe that during

resting time, people keep on using *social* Apps while during working time, the most frequently used Apps are unexpectedly *stock*. We have also found eight dynamic patterns which are composed of these typical states. **Second**, compared with existing data-driven urban dynamics revealing systems [2, 24], our proposed state-sharing HMM can model the aggregated activities in a concise and probabilistic way, which means dynamic prediction is also achieved at the same time. Take pattern #2 as an example, as shown in Figure 11 (a) and (b), our model can predict the volume of arriving, leaving and staying as well as the popular Apps in the next time slot. The prediction error is always less than 40% and the prediction accuracy is always higher than 50%. Based on our prediction, city planners can make better urban management and resource allocation. **Third**, we demonstrate that urban dynamics represented by online and offline behaviours can be used to infer urban functions. Compared with previous urban function discovering works only using offline human mobility data and statistic POIs [43, 44, 53], we detect eight dynamic patterns corresponding to different functions. More importantly, we unexpectedly find that among these dynamic patterns, some indicates single function such as pattern #1, which is simple residence, while the others indicate dynamic and compound functions such as pattern #2 and #5, which play the role as residence and office alternately. All of these are our new contributions and make our findings significant.

6.2 Applications

Our work facilitates urban planning from different aspects. **First and foremost**, it provides more comprehensive understanding for urban dynamics and urban functions. For architects, planners, and urban designers, neighborhood activity patterns from intensive ethnographic surveys that take years to conduct and given the speed at which neighborhoods change, can be out of date quickly. The features provided by our work helps examine the diversity, distribution, and intensity of human activity within a given neighborhood, thus offering insight to the functioning of the entire neighborhood and supporting the government make better land use planning. **Second**, understanding and predicting the pattern of the crowd flow, the offline behaviours we utilize, is of great significance for traffic dispatching, transportation infrastructure construction, etc. For example, according to our analysis, there is usually a leaving rush at 7-9am in the residence regions but an arriving rush at 8-10am in the office regions, which indicates different traffic demand in different regions and different time slots. This also reminds the traffic management department to dynamically dispatch staffs and resources. **Third**, analyzing and predicting the App usage is also necessary for network operator to make better network infrastructure construction and resource allocation. For the regions with Apps used diversely and frequently, more base stations and network resource are needed. Considering the dynamics of the usage of different App categories, a reasonable allocation strategy can improve the utilization of the existing network. In summary, our work sheds light on the fundamental dynamics in the city, which contributes to solving many urban issues.

6.3 Limitations

This study is the first step to utilize the hidden Markov model to describe the urban dynamics. By sharing states among different regions, the difference, as well as similarities, could be learned around the city. Our work has the following limitations. First, the length of our data is one week, and we aggregate it to one working day and one non-working day. Thus, we only exhibit the regular dynamics of working and non-working day. We leave investigating more detailed daily, weekly and monthly dynamics as future work. Second, due to the data limitation, we only evaluate the performance of prediction for the working day. References have demonstrated that the most important time context for population [34, 39] as well as App usage [18, 36] prediction is the hour of the day and the day of the week. Therefore, one week's data is enough for model evaluation [35, 37]. The performance of the baselines could be improved if more data are available. However, it would be limited, and our model can still outperform the baselines as state-sharing strategy can make full use of the data and learn robust results for urban dynamics.

7 CONCLUSION

In this paper, we reveal urban dynamics by learning human online and offline behaviours together via a large-scale dataset of mobile cellular network accessing. We propose a state-sharing HMM system that models urban dynamics and infers functions by extracting human activities of App usage and mobility. The evaluations demonstrate that our system can effectively detect dynamic patterns. Our work opens a new angle to reveal urban dynamics using online and offline behaviours together, and paves the way for extensive applications including urban demand analysis, land use planning, and activity prediction.

APPENDIX I: MODEL TRAINING

The specific process of the state-sharing HMM training and state sequence learning is described as below.

