Discovering Semantic Mobility Pattern from Check-in Data

Ji Yuan, Xudong Liu, Richong Zhang, Hailong Sun,
Xiaohui Guo, and Yanghao Wang

School of Computer Science and Engineering, Beihang University,
Beijing, 100191 China
{yuanji,liuxd,zhangrc,sunhl,guoxh,wangyh}@act.buaa.edu.cn

Abstract. The wealth of check-in data offers new opportunities for better understanding user movement patterns. Existing studies have been focusing on mining explicit frequent sequential patterns. However, the sparseness of check-in data makes it difficult that all explicit patterns be precisely discovered. In addition, due to the weakness in expressing semantic knowledge of explicit patterns, the need for discovering semantic pattern rises. In this paper, we propose the Topical User Transition Model (TUTM) to discover the semantic mobility patterns by analyzing topical transitions. Via this model, we can discover semantic transition properties and predict the user movement preferences. Furthermore, Expectation-Maximization (EM) algorithm incorporating with Forward-Backward algorithm is provided for estimating the model parameters. To demonstrate the performance of TUTM model, experimental studies are carried out and the results show that our model can not only reasonably explain user mobility patterns, but also effectively improve the prediction accuracy in comparison with traditional approaches.

Keywords: Mobility pattern, Spatial semantics, Check-in data

1 Introduction

The location-based data is growing exponentially with the pervasive usage of GPS-enabled mobile devices and mobile applications. These location data contains a large quantity of mobility patterns, which represent the common regulations that people follow when they move among locations in daily life. In practice, user mobility patterns can be explained through many forms, such as destination preference [1], frequent transitions [2], periodic regulations [3], etc. These existing studies mainly focus on interpreting the mobility patterns as explicit frequent sequential movements, such as frequent visiting regulations [4] and trajectory patterns [5]. However, the explicit patterns does not carry any semantic descriptions. In addition, the sparseness, incompleteness and randomness of the user-generated data makes it impossible that all transitional patterns be precisely characterized by statistically analyzing the explicit data. Therefore, the requirement for discovering semantic mobility pattern naturally rises. As a simple example, the semantic pattern of *Office - Food* can be abstracted from the explicit sequential movement "*Office Building A - Restaurant B*". Obviously, this semantic pattern generalizes the human explicit sequential movements and overcomes the data uncertainties.

To model the semantic patterns, researches propose probabilistic approaches [6] [7] by exploiting latent factors for discovering the user hidden mobility preferences. Yet, the transition patterns between locations are not captured. There are some researches that combine the mobility transition and user preferences. For instance, Kustrama [1] *et al.* combine topic model and Markov model for recommending possible destinations. Still, the semantic transitional pattern is not modeled by this hybrid approach, which simply considers the explicit transition as a calculation factor.

In order to discover the semantic mobility pattern from location data, we assume that there are latent topics corresponding to each location, which represents the implicit semantic attributes of location itself and when people transiting from one place to another, semantic transitions between latent topics are generated accordingly. Our objective is to find these common semantic transitions. As the previous example "many people would go to restaurants after work" indicates, such phenomena represents a coarse-gained transition among the latent semantic topics in adjacent locations. These transitions reflect the latent spatial topical semantic pattern.

Based on the previous analysis, in this study we propose the Topical User Transition Model (TUTM): a probabilistic model which leverages the topical transitions among gathered location sequences to model the user mobility patterns in daily life. Specifically, we synthesize the strong dependencies between the consecutive locations and exploit the transition between locations in building this model. Since transitions between locations obviously reveal knowledge about users' latent mobility regulation, the transitional relations are established between consecutive latent topics rather than explicit locations. Generally, evolving from LDA, our model takes into consideration the dependencies between topics, which will be specified in Section 3.1. Parameters of this model are estimated through the widely-used Expectation-Maximization algorithm.

Experimental studies are carried out on a real-world dataset shared by Cheng *et al.* [8] and the results confirm the effectiveness of our model on inferring the semantic mobility patterns by analyzing the topical transitions of the model. Moreover, the applicability of our model on predicting the user destinations is also demonstrated through the comparison with existing methods.

The rest of the paper is organized as follows. We indicate the existing studies related to the problems of mobility patterns modeling firstly in Section 2. Through augmenting the first-order Markov property to the locations' topic assignments, we formulate the topical location transition problem as a probabilistic generative process in section 3. The EM-styled inference method and the next-place predictive probability follows up. Section 4 contains the experiments carried out for verifying model performance. Finally, we conclude our contribution and the future work.

2 Related Work

Mining the knowledge from the location based services and social networks is an appealing direction in the research domain. Cheng *et al.* [8] and Daniel *et al.* [4] provide some large scale statistical quantitative analysis of human mobility pattern in check-in data, considering the temporal-spatial influence and even economic factors. Besides, the work of GPS trajectory pattern mining which is considered by Zheng *et al* [9], develop a framework to mine interesting locations and travel sequences from GPS trajectories. Travel patterns are discovered by Zheng *et al.* [2] based on geo-tagged photo sharing community. They leverage the geo-tags and textual features annotated on the photos to detect the travel patterns at local city level using a markov chain model. Similarly, the frequent trip patterns are also mined in [10]. In the work by Bayir *et al* [11], the spatio-temporal mobility patterns and profiles of mobile phone users are captured using cell phone log data. In the urban computing, the taxicab traces are also utilized to inferring mobility patterns in [12]. Obviously, a variety of location datasets are analyzed in those researches. But they pay less attention on the semantic aspect of the location transition and deterministically mine the frequent patterns, which usually suffer from highly-cost computation and performance degradation due to the sparse and large scale data.

