Big Data Driven Hidden Markov Model Based Individual Mobility Prediction at Points of Interest

Qiujian Lv, Yuanyuan Qiao, *Member, IEEE*, Nirwan Ansari, *Fellow, IEEE*, Jun Liu, *Member, IEEE*, and Jie Yang

***Abstract*—With the emergence of smartphones and location-based services, user mobility prediction has become a critical enabler for a wide range of applications, like location-based advertising, early warning systems, and citywide traffic planning. A number of techniques have been proposed to either conduct spatio-temporal mobility prediction or forecast the next-place. However, both produce diverse prediction performance for different users and display poor performance for some users. This paper focuses on investigating the effect of living habits on the models of spatio-temporal prediction and next-place prediction, and selects one from these two models for an individual to achieve effective mobility prediction at users' points of interest. Based on the hidden Markov model (HMM), a spatio-temporal predictor and a next-place predictor are proposed. Living habits are analyzed in terms of entropy, upon which users are clustered into distinct groups. With large-scale factual mobile data captured from a big city, we compare the proposed HMM-based predictors with existing state-of-the-art predictors and apply them to different user groups. The results demonstrate the robust performance of the two proposed mobility predictors, which outperform the state of the art for various user groups.**

***Index Terms*—Big data, cellular data network, hidden Markov model (HMM), next-place prediction, spatio-temporal mobility prediction.**

I. Introduction

With the wide deployment of 3G/4G cellular data networks, we have witnessed the tremendous growth of mobile Internet access worldwide. Users access the Internet anywhere with smart mobile devices via cellular data networks to check emails, browse the Web, chat online, and perform various mobile applications. Meanwhile, there is a great potential for service and network providers to capture big and invaluable data [1], [2], particularly those related to user mobility.

Recently, the location information extracted from cellular data networks has been found extremely significant to study human dynamics [3]–[6]. As compared with other popular location recording methods, like global positioning system (GPS) or call detail records (CDRs), passively collecting location information as users access cellular data networks incurs low energy consumption, covers a wide range and a large number of individuals, and yields fine time granularity [7]. Users are becoming more reluctant to share locations by using GPS, because continuously collecting GPS data may drain mobile devices' energy quickly or make people uncomfortable with respect to the privacy issue [7]. To collect locations, a limited pool of volunteers with similar living habits is selected [8], [9]. In addition, GPS signals may easily become unavailable in indoors or underground environments, and some noisy points are recorded. As for CDRs, they record the identities of the connecting cell towers when mobile devices initiate or receive a call or text message. Yet, they are sparse in time and coarse in space. These disadvantages of GPS and CDRs limit the scope of their applications to study human mobility of a citywide population. Investigating user mobility with large-scale users in a big city enables advanced human-centered mobile applications and empowers a smart city to engage with its citizens more effectively and actively.

Existing literature focuses on two kinds of prediction models: spatio-temporal prediction [8], [10] and next-place prediction [11]. The first one predicts where the user would be at a given time in the future. The other model aims at predicting where a user would visit after leaving the current place. Different individuals display different prediction performances [10], [12] in employing either model. For some users, one prediction model may perform poorly. So, when the model is applied to a practical application scenario, these users receive inevitably many recommendations of uninterested places. Thus, we propose to incorporate the effect of users' living habits to enhance the performance of these two models and flexibly implement mobility predictors between the two prediction models for different individuals.

Moreover, it is very challenging to carry out user mobility prediction from a large amount of location information collected from cellular data networks. First, user locations represented by base station (BS) IDs are passively collected as users access mobile Internet. These collected locations exhibit the phenomenon of oscillation [13], which mixes users' static and mobile states.

Manuscript received March 31, 2016; revised July 27, 2016; accepted September 11, 2016. Date of publication September 20, 2016; date of current version June 16, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61671078 and Grant 61601042, the Fundamental Research Funds for the Central Universities (2015RC11), the Director Foundation Project (2015BKL-NSAC-ZJ-01), 111 Project of China (B08004), and the EU FP7 IRSES MobileCloud Project under Grant 612212. The review of this paper was coordinated by Dr. Y. Song.

Q. Lv, Y. Qiao, J. Liu, and J. Yang are with the Beijing Key Laboratory of Network System Architecture and Convergence and the Beijing Laboratory of Advanced Information Networks, School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: lvqiujian@bupt.edu.cn; yyqiao@bupt.edu.cn; liujun@bupt.edu.cn; janeyang@bupt.edu.cn).