We use the Baum-Welch algorithm to train the model. Starting with random initial parameters, the EM algorithm alternates between E-step and M-step round by round until the stop condition is satisfied. In the $(t + 1)$ -th round E-step, it computes the likelihood function $Q(\theta, \theta^t) = \sum_{r=1}^R \sum_S p(S|O_r; \theta^t) \ln p(O_r, S|\theta)$, and in M-step it finds a new estimation $\theta^{(t+1)}$ to maximizes the Q-function. The details of two steps are as follows:

E-step: In this step, $Q(\theta^{t+1})$ based on the old parameter θ^t in t -th round is computed. Given an observation $O_r = O_{r,1}O_{r,2}\dots O_{r,n}\dots O_{r,N}$, the joint probability distribution over both hidden and observation variables is

$$p(O_r, S|\theta) = p(s_{r,1}|\pi) \sum_{n=2}^N p(s_{r,n}|s_{r,n-1}, A) \sum_{m=1}^N p(O_{r,m}|s_{r,m}, \mu, \sigma). \quad (12)$$

In order to compute the Q-function, the forward distribution $\alpha(s_{r,n})$ and backward distribution $\beta(s_{r,n})$ are defined as,

$$\begin{aligned} \alpha(s_{r,n}) &= p(O_{r,1}, O_{r,2}, \dots, O_{r,n}, s_{r,n}|\theta^{(t)}), \\ \beta(s_{r,n}) &= p(O_{r,n+1}, O_{r,n+2}, \dots, O_{r,N}, s_{r,n}|\theta^{(t)}). \end{aligned} \quad (13)$$

Here, $\alpha(s_{r,n})$ can be computed in a forward fashion, and $\beta(s_{r,n})$ can be computed in a backward fashion:

$$\begin{aligned} \alpha(s_{r,n}) &= p(O_{r,n}|s_{r,n}) \sum_{s_{r,n-1}} \alpha(s_{r,n-1})p(s_{r,n}|s_{r,n-1}), \\ \beta(s_{r,n}) &= \sum_{s_{r,n+1}} \beta(s_{r,n+1})p(O_{r,n+1}|s_{r,n+1})p(s_{r,n+1}|s_{r,n}), \end{aligned} \quad (14)$$

Obtaining the forward distribution $\alpha(s_{r,n})$ and backward distribution $\beta(s_{r,n})$, the other two probabilities of hidden states can be computed: (1) $\gamma(s_{r,n} = k)$, i.e., the probability of n -th hidden state to be k ; (2) $\xi(s_{r,n} = j, s_{r,n+1} = k)$, i.e., the probability of two consecutive states to be j and k . These two distributions are given by

$$\begin{aligned} \gamma(s_{r,n}) &= p(s_{r,n}|O_r) = \alpha(s_{r,n})\beta(s_{r,n})/p(O_r), \\ \xi(s_{r,n}, s_{r,n+1}) &= p(s_{r,n}, s_{r,n+1}|O_r) = \alpha(s_{r,n-1})p(s_{r,n}|s_{r,n-1})p(O_{r,n}|s_{r,n})\beta(s_{r,n})/p(O_r), \end{aligned} \quad (15)$$

where $p(O_r) = \sum_{s_{r,N}} \alpha(s_{r,n})$.

With the definitions of γ and ξ , the Q-function can be expressed from (12) as,

$$Q(\theta, \theta^t) = \sum_{r=1}^R \sum_{k=1}^K \gamma(s_{r,1}^k) \ln \pi_k + \sum_{r=1}^R \sum_{n=2}^N \sum_{j=1}^K \sum_{k=1}^K \xi(s_{r,n-1}^j, s_{r,n}^k) \ln A_{j,k} + \sum_{r=1}^R \sum_{n=1}^N \sum_{k=1}^K \gamma(s_{r,n}^k) \ln p(O_{r,n}|\mu, \sigma), \quad (16)$$

where $s_{r,n}^k$ is binary variable, and $s_{r,n}^k = 1$ means the n -th hidden state is k . The whole training process is shown in Algorithm 1.