Other than frequent pattern mining, the semantics knowledge mining of human mobility close to our target are also concerned by academic community. In [13], the social-spatial properties in social network are discussed. They study the connections between spatial properties and users' relations. Eagle *et al.* [14] infer the friendships and the spatio-temporal patterns in their physical proximity and calling patterns based on the observational device mobility data, and compare the results with the self-reported data. Xiao *et al.* [15] estimate the similarity between users according to their GPS trajectories. They map every location in the trajectory to a certain category, and measure the similarity between trajectories. Similar research has been conducted on Facebook data in [16] to analyze the correlations between the distance and social relations, and further utilize it to predict users' current location. These works are closely relevant to our work of analyzing the mobility semantics. However, they mostly focus on discovering the impact of the location based services on the social relations, rather than the human transition behaviors semantics itself.

Actually, the transition pattern mining and semantics analysis serve the purpose of facilitating the diversified realistic applications, such as, convenient routes recommendation, next-place prediction, and city governance. In [17], frequent routes are generated for travel recommendation based on trajectory patterns discovering. Monreale *et al.* [5] presents a model called $WhereNext$, which aims at predicting the next location also by trajectory pattern mining. These approaches based on frequent patterns have limitations on discovering semantic transition patterns for modeling mobility. As for semantics based methods, the prediction algorithms considering several semantic temporal-spatial factors are also used for modeling user mobility. Noulas *et al.* [18] have examined plenty of methods to predict the next location, which are statistical models considering various temporal or spatial factors. Baumann *et al.* [19] also study the influence of temporal and spatial features on prediction algorithms. As the method concerning temporal factors, Wang *et al.* [3] propose a periodicity based prediction model to detect the periods for the next place recommendation problem.

There are also several probabilistic models for the location recommendation problem emerged. In the work [20], a Mobility Markov Chain model is established by generalizing the Hidden Markov Model, in which the hidden states do not correspond only to a single location, but rather a sequence of the n previous visited places. Yu *et al.* [6] use the conventional LDA for location preference analysis and recommend locations for mobile users. A hybrid and approximative approach combining Markov and topic models in [1], named Photographer Behavior Model, is presented to handle the trip planning problem by Kurashima *et al.* It seems to coincide with our modeling intentions; however, we coherently build our model using the consistent language of probabilistic graphical models.

All in all, the mobility pattern and semantics analysis for location data are indeed of great practical significance. Nevertheless, seldom works focus on the topical transition semantic analysis of the location based service data in a probabilistic way.

3 Topical User Transition Model

In this section, we represent Topical User Transition Model as a probabilistic generative model. As we analyzed in previous sections, there are two facts that should be highlighted in modeling the focused problem. One is that user generated trajectories collaboratively and statistically exhibit latent spatial topical semantics. The other is that the sequential venues shift might be governed by some coarse-grained transitions among the latent topics.

3.1 Model Specification

Suppose we have the trajectory (i.e. location sequences without ambiguity) set S with the cardinal $|S| = N$. Each sequence $s \in S$ is composed by N_s successive locations, $l_1, l_2, \cdots, l_{N_s}$, and $l_i \in L$; $i = 1, \cdots, N_s$. Here L is the total involved locations or venues set with the cardinal $|L| = M$. When it comes to the latent topic recognizing issue, the probabilistic hierarchical generative model LDA, shown in Fig.1(a), is competent to a large extent. For t^{th} observed location l_t of some s, one first samples l_t's topic assignment $z_t \in \{1, \cdots, K\}$ from the corresponding multinominal distribution probability vector θ_s of the trajectory s. For the trajectory level topic distributions θ_s, LDA imposes the common Dirichlet prior parameterized by α on them. Then, given the topic z_t, the location l_t could be drawn from the multinomial distribution over locations, denoted by ϕ_{z_t}, which is also regulated by a conjugate Dirichlet prior sharing the common concentration hyperparameter vector β. We pile all the ϕs into a matrix Φ of size $K \times M$.

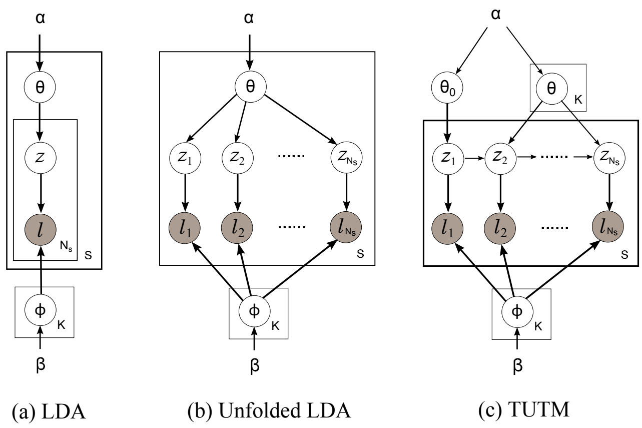


Fig. 1. Graphical Notation Representations of the Models

With the simple "bag of words" assumption, i.e. the topic assignments are independent in a specific "bag", LDA is effective in many domains. But in the user mobility application, the topic of current "word", i.e. location, is usually strongly correlated and dependent on the previous locations in the trajectory "bag", owing to the geographical restriction or users' intentions. For the simplicity and effectiveness, the first-order markov property is employed to address those conditional location transitions probabilistically.