N. Ansari is with the Advanced Networking Laboratory, Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: nirwan.ansari@njit.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2016.2611654

In addition, among these locations, some are visited less often or only sporadically while others are points of interest (POI). These POIs are associated with the semantics of human's latent states, such as home or workplace. Modeling user mobility at POIs can not only improve understanding of general human movement patterns but also support location-based services for practical applications.

To tackle the above challenges, this paper presents an in-depth investigation on the performance of the next-place prediction model and the spatio-temporal prediction model on individuals on a large scale with different living habits. To begin, we cluster adjacent locations to reduce oscillation and identify POIs. Then, we group users based on users' living habits, which are quantified by the randomness of user mobility during different time periods with entropy profiles. In terms of the Markovian property of POI transitions [14], efficient mobility predictors are designed for both next-place prediction and spatio-temporal prediction by leveraging the hidden Markov model (HMM). The performance of predictors on different user groups also indicates that applying different prediction models to users with distinct living habits can achieve better user mobility modeling. Overall, the contributions of this paper can be summarized as follows.

1) We apply the "Leader–Follower algorithm" to cluster locations collected from a cellular data network that helps to minimize the oscillation and identify POIs. In this method, sorting BSs by the number of days that BSs are accessed helps enhance the POI identification. The method has been applied successfully on a large amount of real data and is shown to be effective in analyzing user trajectories extracted from a cellular data networks.
2) We conduct both spatio-temporal mobility prediction and next-place prediction by leveraging HMM. As compared with the existing predictors of NextPlace [10] and the order-2 Markov model [14], our proposed predictors show higher efficiency and can potentially be incorporated into smart life, like preheating the home in anticipation of the owner's arrival or adjusting traffic routes in case of a traffic jam.
3) With available data on a large number of users throughout a city, we study the effectiveness of the two prediction models on different user groups. Given the living habit of a user, a proper predictor from the perspective of spatio-temporal prediction or next-place prediction can be selected to achieve a better user mobility modeling. Particularly, for individuals leading highly mobile lives, next-place prediction shows a significant advantage over spatio-temporal mobility prediction. For the user who lives an orderly life and has a short trace, spatio-temporal prediction performs better.

The rest of the paper is organized as follows. Related works are reviewed in Section II. Section III provides the problem statement for user mobility prediction. In Section IV, we model user mobility as an HMM. Section V further depicts the key enabling technologies for mobility prediction in detail. We evaluate the models in Section VI by implementing a series of tests on voluminous factual data. Finally, our conclusion and future work are presented in Section VII.

II. Related Works

We categorize existing works of understanding user mobility at POI into two parts: identifying POI, and user mobility modeling and prediction.

A. Identifying POI

Identifying POIs from user trajectories is a key task in mining mobility patterns and has been studied extensively. Most prior works attempted to identify POIs by mining individual GPS data [9], [10], [15]–[18] or Wi-Fi beacons [10], [19]. Commonly known methods include time-based clustering [18], [19], density-based clustering [9], [17], and some popular clustering algorithms [19]. The k-means algorithm [19], a popular clustering algorithm, is rather effective provided that the number of clusters is known a *priori*. Actually, researchers would not know how many POIs users have. In addition, though the time-based algorithms [18], [19] are simple and work in an incremental way on mobile devices, it is difficult to discover places that are visited with high frequency but short on dwell time. Detected clusters by using the density-based clustering method [17] exhibit a wide variation in local density. Such clusters can be divided into several smaller clusters, with locations distributed uniformly within each small cluster. Some other studies [20] also identify POIs by using CDRs, which are generated only when a phone engages in a voice call or text messages.

B. User Mobility Modeling and Prediction

Human movement traces, which are collected from real-life human mobility or generated using synthetic mobility models, have been used to explore patterns of user trajectories and yield insight into a variety of issues, such as urban planning, disease spreading, and radio resource optimization [21], [22]. Prevalent synthetic mobility models include the random walk mobility model [23], random waypoint mobility model [24], the Gauss–Markov mobility model [25], and so on. All these synthetic models attempt to mimic the movements of mobile users and simulate their mobility patterns using parametric methods synthetically.

Besides, a number of previous efforts have attempted to model user mobility based on real-life movement traces. To elicit the state of the art, we focus on two kinds of modeling: spatio-temporal modeling and spatial movement modeling. The spatio-temporal modeling aims to discover the prominent daily temporal habits as well as predicting future individual activities [8], [10], [18], [26], [27]. The spatial movement modeling facilitates the next-place prediction [28], [29], residence time prediction [30], [31], life pattern mining in the user traces [9], [28], and identification of departures from mobility routine [32], [33].

