Understanding Urban Dynamics via State-sharing Hidden Markov Model

Tong Xia[1*], Yue Yu[1*], Fengli Xu[1], Funing Sun[2], Diansheng Guo[2], Depeng Jin[1], Yong Li[1]

[1]Beijing National Research Center for Information Science and Technology,
Department of Electronic Engineering, Tsinghua University, Beijing, China

[2]Tencent Inc., Beijing, China

liyong07@tsinghua.edu.cn

ABSTRACT

Modeling people's activities in the urban space is a crucial socio-economic task but extremely challenging due to the deficiency of suitable methods. To model the temporal dynamics of human activities concisely and specifically, we present State-sharing Hidden Markov Model (SSHMM). First, it extracts the urban states from the whole city, which captures the volume of population flows as well as the frequency of each type of Point of Interests (PoIs) visited. Second, it characterizes the urban dynamics of each urban region as the state transition on the shared-states, which reveals distinct daily rhythms of urban activities. We evaluate our method via a large-scale real-life mobility dataset and results demonstrate that SSHMM learns semantics-rich urban dynamics, which are highly correlated with the functions of the region. Besides, it recovers the urban dynamics in different time slots with an error of 0.0793 , which outperforms the general HMM by 54.2%.

CCS CONCEPTS

• **Information systems** → **Spatial-temporal systems**; **Data mining**; • **Human-centered computing** → *Empirical studies in ubiquitous and mobile computing*.

KEYWORDS

Urban Dynamics; Mobility; Hidden Markov Model

1 INTRODUCTION

The rapid urbanization process makes the topic of understanding urban dynamics, i.e. the temporal patterns of urban activities becomes increasingly important. However, traditional approaches rely on expensive manual surveys, yet the understanding is still coarse-grained and limited in geographical scope [1]. Fortunately,

---

* These two authors contributed equally to this research.

---

the advent of the mobile Internet and Location-Based Social Networks (LBSNs) makes it possible to collect population-scale urban mobility data, which sheds new light on this open problem. Previous works demonstrated that the movements of citizens can be utilized to infer the functions of urban regions [20, 22], and urban activities (e.g., working, resting, etc.) are closely correlated with urban mobility patterns [19, 23], which indicates the feasibility of leveraging urban mobility data to model urban dynamics.

In this paper, we aim to harness the power of massive urban mobility data to deepen the understanding of urban dynamics. The research problem is non-trivial mainly for three reasons: (1) Urban mobility behaviour is a *noisy* representation of urban activities. Therefore, it is difficult to robustly and accurately infer the underlying urban activities from the empirical observation of urban mobility behaviour. (2) The semantic-rich mobility data, i.e., check-in data on LBSNs, is sparse in urban space, especially in sparsely populated areas which poses challenges to extract reliable patterns of urban dynamics for different regions in urban systems. (3) Modern cities are complicated socio-economic systems, where each urban region possesses different urban dynamics due to its unique activities. Therefore, it is challenging to interpret the identified urban dynamics and reveal the underlying mechanisms.

Motivated by these challenges, we propose a novel State-sharing Hidden Markov Model (SSHMM), which considers the mobility behavior of an urban region to reveal the urban dynamics. Specifically, the model learns the urban dynamics of a region as transitions between hidden states, where each state maps to a certain urban activity. The corresponding mobility behavior is generated from the hidden state, which allows the same urban activities to be mapped to slightly different mobility behavior and effectively addresses the problem of data noisy. More importantly, SSHMM facilitates different urban regions to share the same set of hidden states. Therefore, it addresses the challenge of data sparsity , which fully exploits the correlation between different regions. Finally, as a generative model, SSHMM can also characterize the urban dynamics as hidden state sequences. Based on the identified state sequences, we further design an clustering analysis technique to reveal their correlation with urban functions (i.e., land use) and provides a meaningful interpretation of the urban dynamics.

The contributions of our research are three-fold:

(1) We propose a novel urban dynamics revealing model SSHMM. It can robustly and accurately infer the underlying urban activities from noisy and sparse urban mobility data by sharing model parameters across different regions. In addition, it achieves qualitative representations of urban dynamics as the transition patterns between urban activities.

(2) We propose an effective and efficient algorithm to infer the parameters of our model. By splitting the long observations into shorter ones and updating the parameters in parallel, we reduce the training time of learning R groups of parameters to that of only one group, where R is the number of regions.

(3) We evaluate our method using population-scale mobility dataset, which demonstrate that our SSHMM model learns meaningful and explainable urban dynamics.