The word layer plate unfolded version of LDA is given in Fig.1(b), where the latent topics' independence assumption is explicitly expressed. Then a Markov chain is introduced to manifest the conditional transitions among the latent topic assignments of locations, as shown in Fig.1(c). Furthermore, we lift the topic multinomial distributions θs from the trajectory layer to the top layer. Meanwhile, $\theta_0 \sim Dirichlet(\alpha)$ is used to initialize the hidden Markov chains' starting states, and $\theta_{1,\cdots,k} \sim Dirichlet(\alpha)$, are stacked by rows to form the $K \times K$ Markov state transition probability matrix Θ. Specifically, each matrix entry θ_{ij} governs the probability of location's topic assignment i shifting to j $(i, j = 1, \cdots, K)$. Thus for two successive locations l_t and l_{t+1} with topic assignments z_{l_t} and $z_{l_{t+1}}$, the corresponding transition probability can be directly determined by $p(z_{l_{t+1}} = j | z_{l_t} = i) = \Theta_{ij}$.

Totally, the proposed Topical User Transition Model (TUTM) can be formulated as following generative process.

1. For each topic $z = 1, \cdots, K$
   draw $\phi_z \sim Dir(\beta)$
2. For the initial topics and the Markov topic transition matrix
   (a) draw $\theta_0 \sim Dir(\alpha)$
   (b) draw $\theta_k \sim Dir(\alpha),\ k = 1, \cdots, K$
3. For each trajectory $s \in S$
   (a) For the first location
      i. draw its topic assignment $z_1 \sim Mul(\theta_0)$

ii. draw $l_1 \sim Mul(\phi_{z_1})$

(b) For the subsequent locations $l_t,\ t = 2, \cdots, N_s$

i. draw its topic assignment $z_t \sim Mul(\Theta_{z_{t-1}})$

ii. draw $l_t \sim Mul(\phi_{z_t})$

Collaboratively, the above generative process is actually encoding following joint probability distribution over latent variables, $Z = \cup_{s=1}^{N}\{z_1^s, z_2^s, \cdots, z_{N_s}^s\}$, and observed variables, $L = \cup_{s=1}^{N}\{l_1^s, l_2^s, \cdots, l_{N_s}^s\}$, and the given model parameters, $\Lambda = \{\theta_0, \Theta, \Phi\}$, For simplicity, the hyperparameters α and β are omitted and assumed to be explicitly dependent.

$$p(Z, L|\Lambda) = \prod_{s=1}^{N} p(z_0^s|\theta_0)\Big[\prod_{t=2}^{N_s} p(z_t^s|z_{t-1}^s, \Theta)\Big]\prod_{r=1}^{N_s} p(l_r^s|z_r^s, \Phi) \tag{1}$$

Compared with traditional LDA, TUTM can obviously and felicitously characterize the human daily mobility or migration, through introducing the Markov property to express the transitional dependency of the successive location's latent topics.

3.2 Inference

The posteriors of the probabilistic models, embedding the LDA as a building block, are usually intractable, and inferred by variational methods, Markov chain Monte Carlo, and expectation propagation, etc. Instead, we employ the widely used Expectation-Maximization (EM) algorithm to estimate our TUTM's model parameters, for the fact that we can incorporate the endowments of the off-the-shelf Forward-Backward algorithm to treat the Markov chain's estimation at the expectation step. Specifically, the Expectation and Maximization steps are formulated as follows:

E-step: We take the initial or previously estimated parameters Λ^{old} values to calculate the posterior of the latent variables $p(Z|X, \Lambda^{old})$. We then use this posterior distribution to evaluate the expectation of the logarithm of the complete-data likelihood function parameterized by Λ of the current iterative round, by summing up all the configuration of the latent variables Z as follows,

$$Q(\Lambda, \Lambda^{old}) = \sum_{Z}\big[p(Z|L, \Lambda^{old}) \ln p(L, Z|\Lambda)\big] \tag{2}$$

We subtitute the joint distribution Eq.1 into Eq.2, then yields,

$$\begin{aligned} Q(\Lambda, \Lambda^{old}) = \sum_{s=1}^{N}\Big\{&\sum_{k=1}^{K} p(z_1^s = k|L, \Lambda^{old}) \ln\theta_0 + \sum_{t=1}^{N_s}\sum_{k=1}^{K} p(z_t^s = k|L, \Lambda^{old}) \ln p(l_t^s|\phi_k) \\ &+ \sum_{t=2}^{N_s}\sum_{j=1}^{K}\sum_{k=1}^{K} p(z_t^s = k, z_{t-1}^s = j|L, \Lambda^{old}) \ln\Theta_{jk}\Big\} \end{aligned} \tag{3}$$