Alvarez-Lozano *et al.* [18] proposed a medium-term spatio-temporal prediction model based on HMM. However, this model was tested with only 63 limited users. Zheng and Ni [8] automatically uncovered and quantified spatio-temporal behavior patterns in users' daily lives with 95 users from the MIT Media Lab. They pointed out that users leading highly mobile lives could not be well described by the proposed model. Scellato *et al.* [10] proposed the NextPlace method to predict

spatio-temporal behavior based on the nonlinear time series analysis of arrival time. The method produced different prediction accuracies for various datasets, largely owing to the differences in the number of POIs and the total residence time in POIs of users. Hence, for spatio-temporal prediction, it is meaningful to use a large pool of users to systematically study the effect of living habits on prediction performance.

With regard to next-place prediction by spatial movement modeling, in [34] and [35], HMM for mobile prediction was adopted and the performance with real-world data was tested. Si *et al.* [34] indicated that for a given HMM, the probabilities of observable variables represented by the historical traces decline, as the length of historical traces increase. The method of probability normalization solved the problem at the cost of increased computational complexity. A weak prediction accuracy of 13.85% was obtained with another HMM-based predictor [35]. In [14] and [36], a series of Markov-based models to perform next location prediction has been implemented. Their experimental results revealed that the performance of the order-2 Markov model was better than the other complex predictors [14], and the maximum predictability could be approached with Markov-based models [36]. However, users tested in [14] were limited in a small region of a college, and Si *et al.* [34] adopted only one person as the research object.

Apart from the previous works, we distinguish our work as follows.

1) The dataset we used is real-world control-plane traffic collected from a long-term evolution (LTE) cellular data network, other than global system for mobile communications, universal mobile telecommunications system, GPS, or Wi-Fi. Moreover, test users are distributed in a famous city in Southern China and differ in living habits and working conditions.
2) To identify POIs from the cellular data network, a clustering method based on the Leader–Follower algorithm is developed. Distinguishing different BSs by sorting as well as selecting reasonable self-defined thresholds enhance the identification of POIs.
3) By leveraging HMM, we conduct spatio-temporal prediction and propose a new next-place predictor. We also analyze the living habits of a large pool of users on the basis of entropy and cluster them into distinct groups. We have applied these two prediction models on different user groups and analyzed their performance.

III. Individual Mobility Prediction

For different individuals, the prediction accuracies of spatio-temporal prediction are diverse [8], [10]. In terms of next-place prediction, both the degree of movement randomness [12] and the size of movement history [11] affect prediction performance. Given the diversity in prediction performance of different users, we consider applying different prediction models (spatio-temporal prediction and next-place prediction) to users with distinct living habits. In this way, better prediction performance can be achieved for practical applications.

This section presents the problem of individual mobility prediction by first defining some basic terms, followed by the problem statement.

A. Basic Terms

Spatial trajectories can be generated unintentionally when an individual accesses mobile Internet and moves from one BS to another. These trajectories are represented by sequences of BS IDs with the corresponding transition time. To gain practical insight into user mobility, some basic terms and definitions are given as follows.

Definition 1: Trajectory. A user's trajectory Traj is represented by a sequence of time-stamped locations $\text{Traj} = l_0, l_1, \ldots, l_k$, where $l_i = (x_i, y_i, t_i)$ $(i = 0, 1, \ldots, k)$; (x_i, y_i) corresponds to the latitude and longitude coordinate of the BS, which starts to serve the user at the timestamp t_i ($\forall 0 \leq i < k$, $t_i < t_{i+1}$). $\text{Dist}(l_i, l_j)$ represents the geospatial distance between two locations l_i and l_j and is calculated by using Vincenty's formulae [37] with the latitude and longitude coordinates (x_i, y_i) and (x_j, y_j) of the two locations, respectively.

Definition 2: Place. A place c is a geographical region within a radius threshold of D_r. Each place carries a semantic meaning, like home, work, gym, and market. In a user's trajectory, c is characterized by a set of consecutive locations $L = l_m, l_{m+1}, \ldots, l_n$, where $\forall m < i \leq n$, $\text{Dist}(l_m, l_i) \leq D_r$ and $\text{Dist}(l_m, l_{n+1}) > D_r$.

Definition 3: Points of Interest. A POI is a place c where a person has visited for more than a threshold in terms of the number of days T_r. Otherwise, it is labeled as "non-POI." Section V will detail the method of POI identification from user trajectories, respectively.

Definition 4: Travel Sequence. A travel sequence $\text{TS} = \text{ts}_0, \text{ts}_1, \text{ts}_2, \ldots, \text{ts}_n$ is a sequence of time-stamped places in which all except one are POIs. The "non-POI" is used when the user is at a place that is not a POI.