2 RELATED WORK

Urban dynamics modeling. Urban dynamics, generally defined as how sociological indicators change over time [5] which focuses on characterizing human daily activity rhythms in the city [1, 14]. Time is an important dimension to understand the city. Abbar et al. [1] built activity time series in cities to identify different dynamic patterns via the geo-tagged data from Twitter. Zhang et al. [23] utilized the geo-tagged social media to model people's activities in the urban space via representative learning. Zong et al. [27] investigated to generate dynamic population distributions from static PoIs. Lin et al. [12] considered the impact of PoIs in different time of one day to achieve better crowd flow prediction. Besides, Yuan et al. [22] proposed an LDA model to detect the existence of different functional regions in a city through the GPS trajectory and PoI datasets, and Yao et al. [20] presented a city zone embedding framework using human mobility patterns to infer urban functions. Different from the existing works based on statistical analysis [1] and data visualization [14], we propose a specific model to learn hidden state from human mobility and check-ins to represent urban dynamics.

Hidden Markov Model and its application. Hidden Markov Model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (i.e., hidden) states [16]. In HMM, model mixture and parameter sharing are very helpful to deal with increasingly complex tasks [6]. One of the well-known mixture methods is the Gaussian mixture model based HMM (GMM-HMM). The second example is shared-distribution HMM, where clustering is carried out for parameters sharing and output distributions are shared with each other if they exhibit acoustic similarity [8]. Another model tied-mixture HMM uses both mixture and parameter sharing, which belongs to semi-continuous HMM [2, 7, 11].

Our proposed SSHMM is also a kind of parameters sharing HMM. While different from previous works, we design it to automatically learn a set of states for the continuous observations and no following clustering is conducted to force the parameters shared. Both based on the Gaussian emission function, we use multi-dimentional Gaussian instead of GMM. In addition, though HMM has been successfully applied into the topic of mobility modeling, most of the work concentrated on individual mobility prediction [13, 24, 26], and we are the first to apply it in urban dynamics modeling.

3 OVERVIEW

3.1 Problem Description

Before formally define our investigated problem, we give the definition of mobility behaviour observation as follows,

Definition 1 (Mobility Behaviour Observation). *The mobility behaviour sequences of region r is a time-ordered sequence $O_r = [O_{r,1}, O_{r,2}, ..., O_{r,N}]$, where $O_{r,n}$ is a tuple of length L, standing for the observation in n-th time slot. It contains two parts: (1) $\{o_{r,n,1}, o_{r,n,2}, o_{r,n,3}\}$ denotes the number of people who arrive at, leave from and stay in this region in this time slot. (2) $\{o_{r,n,4}, o_{r,n,5}, ..., o_{r,n,L}\}$ denotes the check-in frequency of different categories of PoIs.*

Now, we formulate the urban dynamics understanding problem. Given the mobility behaviour observations of a group of regions in the city, we aim to (1) discover hidden states represented by the volume of population and the visited frequency of different PoIs to understand the intrinsic life modes in the city; (2) reveal urban dynamics represented by hidden state sequences to understand the daily life rhythms and their correlation with urban functions; (3) characterize dynamic regularity by learning state transition probability to achieve dynamic prediction.

3.2 Model Description
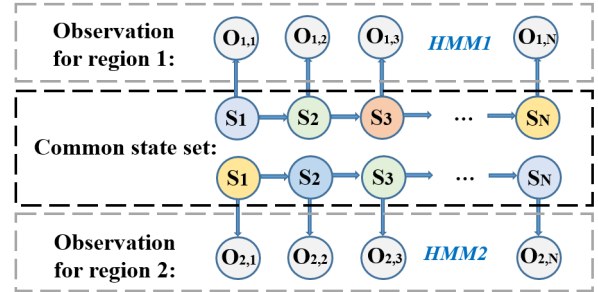
Figure 1: An illustration of the SSHMM, where two kinds of dynamics for region 1 and 2 are generated from the same set of states and each state is presented by a unique color.

Definition 2 (SSHMM). *HMM is parameterized by two parts, one is the state transition parameter characterizing how the states transit and the other is the state emission parameter characterizing how the observation generated by the state. SSHMM contains R groups of HMMs for R groups of observation sequences. However, these HMMs share the same parameters, which means all the observations are generated by the same set of hidden states.*

Our model is based on the general assumptions of HMM, where each observation $O_{r,n}$ is generated from one hidden state s_n, which has L dimensional features, and the n-th hidden state s_n merely depends on the previous state s_{n-1}. We use one HMM denoted by θ_r to model the dynamics of r-th region, therefore we need to learn R groups of transition parameters for R regions in the city. Figure 1 gives the illustration of our model. There are two observation sequences for region 1 and 2, so we learn two HMMs for them respectively. These observations are generated by the same set of states, since life modes (i.e., sleeping, working, etc.) in a city are limited and there exists some similarities between different regions with similar functions. Moreover, the massive mobility data is often sparse and mobility behaviour is a noisy representation of urban activities, which indicates that we can achieve more robust and accurate models by sharing states. These states appear in different time slots, which reveals different dynamics of these two regions.