Utilizing the Forward-Backward algorithm, we achieve the posteriors of latent variables, $p(z_t^s|L, \Lambda^{old})$, and their co-occurrences, $p(z_t^s = k, z_{t-1}^s = j|L, \Lambda^{old})$, in Eq.3. Firstly, the two posterior are factorized as follows:

$$p(z_t^s|L, \Lambda^{old}) \propto \alpha(z_t^s)\beta(z_t^s); \quad p(z_t^s{=}k, z_{t-1}^s{=}j|L, \Lambda^{old}) \propto \alpha(z_{t-1}^s)\phi_k \Theta_{jk}\beta(z_t^s)$$

Where $\alpha(z_t^s) = p(l_1, \cdots, l_t^s, z_t^s)$ atd $\beta(z_t^s) = p(l_{t+1}^s, \cdots, l_{N_s}|z_t^s)$ are the forward factors and the backward factors respectively. Then, the algorithm uses a two-stage message passing procedure to exactly calculate the factors through these recursion relations.

$$\alpha(z_t^s) = \sum_{z_{t-1}^s} \alpha(z_{t-1}^s)\Theta^{old}_{z_{t-1}^s, z_t^s}\Phi^{old}_{z_t^s, l_t^s}; \quad \beta(z_t^s) = \sum_{z_{t+1}^s} \beta(z_{t+1}^s)\Theta^{old}_{z_t^s, z_{t+1}^s}\Phi^{old}_{z_{t+1}^s, l_{t+1}^s}$$

Thus, the Eq.3 can be easily evaluated, so as to be used to feed the maximization step. During the Forward-Backward algorithm, we also collect the expectation number of the topic transition pairs (i, j): $i, j = 1, \cdots, K$, and topic-location pairs (p, q): $p = 1, \cdots, K$, and $q = 1, \cdots, M$, i.e.

$$\begin{aligned}
C_{.k} &= \mathbb{E}(z_1 = k) = \sum_{s=1}^{N} p(z_1^s = k|L, \Lambda^{old});\\
C_{ij} &= \mathbb{E}(z_t = i, z_{t+1} = j) = \sum_{s=1}^{N}\sum_{t=1}^{N_s-1} p(z_t^s = i, z_{t+1}^s = j|L, \Lambda^{old});\\
C_{pq} &= \mathbb{E}(z_t = p, l_t = q) = \sum_{s=1}^{N}\sum_{t=1}^{N_s} p(z_t^s = p, l_t^s = q|L, \Lambda^{old}).
\end{aligned}$$

M-step: We maximize the evaluated Eq.3 with respect to the nondeterministic parameters $\Lambda = \{\theta_0, \Theta, \Phi\}$. Lagrange Multiplier methods is employed to determine the new parameters θ_0, Θ, and Φ.

$$\theta_{0k} \propto C_{.k} + \alpha; \quad \Theta_{ij} \propto C_{ij} + \alpha; \quad \Phi_{pq} \propto C_{pq} + \beta$$

Thereby, the EM algorithm starts with some initial selection for the model parameters, Λ^{init}, and then alternates between E and M steps until some convergence criterion is satisfied, for instance when the change in the complete-data likelihood function is below some threshold. Besides in the later experiment sections, we empirically assume that the α and β are fixed and heuristically set $\alpha = 50/K$, $\beta = 0.01$.

3.3 Prediction

Due to the TUTM's powerful ability of modeling user mobility, it is suitable for us to utilize the model for the destination recommendation or prediction. We first define the next-place prediction problem. When predicting the next place, we aim at computing the most possible location conditioned user's visiting history.

Therefore, a set of training sequences S and a set of locations L are defined as input of the focused problem. There is a check-in set C_s corresponding to each $s \in S$, which is composed by a list of check-ins $c_{l,t}$, where $l \in L$ and t represents the timestamp. The order of location list $l_1, l_2, \cdots, l_{N_s}$ is indicated by the timestamp. The prediction problem is formalized as: Given a sequence $l_1, l_2, \cdots, l_t$, how to rank the locations in L based on $P(l_{t+1}|l_1, l_2, ..., l_t)$ in order that the actual visited location is top-ranked.

The probability of the next place is calculated using Θ, Φ and $\alpha(z_t)$:

$$p(l_{t+1}|l_1, \cdots, l_t) \propto \sum_{z_{t+1}} \Phi_{z_{t+1}, l_{t+1}} \sum_{z_t} \Theta_{z_t, z_{t+1}} \alpha(z_t)$$

4 Experiments

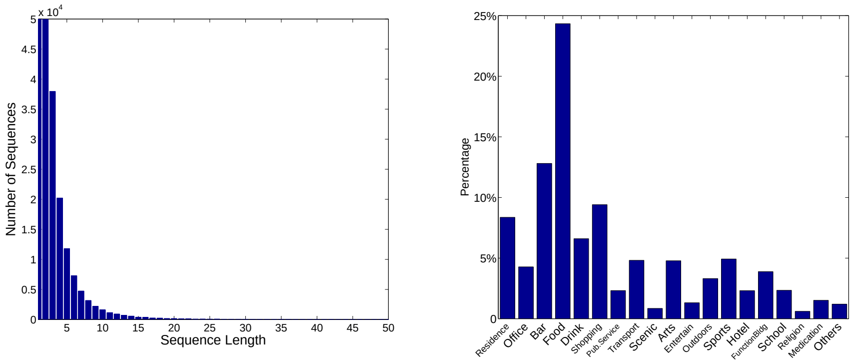
In order to demonstrate the performance of TUTM, the following two experiments are carried out. Firstly, as the main concern of this work, the topical transition probabilities are inferred to discover the semantic mobility patterns of the check-in data. The categorical label is elaborately used to show the conformance between the transition semantics and common sense. Secondly, we assess the next-place prediction accuracy of TUTM, and compare it with some relevant benchmark models to exhibit the feasibility and effectiveness of our model.