M-step: From (16), it's interesting to see that the initial distribution π , the transition probabilities A and the mean and variance of observation probability μ, σ are independent. Using Appropriate Lagrange multipliers, the parameters are easily to be achieved as follows:

$$\begin{aligned}
\pi_k^{(t+1)} &= \frac{1}{R} \sum_{r=1}^R \gamma(s_{r,1}^k), \\
A_{j,k}^{(t+1)} &= \frac{1}{\Xi_j} \sum_{r=1}^R \sum_{n=2}^N \xi(s_{r,n-1}^j, s_{r,n}^k), \\
\mu_{k,l}^{(t+1)} &= \frac{1}{\Gamma_K} \sum_{r=1}^R \sum_{n=1}^N \gamma(s_{r,n}^k) o_{r,n,l}, \\
\sigma_{k,l}^{(t+1)} &= \frac{1}{\Gamma_K} \sum_{r=1}^R \sum_{n=1}^N \gamma(s_{r,n}^k) (o_{r,n,l} - \mu_{k,l}^{(t+1)})^2,
\end{aligned} \tag{17}$$

where $\Gamma_K = \sum_{r=1}^R \sum_{n=1}^N \gamma(s_{r,n}^k)$, $\Xi_j = \sum_{r=1}^R \sum_{n=2}^N \sum_{i=1}^K \xi(s_{r,n-1}^j, s_{r,n}^i)$.

ALGORITHM 1: HMM training by EM algorithm

Input: Observation dataset $O = \{O_1, O_2, \dots, O_R\} (1 \leq r \leq R)$, Maximum Iterations $MaxIter$;

Output: HMM parameters $\theta = \{\pi, A, \mu, \sigma\}$;

Procedure:

Initialization: $t = 0$, initial $\pi_k^{(0)} = 1/K$, $A_{j,k}^{(0)} = 1/K$, $\mu_{k,l}^{(0)} = random(0, 1)$, $\sigma_{k,l}^{(0)} = 0.01$, $\forall 1 \leq j, k \leq K, 1 \leq l \leq L$.

while $t < MaxIter$ **do**

E-step: Calculate $\alpha(s_{r,n})^{(t+1)}$, $\beta(s_{r,n})^{(t+1)}$, $\gamma(s_{r,n})^{(t+1)}$, $\xi(s_{r,n})^{(t+1)}$, $\forall 1 \leq r \leq R, 1 \leq n \leq N$ utilizing old parameters $\theta^{(t)}$ by (14) - (15).

M-step: Update $\pi_k^{(t+1)}$, $A_{j,k}^{(t+1)}$, $\mu_{k,l}^{(t+1)}$, $\sigma_{k,l}^{(t+1)}$, $\forall 1 \leq j, k \leq K, 1 \leq l \leq L$ utilizing $\gamma(s_{r,n})^{(t+1)}$, $\xi(s_{r,n})^{(t+1)}$ by (17).

Update $t: t = t + 1$

end

APPENDIX II: EVALUATION METRICS

For each region r in each time slot n into the test set, we denote the truth app usage $\hat{X}(r, n, l) (1 \leq l \leq L_a)$ as a vector $U_{r,n}$ and the corresponding prediction vector as $U_{r,n}^{pre}$. We also denote the truth population as $L_{r,n}$, $A_{r,n}$ and $S_{r,n}$ and the corresponding prediction as $L_{r,n}^{pre}$, $A_{r,n}^{pre}$ and $S_{r,n}^{pre}$. Then the *TopN-hitrate*, *TopN-accuracy* and runs as follows:

$$TopN - hitrate = \left(\sum_r \left(|U_{r,n} \cap U_{r,n}^{pre}| \geq 1 \right) \right) / R, \tag{18}$$

$$TopN - accuracy = \left(\sum_r \frac{|U_{r,n} \cap U_{r,n}^{pre}|}{N} \right) / R, \tag{19}$$

$$Error \ for \ leaving = \left(\sum_r \frac{|L_{r,n}^{pre} - L_{r,n}|}{L_{r,n}} \right) / R, \tag{20}$$

where the error for arriving and staying is the same with the formula for leaving.