B. Problem Statement

The structure of our model is illustrated in Fig. 1. The five main parts of the model are POI identification from user trajectories, user clustering, spatio-temporal mobility prediction, next-place prediction, and model selection, respectively.

POI identification: In this component, POIs are extracted by clustering the adjacent locations in user trajectory *Traj*, and then the travel sequence $S = s_0, s_1, s_2, \ldots, s_n$ is obtained.

User clustering: We analyze users' living habits characterized by entropy profiles, by exploiting the randomness of user mobility in different time periods. Then, users are clustered into different groups according to their living habits.

Spatio-temporal mobility prediction: Based on the correlation between time and places, we predict where a user would appear at a specific time in the future.

Next-place prediction: By mining spatial movement patterns, this part aims at predicting the next place a user would visit after leaving the current place.

Model selection: The prediction accuracies of spatio-temporal mobility prediction and next-place prediction are evaluated with regard to distinct user groups. We will verify whether choosing different prediction models for users with distinct living habits can improve the prediction accuracy, especially when one model performs poorly for a certain group.

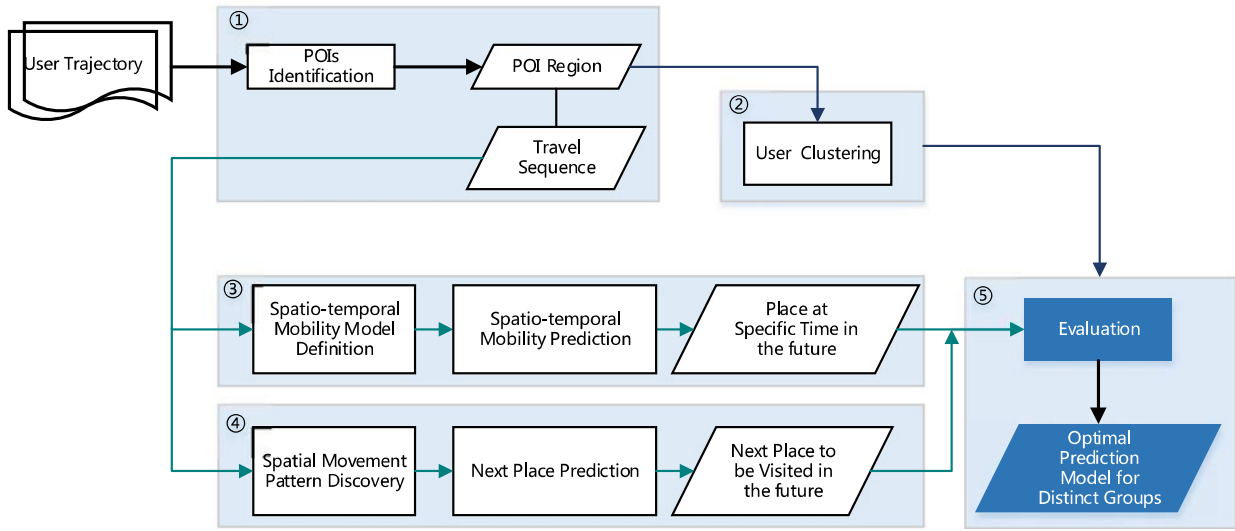


Fig. 1. Structure of our model for user mobility prediction: ① POI identification, ② user clustering, ③ spatio-temporal mobility prediction, ④ next-place prediction, and ⑤ model selection.

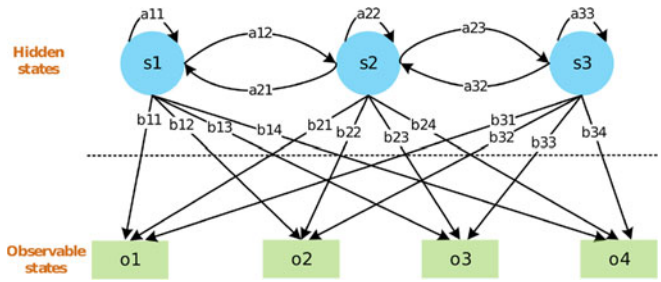


Fig. 2. Example of an HMM.

IV. Modeling User Mobility by HMM

As a user's movement can be characterized by a Markovian stochastic process [14], we model user mobility as an HMM. We shall next build HMM-based predictors for spatio-temporal mobility prediction and next-place prediction.