4.1 Dataset Description and Analysis

The dataset used in this paper is contributed by Cheng [8], which contains check-in data crawled from the location sharing status on Twitter . The original dataset is composed by 22 million check-ins by over 220 thousand users, each of which is represented as {*UserId, CheckInId, Longitude, Latitude, Timestamp, TwitterContent*}.

We only extract those check-ins occurred in New York City from the above mentioned dataset for our experiments. The total number of check-ins samples is already 840,046, with 12,623 users and over 50,000 venues involved. This data subset, without loss of generality, should be dense enough to cover most of the transition patterns in urban area. Through querying the Foursquare API , we firstly map the check-in's geo-coordinates and twitter contents onto unique predefined Foursquare Venue Entity, and meanwhile download the corresponding category labels of every venue for the later semantic analysis, which categorizes the venues into 19 pre-defined classes. Last but not least, we reorganize the user and theirs check-in points set into consecutive venue sequences according to the timestamps. Specifically, we group every user's check-ins per a given idle time span threshold.

As shown in Fig.2, we conduct a preliminary analysis on the dataset. The histogram of sequence length shown in Fig.2(a) prominently reflects that the distribution of sequence length apparently conforms the long tail effect. Thereby, the number of check-ins included in each sequence is highly heterogeneous: few



(a) Length of Sequences (b) Proportion of Venue Categories

Fig. 2. Preliminary Analysis on Dataset

sequences have large number of check-ins, and more than 50% of sequences have less than 5 check-ins, which causes the average length of sequence reaching down to 5. The proportion of venue categories shown in Fig.2(b) illustrates that the use frequency of location-based services varies in different type of venues. We can clearly observe that check-ins are uploaded more in the places of restaurants or bars, while seldom people update the location status in religious places. This observation is quite consistent with common sense, therefore, we can consider the category distribution as the empirical distribution over topics, which is helpful for topical semantic analysis.

4.2 Topical Transition Analysis

In order to quantitatively demonstrate the TUTM's ability of identifying the semantics transition patterns, we resort to the prepared venues categories information and align them to the Markov transition matrix parameters Θ , so as to manifest our results as intuitively as possible. To this end, firstly, the topic number of TUTM is set to 19, the same as the number of venues' categories. Secondly, we assume that the venues categories' label naturally exhibits the desired topical semantics, and hence, we initialize the starting topic distribution θ_0 with the empirical distribution shown by Fig.2(b). Thirdly, similarly according to the category information, we create a topic-location concurrence binary matrix and normalize it as per topic rows, so as to use it as prior for the model parameter Φ. Thus after the EM algorithm converges, the Markov transition matrix Θ and final venues distribution Φ will be fitted under these configurations.

Fig.3 visualize the topic transition matrix, i.e. Θ learned by the above procedure, in a gray-scale image. Every transition in the matrix is presented as the source being on the vertical axis and the destination on the horizontal axis. And, the darker the matrix cell is, the higher the transition probability is. Furthermore, it is prominently observed that the self-transition probabilities on the

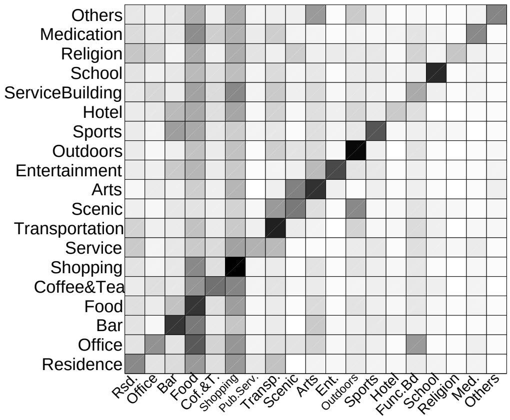


Fig. 3. Topical Transition Matrix

diagonal are generally higher than the probabilities off the diagonal, which is obviously conformed to realities. Because the two adjacent locations are usually generated under the same user intention. We will detail the inferred mobility patterns displayed by this matrix in two aspects: 1) self-transitions; 2) inter-topic transitions.

Table 1. Lowest and Highest Self-Transitions

Topic	**Probability**
Hotel	0.0828
Religion	0.0901
PublicService	0.1187
FunctionalBuilding	0.1338
Office	0.1678
Medication	0.1792
Residence	0.1836

Topic	**Probability**
Shopping	0.3952
Outdoors	0.3770
Transportation	0.3399
School	0.3296
Arts	0.3164
Bar	0.3127
Food	0.3119

The self-transitions on the diagonal of topic transition matrix, represent that one transits between two locations of the same topic. As shown in Table.1, the self-transition of the Hotels and Religious venues have the lower probabilities. Because one venue from these topics could be enough to accomplish people's intended tasks. For instance, the user who has checked in to a hotel will not visit another hotel next. While the self-transitions of Shopping, Outdoors and Transportation topics are with the higher probabilities. Multiple venues of these topics are combined to reach the users' aims. This fact happens frequently in real life. Taking the Shopping topic as an example, people usually go to several stores to achieve the more opportunities to compare the commodities and make purchase decisions. Moreover, as for the transportation topics, it also make sense that people regularly transform among several stops to reach their destination.