4 METHODOLOGY

4.1 Data Pre-processing

We first adopt the map segment method [22] to obtain the geographical boundaries of each region formed by the road network. In view of the urban mobility patterns and life-styles, those regions are used as an unit to reveal the dynamics instead of simply dividing the city into grids. To share the states in the city, we normalize mobility behaviour observations to eliminate the problem regarding the difference of population between regions. For mobility, i.e., the number of arriving, leaving and staying, we directly conduct min-max normalization for each region over different time slots respectively. For check-ins, i.e., the visit frequency of PoIs, we first compute the TF-IDF [15] weights based on the region-PoI matrix in each time slot to indicate how popular a PoI category is in a certain time slot. Then, for each region we conduct the min-max normalization on the TF-IDF weights over different time slots to re-scale the mobility behaviour observations to $0 - 1$. After pre-processing, we regard the aggregated activities in the city as normalized time series, which is denoted by $O = \{O_1, O_2, ..., O_r, ..., O_R\}$ as **Definition 1**.

4.2 Model Definition

When generating $o_{r,n}$ from state s_n, we assume that the emission probability is Gaussian distribution, i.e., $p(o_{r,n,l}|s_n) = N(o_{r,n,l}|\mu_n, \sigma_n)$. The reason to chose the Gaussian is intuitive: μ_n describes the fundamental feature of s_n, while the corresponding observations generated by s_n is slightly different from the mean μ_n with the difference controlled by the variance σ_n. Therefore, we build the SSHMM parameterized by R groups of parameters $\theta = \{\theta_1, \theta_2, ..., \theta_r, ..., \theta_R\}$ with $\theta_r = \{\pi_r, A_r, \mu, \sigma\}$ denoted for r-th region, where
(1) $\pi_r \in \mathbb{R}^{K \times 1}$ denotes the initial distribution over K hidden states, i.e., $\pi_{r,k} = p(s_1 = k)(1 \le k \le K)$;
(2) $A_r \in \mathbb{R}^{K \times K}$ 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 $A_{r,j,k}$, i.e.,$p(s_n = k|s_{n-1} = j) = A_{r,j,k}$;
(3) $\mu, \sigma \in \mathbb{R}^{K \times L}$ denotes the mean and variance of observation probability, i.e., $p(O_{r,n}|s_n = k) = \prod_{l=1}^{L} \frac{1}{\sqrt{2\pi}\sigma_{k,l}} \exp\left(-\frac{(o_{r,n,l}-\mu_{k,l})^2}{2\sigma_{k,l}}\right)$.

It is worth noting that we do not use the subscript r to distinguish μ, σ of different HMMs, because in SSHMM they are determined by the common state set.

4.3 Model learning

4.3.1 Parameter Inference. To infer the parameters, we use Expectation-Maximization method (EM) [4] as the solution. Since SSHMM shares the same set of hidden states, the existing famous Baum-Welch algorithm [18] cannot be applied directly. The insight of EM method is to maximize log likelihood, parameters can be updated by maximizing $Q(\theta_r, \theta_r^t)$ step by step. Therefore, we derive the Q-function for all regions as Eq.1, and we can maximize it to enable the log likelihood converge to its maximum.

$$Q(\theta, \theta^t) = \sum_{r=1}^{R} \sum_{S} p(S|O_r, \theta_r^t) \ln \pi_{r,k} + \sum_{r=1}^{R} \sum_{S} \sum_{n=1}^{N-1} p(S|O_r, \theta_r^t) \ln p(s_{n+1}|s_n)$$
$$+ \sum_{r=1}^{R} \sum_{S} \sum_{n=1}^{N} p(S|O_r, \theta_r^t) \ln p(O_{r,n}|s_{r,n}),$$
$$(1)$$

where $p(O_r|S, \theta_r^t)$ can be calculated as follows,

$$p(O_r|S, \theta_r^t) = \pi_{r,k} p(O_{r,1}|s_{r,1}) \cdot p(s_{r,2}|s_{r,1}) p(O_{r,2}|s_{r,2})$$
$$\cdots (s_{r,N}|s_{r,N-1}) p(O_{r,N}|s_{r,N}).$$
$$(2)$$

From Eq. 1, we observe that optimization parameters π, A (i.e., $p(s_{n+1}|s_n)$) and $\{\mu, \sigma\}$ (i.e., $p(O_{r,n}|s_{r,n})$) appear independently, so we maximize each item separately. Consequently, the forward distribution $\alpha(s_{r,n})$ and backward distribution $\beta(s_{r,n})$ are defined as

$$\alpha(s_{r,n}) = p(O_{r,n}|s_{r,n}) \sum_{s_{r,n-1}} \alpha(s_{r,n-1}) p(s_{r,n}|s_{r,n-1}),$$
$$\beta(s_{r,n}) = \sum_{s_{r,n+1}} \beta(s_{r,n+1}) p(O_{r,n+1}|s_{r,n+1}) p(s_{r,n+1}|s_{r,n}),$$
$$(3)$$