REFERENCES

- [1] 2018. Mobile Marketing Statistics compilation. <https://www.smartinsights.com/mobile-marketing/mobile-marketing-analytics/mobile-marketing-statistics/>. (2018).
- [2] Sofiane Abbar, Tahar Zanouda, Noora Al-Emadi, and Rachida Zegour. 2018. City of the People, for the People: Sensing Urban Dynamics via Social Media Interactions. In *Social Informatics*, Steffen Staab, Olessia Koltsova, and Dmitry I. Ignatov (Eds.). Springer International Publishing, Cham, 3–14.
- [3] J. R. Bellegarda and D. Nahamoo. 1990. Tied mixture continuous parameter modeling for speech recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing* 38, 12 (Dec 1990), 2033–2045.
- [4] M. Bishop Christopher. 2016. PATTERN RECOGNITION AND MACHINE LEARNING. In *Springer-Verlag New York*. 605–652.
- [5] Demissie M D, Correia G, and Bento C. 2015. Analysis of the pattern and intensity of urban activities through aggregate cellphone usage. *Transportmetrica A: Transport Science* 11, 6 (2015), 502–524.
- [6] D. L. Davies and D. W. Bouldin. 1979. A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-1, 2 (April 1979), 224–227.
- [7] Adriana de Souza e Silva. 2006. From Cyber to Hybrid: Mobile Technologies as Interfaces of Hybrid Spaces. *Space and Culture* 9, 3 (2006), 261–278.
- [8] Riccardo Di Clemente, Miguel Luengo-Oroz, Matias Travizano, Sharon Xu, Bapu Vaitla, and Marta C González. 2018. Sequences of purchases in credit card data reveal lifestyles in urban populations. *Nature communications* 9 (2018).
- [9] Manuel García Docampo. 2014. Theories of Urban Dynamics. In *International Journal of Population Research*, Vol. 2014.
- [10] Jacques Gabarro-Arpa and Roger Revilla. 2000. Clustering of a molecular dynamics trajectory with a Hamming distance. *Computers & Chemistry* 24, 6 (2000), 693 – 698.
- [11] Marta C. González, César A. Hidalgo, and Albert-László Barabási. 2008. Understanding individual human mobility patterns. *Nature* 453, 7196 (2008), 779–783.
- [12] Xuedong Huang, Fileno Allewa, Hsiao-Wuen Hon, Mei-Yuh Hwang, Kai-Fu Lee, and Ronald Rosenfeld. 1993. The SPHINX-II speech recognition system: an overview. *Computer Speech & Language* 7, 2 (1993), 137 – 148.
- [13] X. D. Huang. 1992. Phoneme classification using semicontinuous hidden Markov models. *IEEE Transactions on Signal Processing* 40, 5 (May 1992), 1062–1067.
- [14] Mei-Yuh Hwang and Xuedong Huang. 1993. Shared-Distribution Hidden Markov Models for Speech Recognition. 1 (11 1993), 414 – 420.
- [15] Maxime Lenormand, Miguel Picornell, Oliva G Cantú-Ros, Thomas Louail, Ricardo Herranz, Marc Barthelemy, Enrique Frias-Martínez, Maxi San Miguel, and José J Ramasco. 2015. Comparing and modelling land use organization in cities. *Royal Society open science* 2, 12 (2015), 150449.
- [16] Maxime Lenormand and José J Ramasco. 2016. Towards a better understanding of cities using mobility data. *Built Environment* 42, 3 (2016), 356–364.
- [17] Kenneth Rose Liang Gu, Jayanth Nayak. 2000. Discriminative training of tied-mixture HMM by deterministic annealing. *ICSLP-2000* 4 (2000).
- [18] Zhong-Xun Liao, Yi-Chin Pan, Wen-Chih Peng, and Po-Ruey Lei. 2013. On Mining Mobile Apps Usage Behavior for Predicting Apps Usage in Smartphones. In *Proceedings of the 22Nd ACM International Conference on Information & Knowledge Management (CIKM '13)*. ACM, New York, NY, USA, 609–618.
- [19] Thomas Louail, Maxime Lenormand, Miguel Picornell, Oliva García Cantú, Ricardo Herranz, Enrique Frias-Martínez, José J Ramasco, and Marc Barthelemy. 2015. Uncovering the spatial structure of mobility networks. *Nature Communications* 6 (2015), 6007.
- [20] Thomas Louail, Maxime Lenormand, Oliva G Cantu Ros, Miguel Picornell, Ricardo Herranz, Enrique Frias-Martínez, José J Ramasco, and Marc Barthelemy. 2014. From mobile phone data to the spatial structure of cities. *Scientific reports* 4 (2014), 5276.
- [21] Q. Lv, Y. Qiao, N. Ansari, J. Liu, and J. Yang. 2017. Big Data Driven Hidden Markov Model Based Individual Mobility Prediction at Points of Interest. *IEEE Transactions on Vehicular Technology* 66, 6 (June 2017), 5204–5216.
- [22] Ping Wang Marcus Berliant. 2015. Dynamic Urban Models: Agglomeration and Growth. In *Urban Dynamics and Growth: Advances in Urban Economics* (Mar 2015), 533–581.
- [23] Wesley Mathew, Ruben Raposo, and Bruno Martins. 2012. Predicting future locations with hidden Markov models. In *Proceedings of the 2012 ACM conference on ubiquitous computing*. ACM, 911–918.
- [24] F. Miranda, H. Doraiswamy, M. Lage, K. Zhao, B. Gonçalves, L. Wilson, M. Hsieh, and C. T. Silva. 2017. Urban Pulse: Capturing the Rhythm of Cities. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (Jan 2017), 791–800.
- [25] JEFFREY D. MORENOFF, ROBERT J. SAMPSON, and STEPHEN W. RAUDENBUSH. [n. d.]. NEIGHBORHOOD INEQUALITY, COLLECTIVE EFFICACY, AND THE SPATIAL DYNAMICS OF URBAN VIOLENCE*. *Criminology* 39, 3 ([n. d.]), 517–558.