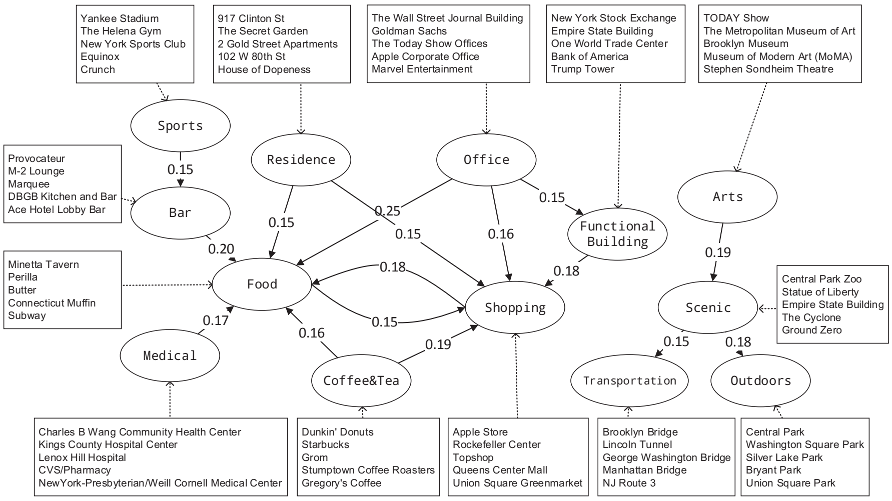


Fig. 4. Inter-Topic Transitions

We then analyze the inter-topic transition probabilities which are shown in Fig.4. In this graph, the transitions with probability less than 0.15 are discarded in order to clearly discover significant patterns. Besides, isolated nodes are discarded too. Notes are attached to each topic z showing the top 5 locations with highest probability according to the location distribution Φ_z. As observed, it is common for users transiting from residences or offices to restaurants, and bars are always attractive after sports games. Moreover, tourists usually take transportation after visiting scenic places. Another fact can be easily observed is that the Food and Shopping have the most in-degrees, indicating that the places under topic Food and Shopping are the most popular places to be transited to. Besides, the hottest places are mined through the location probabilities Φ. As we can see in the attaching notes, there are Apple Store, Rockefeller Center, Topshop under topic Shopping, and Wall Street Journal Building, Goldman Sachs under Office.

The previous analysis demonstrates that our model can reasonably explain the user mobility patterns. Due to the expressiveness of this model, we are convinced that TUTM can better understand user mobility features for destination recommendation and prediction. To further evaluate the usefulness of TUTM, we put the model into practice by predicting the next place.

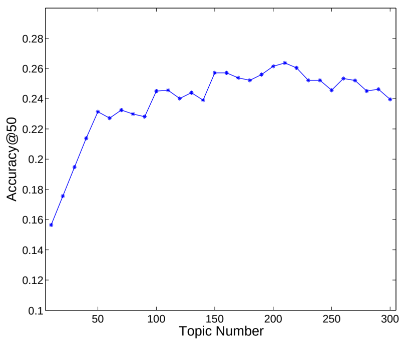
4.3 Next-Place Prediction

Experiments on next-place prediction are carried out based on the aforementioned predicting problem in Section 3.3. We divide the dataset into training and testing set: for each sequence $s \in S$, the former $N_s - 1$ locations $\{l_1, \cdots, l_{N_s-1}\}$ become training sequence, and the last location l_{N_s} to be predicted composes the testing set. Due to the large quantity of the potential target locations in L, the ranking performance cannot be well reflected by the measurement of exact prediction accuracy. In this case, we assess the performance with prediction metric $Accuracy@N$ used in [18],

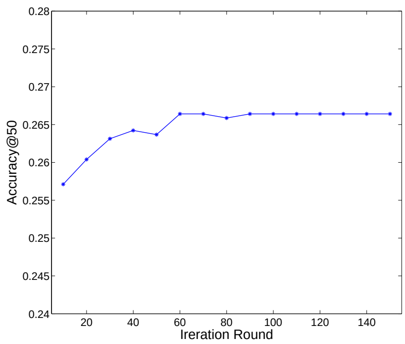
$$Accuracy@N = \frac{|Hit\ Sequences@N|}{|S|} \tag{4}$$

where N denotes the number of candidates chosen with the highest ranking. $Hit\ Sequence@N$ represents the number of sequences with successful prediction, in which the test venue is ranked within top N.

Before training, model parameters are initialized as random probabilities. We first try to discover how performance trends with increasing topic number and iteration round. Fig.5(a) shows that the performance improves overall during the topic number grows until it gets 210. As the influence of iteration round shown in Fig.5(b), the EM algorithm converges approximately at round 100, and the performance will not improve any longer.