where $\alpha(s_{r,1}) = \pi_{r,k} p(O_{r,1}|s_{r,1} = k)$ and $\beta(s_{r,N} = k) = 1$. Then, two probability can be derived from α and β as follows,

$$\gamma(s_{r,n}) = p(s_{r,n}|O_r) = \alpha(s_{r,n})\beta(s_{r,n})/p(O_r),$$
$$\xi(s_{r,n}, s_{r,n+1}) = p(s_{r,n_1}, s_{r,n}|O_r)$$
$$= \alpha(s_{r,n-1}) p(s_{r,n}|s_{r,n-1}) P(O_{r,n}|s_{r,n})\beta(s_{r,n})/p(O_r),$$
$$(4)$$

where $p(O_r) = \sum_{s_{r,N}} \alpha(s_{r,n})$. Based on Eq. 3 and 4, we can thus optimize the parameters π, A and $\{\mu, \sigma\}$ separately as follows.

$$\pi_{r,k}^{(t+1)} = \gamma(s_{r,1}^k),$$
$$A_{r,j,k}^{(t+1)} = \frac{1}{\Xi_j} \sum_{n=2}^{N} \xi(s_{r,n-1}^j, s_{r,n}^k),$$
$$\mu_{k,l}^{(t+1)} = \frac{1}{\Gamma_K} \sum_{r=1}^{R} \sum_{n=1}^{N} \gamma(s_{r,n}^k) o_{r,n,l},$$
$$\sigma_{k,l}^{(t+1)} = \frac{1}{\Gamma_K} \sum_{r=1}^{R} \sum_{n=1}^{N} \gamma(s_{r,n}^k)(o_{r,n,l} - \mu_{k,l}^{(t+1)})^2,$$
$$(5)$$

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

4.3.2 Learning algorithm. Another difficulty is that the process could be inefficient when applied into large-scale datasets due to two aspects. First, the large dataset with a long period N leads the computation of α and β complicated as Eq. 3, and the result may even cause the float number exceeding. Second, the time complexity of parameter inference is $O(RNLK^2)$. which is mainly costed by the calculation of ξ as Eq. 4. The computational complexity of which is quadratic in K, rendering it inefficient for large K. However, if the time period is short or K is not large enough, the model is unable to capture all the dynamic patterns in the city.

To overcome these problems, we first split the long observation sequences into several shorter sub-sequences for each round. By doing this, we not only utilize all the data for parameter learning but also avoid float-point number exceeding. Second, from Eq. 5, we can observe that the updating of $\pi_{r,k}$ and $A_{r,j,k}$ for region r is independent. Thus, we update them in parallel and cuts down the time complexity to $O(NLK^2)$ approximately. The detailed procedure is shown in Algorithm 1. In each iteration, we run EM-steps for W rounds, and in each round, we fed the subsequences of length N into the model. After calculating the $\alpha(s_{r,n})^{(t+1)}, \beta(s_{r,n})^{(t+1)}, \gamma(s_{r,n})^{(t+1)}, \xi(s_{r,n})^{(t+1)}, \pi_{r,k}^{(t+1)}$ and $A_{r,j,k}^{(t+1)}$ **in parallel** for all the regions in each round, we update the state parameters $\mu_{k,l}^{(t+1)}, \sigma_{k,l}^{(t+1)}$ until the convergence is realized.

Algorithm 1: SSHMM parameter learning.

Input: Observations $O = \{O_1, O_2, ..., O_R\}$, Maximum Iterations $MaxIter$;

Output: $\theta_r = \{\pi_r, A_r, \mu, \sigma\} \ \forall 1 \le r \le R$;

Procedure:

    Initialization: $t = 0$, initial $\pi_{r,k}^{(0)} = 1/K$, $A_{r,j,k}^{(0)} = 1/K$,

    $\mu_{k,l}^{(0)} = random(0, 1), \sigma_{k,l}^{(0)} = random(0, 0.1)$,

    $\forall 1 \le j, k \le K, 1 \le l \le L$.

    while $t < MaxIter$ **do**

        for $w = 1, 2, ..., W$ **do**

            E-step: $\forall 1 \le r \le R$, calculate $\alpha(s_{r,n})^{(t+1)}$,

            $\beta(s_{r,n})^{(t+1)}, \gamma(s_{r,n})^{(t+1)}, \xi(s_{r,n})^{(t+1)}$ **in parallel**

            based on old parameters $\theta_r^{(t)}$ utilizing the w-th

            subsequence of O_r.

            M-step: $\forall 1 \le r \le R$, update $\pi_{r,k}^{(t+1)}, A_{r,j,k}^{(t+1)}$ **in parallel**.

            Update $\mu_{k,l}^{(t+1)}, \sigma_{k,l}^{(t+1)}$ utilizing $\gamma(s_{r,n})^{(t+1)}$,

            $\xi(s_{r,n})^{(t+1)} \ \forall 1 \le r \le R$.