- [26] Flávio Nunes. 2006. The Portuguese urban system: An opposition between its hierarchical organization in cyberspace vs. physical space. *Telematics and Informatics* 23, 2 (2006), 74 – 94.
- [27] Jiaul H. Paik. 2013. A Novel TF-IDF Weighting Scheme for Effective Ranking. In *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '13)*. ACM, New York, NY, USA, 343–352.
- [28] Bryan C Pijanowski, Daniel G Brown, Bradley A Shellito, and Gaurav A Manik. 2002. Using neural networks and GIS to forecast land use changes: a Land Transformation Model. *Computers, Environment and Urban Systems* 26, 6 (2002), 553 – 575.
- [29] L. R. Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 2 (Feb 1989), 257–286.
- [30] Hongbo Si, Yue Wang, Jian Yuan, and Xiuming Shan. 2010. Mobility Prediction in Cellular Network Using Hidden Markov Model. (02 2010), 5 pages.
- [31] Jameson L. Toole, Michael Ulm, Marta C. González, and Dietmar Bauer. 2012. Inferring Land Use from Mobile Phone Activity. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing (UrbComp '12)*. ACM, New York, NY, USA, 1–8.
- [32] Zhen Tu, Runtong Li, Yong Li, Gang Wang, Di Wu, Pan Hui, Li Su, and Depeng Jin. 2018. Your Apps Give You Away: Distinguishing Mobile Users by Their App Usage Fingerprints. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 138 (Sept. 2018), 23 pages.
- [33] ANDREW J. VITERBI [n. d.]. *Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm*. 41–50.
- [34] Huandong Wang, Fengli Xu, Yong Li, Pengyu Zhang, and Depeng Jin. 2015. Understanding Mobile Traffic Patterns of Large Scale Cellular Towers in Urban Environment. In *Proceedings of the 2015 ACM Internet Measurement Conference, IMC 2015, Tokyo, Japan, October 28-30, 2015*. 225–238.
- [35] Jing Wu, Ming Zeng, Xinlei Chen, Yong Li, and Depeng Jin. 2018. Characterizing and Predicting Individual Traffic Usage of Mobile Application in Cellular Network. In *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. ACM, 852–861.
- [36] Fengli Xu, Yong Li, Min Chen, and Sheng Chen. 2016. Mobile cellular big data: linking cyberspace and the physical world with social ecology. *IEEE Network* 30, 3 (2016), 6–12.
- [37] Fengli Xu, Yu-Yun Lin, Jiaxin Huang, Di Wu, Hongzhi Shi, Jeungeun Song, and Yong Li. 2016. Big Data Driven Mobile Traffic Understanding and Forecasting: A Time Series Approach. *IEEE Trans. Services Computing* 9, 5 (2016), 796–805.
- [38] Fengli Xu, Tong Xia, Hancheng Cao, Yong Li, Funing Sun, and Fanchao Meng. 2018. Detecting Popular Temporal Modes in Population-scale Unlabelled Trajectory Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 46 (March 2018), 25 pages.
- [39] Fengli Xu, Pengyu Zhang, and Yong Li. 2016. Context-aware real-time population estimation for metropolis. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, UbiComp 2016, Heidelberg, Germany, September 12-16, 2016*. 1064–1075.