(a) Topic Number Selection



(b) Iteration Round Impact

Fig. 5. Parameter Selection

Moreover, in order to manifest the capability of TUTM on predicting or recommending next locations, we compare our model with the following baseline models:

- **Popularity** (PopRec): The popularity recommendation method assume the popular places will be more likely to be visited, thus it utilizes the frequency of location occurrence for predicting, i.e. $p(l_{N+1}|l_1, \cdots, l_N) \propto Count(l_{N+1})$. This approach is commonly used in recommendation field.
- **Markov Model** (MM): Markov Model consider the most likely transitions from current place for prediction. Thus locations are ranked based on the transition probability, i.e. $p(l_{N+1}|l_1, \cdots, l_N) \propto Count(l_N, l_{N+1})$.

– **Latent Dirichlet Allocation** (LDA): As a topic model, LDA extracts the topics to model user preferences to locations. LDA is applied for prediction using user's preference extracting, which is presented as $p(l_{N+1}|l_1,\cdots,l_N) \propto \sum_{z_{N+1}} p(l_{N+1}|z_{N+1})p(z_{N+1}|U)$.
– **Hidden Markov Model** (HMM): HMM has been used for next-place prediction in [20]. HMM predicts the next location by inferring the most likely hidden states of z_{N+1} and sum up the products with the emission probability $p(l_{N+1}|l_1,\cdots,l_N) \propto \sum_{z_{N+1}} p(l_{N+1}|z_{N+1})p(z_{N+1}|l_1,\cdots,l_N)$

PopRec and MM are statistical methods to deterministically calculate the occurrence. As for probabilistic models LDA and HMM, we tune the parameters by selecting their best performance respectively, which are 50 topics for LDA, and 150 hidden states for HMM.

Table 2. Prediction Accuracy

	Acc@10	Acc@50	Acc@100
Pop.Rec	0.0393	0.0973	0.1586
MM	0.1389	0.1734	0.1772
HMM	0.1001	0.2013	0.2598
LDA	0.0979	0.2144	0.2795
TUTM	**0.1493**	**0.2636**	**0.3183**

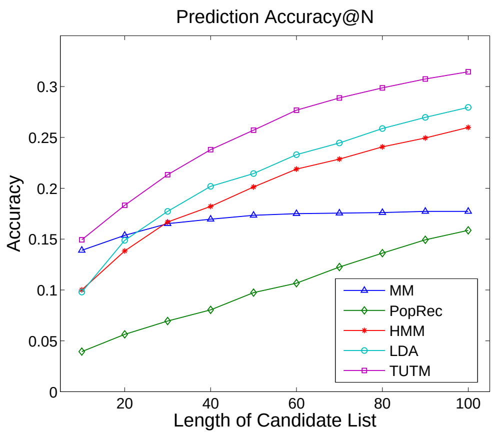


Fig. 6. Accuracy Curves of Different Models

Our model is compared with the baseline models in terms of prediction accuracy specified in Eq.4. As shown in Fig.6, the curves of different approaches demonstrates that our model outperforms the others, and the details are presented in Table.2.

We can observe from the result that the Popularity method perform worst, because it has badly reduced the inter-user diversity. Despite of the high score on top-10, MarkovModel suffers slow growth on accuracy during the increase of candidate set size. It reveals the limits of statistical transition patterns, since only those explicit transitions are recorded. However, TUTM can model the topical transition patterns, which are capable to mine probabilistic relations between locations.

As for the probabilistic models, HMM similarly brings in the transition between hidden states, but the transitions are independent to each other. This deficiency is addressed in TUTM by introducing an unified prior Dirichlet distribution to each transition probabilities. Therefore, as we can observe from the results, an improvement of 24% over HMM has been attained by TUTM on $Acc@50$. The accuracy score of LDA is the closest to our model, since it considers the user's preferences to locations which is widely utilized for recommendation. TUTM improve 20% over LDA in terms of $Acc@50$. In general, TUTM achieves the highest prediction accuracy owing to the great expressiveness in modeling the user mobility pattern.

5 Conclusion and Future Work

In this paper, we propose a probabilistic model TUTM, which utilizes topical transition for semantic mobility pattern analysis. This model can be used for topical transition mining and next-place prediction. Evolving from LDA, TUTM is established by adding transitional dependencies between adjacent latent variables. The Expectation Maximization algorithm incorporating with Forward Backward algorithm is also delivered for estimating the model parameters. We conduct two experiments on real-world check-in dataset to demonstrate the effectiveness of TUTM. Specifically, the topical transition probabilities are analyzed to discover the semantic patterns, which demonstrate that TUTM can reasonably explain the user mobility patterns. Also, in comparison with existing approaches, the next-place prediction performance is confirmed.

As for the future work, we are going to interpret the meaning of the discovered semantics. Since the latent variables corresponding to each location always seem inexplicable, the meaning of them are speculative. In that case, currently we can only temporarily regard them as the attributes or categories of locations. Therefore, reasonably explaining the spatial semantics under the topical transitions are getting on the agenda. Moreover, we are planning to further study the quality of next-place prediction in our model. Currently, the predictive probabilities of our model generated only from the spatial information, without considering the detailed temporal factors. Therefore, we are going to incorporating more context information, such as the distance, time, and even hybrid factors. A more complex model will be expected with more context factors considered. Meanwhile, the social factors are also supposed to be taken into consideration as many existing studies support.