        Update t**:** $t = t + 1$

4.4 Dynamics Revealing and Functions Inferring

With the model $\theta_r = \{\pi_r, A_r, \mu, \sigma\}$ obtained from Algorithm 1, the state sequences can be decoded by Viterbi algorithm [17], which can represent the dynamics of that state. Recall that regions with similar dynamic patterns could have similar functions, we match the dynamics with the functions by clustering the state sequence of different regions via k-medoids algorithm [9]. In k-medoids algorithm, we define the distance of two sequences as the sum of the Euclidean distance of the mean of the two corresponding states. Then, the distance between states is defined as the Euclidean distance between the mean value vector of Gaussian distribution of these two states. Finally, the distance between state sequences S_i, S_j is the average distance of all the corresponding state sequences.

We adopt *Davies-Bouldin index* (DBI) [3] to determine the number of clusters, which reflects the ratio between inter-cluster distance and inter-cluster distance. A smaller DBI usually indicates a more effective division. Finally, similar dynamics in each cluster would present the same kind of function in the city.

4.5 Dynamics Prediction

SSHMM also has the ability of prediction. Through Viterbi algorithm, we can identify the last state of the region based on the observation. With the advantage of a probabilistic model, we can firstly obtain the current state of the region according to the current observation, then predict the next state according to the latest state by maximizing the transition probability. Formally, given the current observation sequence of r-th region denoted by $O_{r,1}O_{r,2}, ..., O_{r,n}$, we decode its corresponding hidden state sequence as $s_{r,1}s_{r,2}, ..., s_{r,n}$. If the latest state $s_{r,n}$ is i-th state, then the state in the next time slot can be predicted as $s_{r,n+1} = \arg\max_{1 \le j \le K} A_{r,i,j}$. This is to say, we can predict the volume of population flows and the percentage of PoIs visited in the next time slot by utilizing the mean value μ, σ of the identified next state.

5 EXPERIMENTS

5.1 Data

The mobility dataset was collected by collaborating with *Tencent*[1], one of the largest Internet service providers in China. The collected dataset covers near 2 million users in Beijing, China from April 1st to 30th, 2018. Each trajectory record is characterized by an anonymized user ID, timestamp and the GPS location. Each check-in record consists of user ID, check-in time, check-in location and PoI, which is divided into nine categories: *Company, Agency, Shopping, Service, Entertainment, Attractions, Education* and *Residence*. We crawl the road network from Map service and divide the the downtown area in Beijing into 665 non-overlapping regions.

Ethics. To protect user privacy, all data is anonymous and stored in Tencent offline servers. We pre-process the data under their supervision and only take the aggregated results of different regions for further analysis. Our research has been reviewed and approved by both the company and Tsinghua University institutional board.

5.2 System Settings

5.2.1 Data Usage. In the experiments, we divide the dataset into two parts. We utilize 21 days of data to generate the mobility behaviour observations to train our model and use the rest for prediction evaluation. We set the length of time slot to 1 hour. The observation in each time slot is 12-dimensional, including 3-dimensional population flow volume and 9-dimensional check-in frequency.