- [40] Hongyi Yao, Gyan Ranjan, Alok Tongaonkar, Yong Liao, and Zhuoqing Morley Mao. 2015. SAMPLES: Self Adaptive Mining of Persistent Lexical Snippets for Classifying Mobile Application Traffic. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking (MobiCom '15)*. ACM, New York, NY, USA, 439–451.
- [41] Donghan Yu, Yong Li, Fengli Xu, Pengyu Zhang, and Vassilis Kostakos. 2018. Smartphone App Usage Prediction Using Points of Interest. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 174 (Jan. 2018), 21 pages.
- [42] Jing Yuan, Yu Zheng, and Xing Xie. 2012. 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 (KDD '12)*. ACM, New York, NY, USA, 186–194.
- [43] N. J. Yuan, Y. Zheng, X. Xie, Y. Wang, K. Zheng, and H. Xiong. 2015. Discovering Urban Functional Zones Using Latent Activity Trajectories. *IEEE Transactions on Knowledge and Data Engineering* 27, 3 (March 2015), 712–725.
- [44] N. J. Yuan, Y. Zheng, X. Xie, Y. Wang, K. Zheng, and H. Xiong. 2015. Discovering Urban Functional Zones Using Latent Activity Trajectories. *IEEE Transactions on Knowledge and Data Engineering* 27, 3 (March 2015), 712–725.
- [45] Chao Zhang, Keyang Zhang, Quan Yuan, Haoruo Peng, Yu Zheng, Tim Hanratty, Shaowen Wang, and Jiawei Han. 2017. Regions, Periods, Activities: Uncovering Urban Dynamics via Cross-Modal Representation Learning. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 361–370.
- [46] Chao Zhang, Keyang Zhang, Quan Yuan, Fangbo Tao, Luming Zhang, Tim Hanratty, and Jiawei Han. 2017. React: Online multimodal embedding for recency-aware spatiotemporal activity modeling. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, 245–254.
- [47] Chao Zhang, Keyang Zhang, Quan Yuan, Luming Zhang, Tim Hanratty, and Jiawei Han. 2016. 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 (KDD '16)*. ACM, New York, NY, USA, 1305–1314.
- [48] Ke Zhang, Qiuye Jin, Konstantinos Pelechrinis, and Theodoros Lappas. 2013. On the Importance of Temporal Dynamics in Modeling Urban Activity. In *Proceedings of the 2Nd ACM SIGKDD International Workshop on Urban Computing (UrbComp '13)*. ACM, New York,

- NY, USA, Article 7, 8 pages.
- [49] Yu Zheng Zhang, Junbo and Dekang Qi. 2017. Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction. In *AAAI*. 1655–1661.
 - [50] Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. 2014. Urban Computing: Concepts, Methodologies, and Applications. *ACM Trans. Intell. Syst. Technol.* 5, 3, Article 38 (Sept. 2014), 55 pages.
 - [51] M. Zhou, M. Wang, and Q. Hu. 2013. A POI data update approach based on Weibo check-in data. In *2013 21st International Conference on Geoinformatics*. 1–4.
 - [52] Wanzheng Zhu, Chao Zhang, Shuochoao Yao, Xiaobin Gao, and Jiawei Han. 2018. A Spherical Hidden Markov Model for Semantics-Rich Human Mobility Modeling. In *AAAI*.
 - [53] Bin Liu Wangsu Hu Hui Xiong Zijun Yao, Yanjie Fu. 2018. Representing Urban Functions through Zone Embedding with Human Mobility Patterns. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*.

Received August 2018; revised October 2018; accepted January 2019