Acknowledgments. This work was supported partly by National Natural Science Foundation of China (No. 61300070, No. 61103031), partly by China 863 program (No. 2012AA011203), China 973 program (No. 2014CB340305), partly by the State Key Lab for Software Development Environment (SKLSDE-2013ZX-16), partly by A Foundation for the Author of National Excellent Doctoral Dissertation of PR China(No. 201159) and partly by Program for New Century Excellent Talents in University.

References

1. Kurashima, T., Iwata, T., Irie, G., Fujimura, K.: Travel route recommendation using geotags in photo sharing sites. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management. CIKM '10, New York, NY, USA, ACM (2010) 579–588
2. Zheng, Y.T., Zha, Z.J., Chua, T.S.: Mining travel patterns from geotagged photos. ACM Trans. Intell. Syst. Technol. **3**(3) (May 2012) 56:1–56:18
3. Li, Z., Wang, J., Han, J.: Mining event periodicity from incomplete observations. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '12, New York, NY, USA, ACM (2012) 444–452
4. Preotiuc-Pietro, D., Cohn, T.: Mining user behaviours: a study of check-in patterns in location based social networks. In Davis, H.C., Halpin, H., Pentland, A., Bernstein, M., Adamic, L.A., eds.: WebSci, ACM (2013) 306–315
5. Monreale, A., Pinelli, F., Trasarti, R., Giannotti, F.: Wherenext: A location predictor on trajectory pattern mining. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '09, New York, NY, USA, ACM (2009) 637–646
6. Yu, K., Zhang, B., Zhu, H., Cao, H., Tian, J.: Towards personalized context-aware recommendation by mining context logs through topic models. In Tan, P.N., Chawla, S., Ho, C.K., Bailey, J., eds.: PAKDD (1). Volume 7301 of Lecture Notes in Computer Science., Springer (2012) 431–443
7. Liu, Q., Ge, Y., Li, Z., Chen, E., Xiong, H.: Personalized travel package recommendation. In: Proceedings of the 2011 IEEE 11th International Conference on Data Mining. ICDM '11, Washington, DC, USA, IEEE Computer Society (2011) 407–416
8. Cheng, Z., Caverlee, J., Lee, K., Sui, D.Z.: Exploring millions of footprints in location sharing services. In Adamic, L.A., Baeza-Yates, R.A., Counts, S., eds.: ICWSM, The AAAI Press (2011)
9. Zheng, Y., Zhang, L., Xie, X., Ma, W.Y.: Mining interesting locations and travel sequences from gps trajectories. In: Proceedings of the 18th International Conference on World Wide Web. WWW '09, New York, NY, USA, ACM (2009) 791–800
10. Arase, Y., Xie, X., Hara, T., Nishio, S.: Mining people's trips from large scale geotagged photos. In: Proceedings of the International Conference on Multimedia. MM '10, New York, NY, USA, ACM (2010) 133–142
11. Bayir, M.A., Eagle, N., Demirbas, M.: Discovering spatiotemporal mobility profiles of cellphone users. In: in Proceedings of the 10th IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM 2009. (2009) 1–9

12. Ganti, R., Srivatsa, M., Ranganathan, A., Han, J.: Inferring human mobility patterns from taxicab location traces. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing. UbiComp '13, New York, NY, USA, ACM (2013) 459–468
13. Scellato, S., Noulas, A., Lambiotte, R., Mascolo, C.: Socio-spatial properties of online location-based social networks. In Adamic, L.A., Baeza-Yates, R.A., Counts, S., eds.: ICWSM, The AAAI Press (2011)
14. Eagle, N., Pentland, A.S., Lazer, D.: Inferring friendship network structure by using mobile phone data. Proceedings of the National Academy of Sciences **106**(36) (2009) 15274–15278
15. Xiao, X., Zheng, Y., Luo, Q., Xie, X.: Finding similar users using category-based location history. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems. GIS '10, New York, NY, USA, ACM (2010) 442–445
16. Backstrom, L., Sun, E., Marlow, C.: Find me if you can: Improving geographical prediction with social and spatial proximity. In: Proceedings of the 19th International Conference on World Wide Web. WWW '10, New York, NY, USA, ACM (2010) 61–70
17. Lu, X., Wang, C., Yang, J.M., Pang, Y., Zhang, L.: Photo2trip: Generating travel routes from geo-tagged photos for trip planning. In: Proceedings of the International Conference on Multimedia. MM '10, New York, NY, USA, ACM (2010) 143–152
18. Noulas, A., Scellato, S., Lathia, N., Mascolo, C.: Mining user mobility features for next place prediction in location-based services. In: ICDM. (2012) 1038–1043
19. Baumann, P., Kleiminger, W., Santini, S.: The influence of temporal and spatial features on the performance of next-place prediction algorithms. In: Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing. UbiComp '13, New York, NY, USA, ACM (2013) 449–458
20. Gambs, S., Killijian, M.O., del Prado Cortez, M.N.n.: Next place prediction using mobility markov chains. In: Proceedings of the First Workshop on Measurement, Privacy, and Mobility. MPM '12, New York, NY, USA, ACM (2012) 3:1–3:6