5.2.2 Model evaluation. To evaluate the effectiveness of our model and the inferred parameters, we utilize the obtained states to recover the observations by concatenating the mean value of corresponding state in the hidden state sequence in chronological order. We adopt $\epsilon = \sqrt{\frac{\sum_{r=1}^{R} \sum_{n=1}^{N} \sum_{l=1}^{L} (\mu_{n,l} - s_{r,n,l})^2}{R \times N \times L}}$, the error between all of the observations and the activities recovered by the mean value of the corresponding hidden state as an evaluation metric.
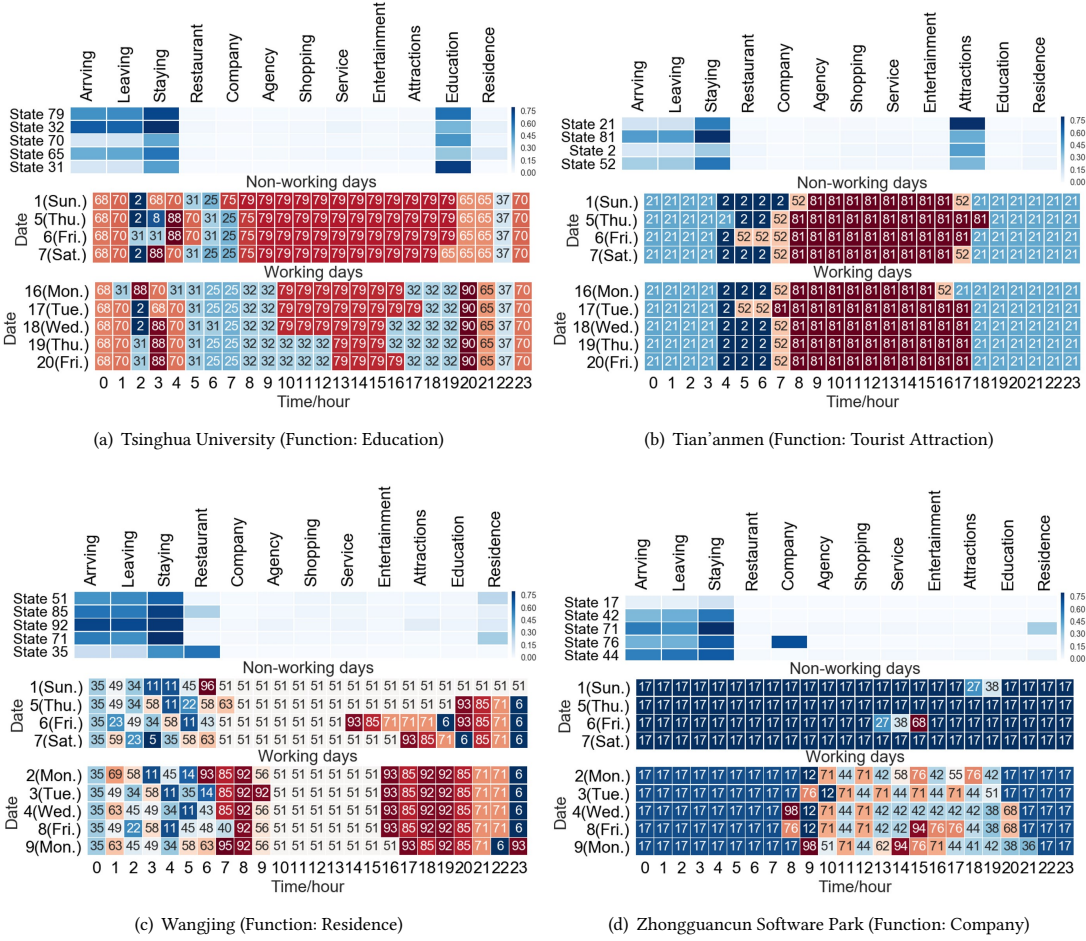
To determine the number of hidden states and the length of time slots for training, we searched different values. From fig. 3(a), we can observe that when K increases, ϵ decreases while the training time increases rapidly. Thus, we set $K = 100$ as a trade-off between model complexity and accuracy. As Figure 3(b) show, we find the length of time window has almost no impact on the training time and ϵ. We set is as 24, i.e., train the model day by day.

5.2.3 Prediction Performance. In terms of prediction, we use the rest 9 days' data for evaluation. We compare the mean value of the predicted next state with the ground-truth observations, and the baseline is general HMM. For mobility prediction, we adopt *RMSE*, the root-mean-square error between the normalized the population flow of predicted and the ground truth [25] for measurement. For check-ins, we adopt *TopM-accuracy* to reflect the accuracy on topN frequently-visited PoI prediction of all regions [21]. A lower *RMSE* or a higher *TopM-accuracy* indicates better prediction performance.

5.3 Results

5.3.1 Model Effectiveness. Figure 4 shows the relationship between recovering error ϵ and the number of iterations. We compare it with

---

(a) Tsinghua University (Function: Education)

(b) Tian'anmen (Function: Tourist Attraction)

(c) Wangjing (Function: Residence)
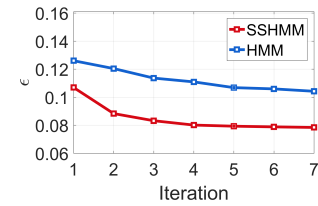
(d) Zhongguancun Software Park (Function: Company)

Figure 2: Visualization for representative states and state sequences for different regions, where each row in the state sequences heatmap exhibits the state transition process of one day with the number indicating the corresponding state.

(a) Effects of K
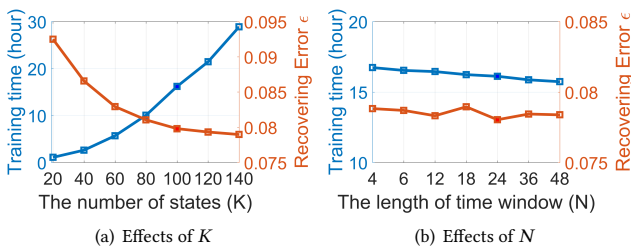
(b) Effects of N

Figure 3: Effects of parameters.

Figure 4: Model Effectiveness comparison.

the result of training independent HMM for each region with the same number of states. From Figure 4, we can observe that as the number of iterations increases, the recovering error ϵ decreases first, then tends to remain unchanged. Furthermore, our representative ability for human activities is better than HMM: When the model converges, our ϵ is 0.0793, which outperforms HMM by 54.2%.

5.3.2 *Visualization of the Hidden States and Urban Dynamics.* In order to demonstrate the ability of our model in discovering hidden states and revealing urban dynamics, we give a series of special examples and detailed explanations. Figure 2 visualizes the results for 6 regions. We first plot the mean value of the states with frequent occurrence, then show the state transition process in working day and non-working day (Apr. 5th (Thur.), 6th (Fri.) and 7th (Sat.) are the Qingming Festival), respectively.

States: Each state represents two aspects: (1) The density of the population flow. For example, **state 32** presents a large volume of flows and high density of populations, **state 21** presents a small volume of flows while a high density of populations. (2) The visit frequency of different PoIs. For example, **state 31** indicates the most frequently visited PoI is *Education*, while **state 21** indicates the most frequently visited PoI is *Attractions*. As shown in Figure 2(a), *Tsinghua University* have the **state 79** during the day, and *Tian'anmen* has **state 81**. By combining these two semantics, we can infer the activity level and lifestyle of the region. For example, **state 79** shows active school status while **state 70** shows quiet school status.

Dynamics: Dynamics are represented by the state transition processes. Take the dynamics of *Tsinghua University* as an example. As shown in Figure 2(a), during the night, there are fewer people than the day as **state 70**. Besides, in working days, there is a sudden increase in crowd flow as **state 32** appear in 8:00-9:00 and 17:00-19:00. The transition from **state 70** to **32**, from **32** to **79** in working days reveals the dynamics that only students live in the region at night and many teachers go to school in the morning, causing the population denser than night. Compared with 2(a), 2(b) shows that both in working day and non-working day, the density of population is consistently high and the PoI visited most frequently is *Attractions*, as *Tian'anmen* is one of the most famous tourist spots in China. Another interesting finding is that for residential areas as Figure 2(c), non-working days are more prosperous and lively than working days, but working areas are quiet and peaceful as Figure 2(d) shown. To conclude, Figure 2 gives the insights as follows:

(1) From the state sequence heatmap, we notice that the state sequences are aligned by day, which indicates that the dynamics in the city have a period of one day, as the states in the same time slot but with different dates are usually the same.

(2) The dynamic patterns within working days or no-working days are very similar, while the patterns between working and non-working days are determined by the function of the region.

In summary, SSHMM not only detects the urban states with different activity patterns but also reveal the rhythm of daily life.

5.3.3 Visualization of Urban Functions. To further evaluate our model to infer the functions, we cluster the state sequences by the method introduced in Section 4.4 and found that 8 types is most suitable as the minimum DBI is achieved. We compare the result with the state-of-the-art functional zones discovering method, i.e., an LDA model by utilizing the static PoIs and mobility together [22]. The geographical distribution of the regions with their function types is shown in Figure 5, where different colors present different functions. After clustering, we label their functions by the revealed semantic dynamics and manually check some regions in Figure 2 to verify their functions, which shows most of the regions with the same functions are divided into one cluster. Besides, clusters obtained by our model and LDA model have with the Normalized Mutual Information (NMI) of 0.25, which measures the similarity of two divisions with the range from -0.5 to 1 [10]. Therefore, the functions we infer from dynamic patterns are validated by the labeled regions and the state-of-the-art method.

5.3.4 Performance of Activity Prediction. In order to evaluate the prediction accuracy of our model, we first show the prediction of the

(a) SSHMM   (b) LDA
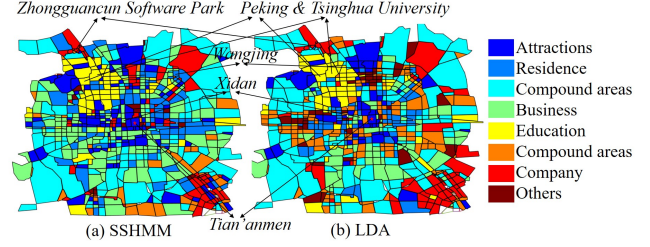
Figure 5: Visualization of the distribution of the regions with similar dynamics and functions.

(a) Prediction for the number of staying of *Tsinghua University*

(b) Evaluation metrics
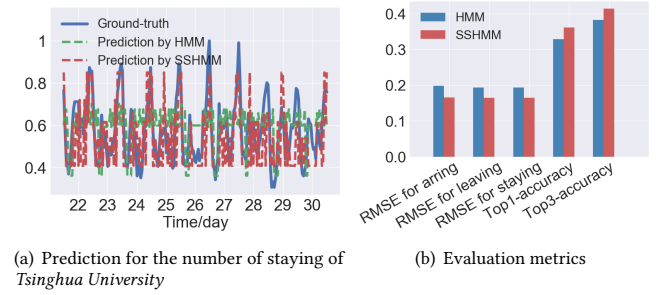
Figure 6: Prediction performance compared with HMM.

number of staying for 9 days in *Tsinghua University* in Figure 6(a) as an example. Compared with the generate HMM, the recovered observations of our model are closer to the ground-truth in different time slots. The results for all metrics are shown in Figure 6(b), where the average *RMSE* of population flow prediction is 0.195 and the *Top3-accuracy* for PoI popularity prediction is 41.4%, outperforming the HMM by 16% and 8% respectively. These results demonstrate our SSHMM outperforms HMM in urban dynamics prediction problem.

6 CONCLUSION

In this paper, we study the problem of understanding urban dynamics. We propose a State-sharing Hidden Markov Model (SSHMM), where all the regions share a common state set, but each region has it own transition regularity. To make it practical, we not only derive the inference of the parameters but also give an efficient algorithm to update them. We evaluate our method via a real-life dataset in Beijing, which demonstrates that SSHMM learns semantics-rich urban dynamics model, recovers different activity regularities by a limited number of states and incurs low training cost. Our work also opens a new angle for dynamic urban region representation by learning the vector for both urban states (i.e, the mean of different feature dimension) and urban dynamics (i.e, the state sequence).

REFERENCES

[1] 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*. 3–14.

[2] 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.

[3] D. L. Davies and D. W. Bouldin. 1979. A Cluster Separation Measure. *T-PAMI* (April 1979), 224–227.

[4] Arthur P Dempster, Nan M Laird, and Donald B Rubin. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)* (1977), 1–38.

[5] Manuel Garcia Docampo. 2014. Theories of Urban Dynamics. In *International Journal of Population Research*, Vol. 2014.

[6] Xuedong Huang, Fileno Alleva, 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. http://www.sciencedirect.com/science/article/pii/S0885230883710077

[7] X. D. Huang. 1992. Phoneme classification using semicontinuous hidden Markov models. *IEEE Transactions on Signal Processing* 40, 5 (May 1992), 1062–1067.

[8] Mei-Yuh Hwang and Xuedong Huang. 1993. Shared-Distribution Hidden Markov Models for Speech Recognition. 1 (11 1993), 414 – 420.

[9] Xin Jin and Jiawei Han. 2011. K-medoids clustering. In *Encyclopedia of Machine Learning*. Springer, 564–565.

[10] Zeger F Knops, JB Antoine Maintz, Max A Viergever, and Josien PW Pluim. 2006. Normalized mutual information based registration using k-means clustering and shading correction. *Medical image analysis* 10, 3 (2006), 432–439.

[11] Kenneth Rose Liang Gu, Jayanth Nayak. 2000. Discriminative training of tied-mixture HMM by deterministic annealing. *ICSLP-2000* 4 (2000).

[12] Ziqian Lin, Jie Feng, Ziyang Lu, Yong Li, and Depeng Jin. 2019. DeepSTN+: Context-aware Spatial-Temporal Neural Network for Crowd Flow Prediction in Metropolis. In *AAAI*.

[13] Wesley Mathew, Ruben Raposo, and Bruno Martins. 2012. Predicting future locations with hidden Markov models. In *UbiComp*. ACM, 911–918.

[14] F. Miranda, H. Doraiswamy, M. Lage, K. Zhao, B. GonÃğalves, 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 (2017), 791–800.

[15] Jiaul H. Paik. 2013. A Novel TF-IDF Weighting Scheme for Effective Ranking. In *SIGIR*. ACM, 343–352.

[16] L. R. Rabiner. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* 77, 2 (Feb 1989), 257–286.

[17] A. Viterbi. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory* 13, 2 (April 1967), 260–269. https://doi.org/10.1109/TIT.1967.1054010

[18] Lloyd R Welch. 2003. Hidden Markov models and the Baum-Welch algorithm. *IEEE Information Theory Society Newsletter* 53, 4 (2003), 10–13.

[19] Fengli Xu, Tong Xia, Hancheng Cao, Yong Li, Funing Sun, and Fanchao Meng. 2018. Detecting Popular Temporal Modes in Population-scale Unlabelled Trajectory Data. *IMWUT* 2, 1 (March 2018), 46:1–46:25.

[20] Zijun Yao, Yanjie Fu, Bin Liu, Wangsu Hu, and Hui Xiong. 2018. Representing Urban Functions through Zone Embedding with Human Mobility Patterns.. In *IJCAI*. 3919–3925.

[21] 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.

[22] Nicholas Jing Yuan, Yu Zheng, Xing Xie, Yingzi Wang, Kai Zheng, and Hui Xiong. 2015. Discovering urban functional zones using latent activity trajectories. *IEEE Transactions on Knowledge and Data Engineering* 27, 3 (2015), 712–725.

[23] 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 *WWW*. 361–370.

[24] 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 *KDD*. ACM, 1305–1314.

[25] Junbo Zhang, Yu Zheng, and Dekang Qi. 2017. Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction.. In *AAAI*. 1655–1661.

[26] Wanzheng Zhu, Chao Zhang, Shuochao Yao, Xiaobin Gao, and Jiawei Han. 2018. A Spherical Hidden Markov Model for Semantics-Rich Human Mobility Modeling.. In *AAAI*.

[27] Zefang Zong, Jie Feng, Kechun Liu, Hongzhi Shi, and Yong Li. 2019. DeepDPM: Dynamic Population Mapping via Deep Neural Network. In *AAAI*.