HMM, as a classic dynamic Bayesian network, is suitable for recognizing temporal patterns of data sequences generated by a Markov process with unobservable (i.e., hidden) states [38]. HMM has two kinds of stochastic variables: state variables (hidden variables) and output variables (observable variables). Fig. 2 shows the general architecture of an instantiated HMM, in which each node represents a random variable. $s(t)$ represents the hidden state S at time t, and $s(t) \in \{s1, s2, s3\}$. $o(t) \in \{o1, o2, o3, o4\}$ is the observable state O at time t. The structure of HMM also implicates two kinds of conditional probabilities: state transition probability $a_{ij}=p(s_{j:t+1}|s_{i:t})$, $1 \le i, j \le 3$ and output probability $b_{ij}=p(o_{i:t}|s_{j:t})$, $1 \le i \le 4, 1 \le j \le 3$. The joint probability distribution of all variables can be simplified as $p(s_{1:T}, o_{1:T}) = \prod_{t=1}^{T} p(s_t|s_{t-1})p(o_t|s_t)$.

$\lambda = \{A, B, \pi\}$ is introduced to characterize HMM, where transition matrix A is an $N \times N$ matrix and $A_{ij} = P(s_{j:t+1}|s_{i:t}), 1 \le i, j \le N$; confusion matrix B is an $N \times M$ matrix and $B_{ij} = P(o_{i:t+1}|s_{j:t}), 1 \le i \le M, 1 \le j \le N$; π is a $1 \times N$ vector and $\pi = [p(s_1), p(s_2), ...p(s_N)]$ (N is the number of hidden states and M is the number of observable states).

In general, HMM associates with several inference problems [38], which are also basically concerned in mobility prediction. One of them is evaluation, which is to efficiently compute the likelihood of an output-sequence $o_{i:t}$ for a particular HMM λ. Applying the principle of dynamic programing, this problem can be handled efficiently by the forward algorithm [38].

The second problem is decoding. It is to identify the most likely sequence of hidden states $s_{i:t}$, for a given output-sequence $o_{i:t}$ and model $\lambda = \{A, B, \pi\}$. This task finds the maximum value of $p(s_{i:t}|\lambda, o_{i:t})$ over all possible hidden state sequences, and is solved efficiently by the Viterbi algorithm [38].

The third problem is HMM parameter learning. The target of learning is utilizing existing data to adjust model parameters so that the resulting parameters $\lambda = \{A, B, \pi\}$ can describe the system structure better. In the process of parameter learning, an iterative algorithm called the Baum–Welch algorithm [38] is used. Each iterative procedure calculates a group of parameters $\lambda^* = \{A, B, \pi\}$ and the likelihood $p(o_{1:T}|\lambda^*)$. The algorithm is considered to have converged if the difference between $p(o_{1:T}|\lambda^*)$ and the likelihood of the previous iteration $p(o_{1:T}|\lambda_{prev})$ is less than a desired threshold, upon which the optimal parameters λ are acquired.

V. Key Methodologies

To address the challenges of user mobility prediction in our framework, we first propose the method of POI identification, followed by users clustering with entropy profiles. Then, HMM-based mobility models are developed for spatio-temporal prediction and next-place prediction, respectively.

A. POI Identification

In this section, we present the "Leader–Follower Clustering" [39] for constructing location clusters to reduce the oscillation

effect and identify POIs from user trajectories. The technique depends on two parameters: a radius threshold D_r and a threshold of the number of days T_r.

First, we calculate the number of days that BSs were accessed (access-day) based on users' trajectories Traj, and sort BSs of every user by "access-day" in descending order. The BS with the most "access-day" ranks first in the returned sorted list $sortedlocList$. Sorting by "access-day" rather than duration or occurrence is meaningful because it decreases the influence of vacations or other locations that were visited only on few days but had relatively long duration or high occurrence. The distinction of BSs by sorting also modestly enhances the leader (centroid) detection of each cluster in step 2.

Second, we cluster locations from the sorted list $sortedlocList$ and return a set containing all "places." The first BS in $sortedlocList$ is the centroid or leader of the first place. The location of following BSs will be compared with the centroids of existing places. If a BS is away from all existing places, it becomes the centroid of a new place. Otherwise, if the BS falls within the threshold radius D_r of an existing place c_i, it is added to c_i as a follower. The centroid of c_i is adjusted to be the average position of all the BSs in place c_i.

The third step is POI determination. A set of POIs is formed by excluding those places with visiting days less than threshold T_r.

After identifying POIs, a user's trajectory is converted into a travel sequence S, which is the input to the subsequent spatio-temporal prediction and next-place prediction.

B. User Clustering

In this section, we distinguish users' living habits by clustering. To obtain knowledge of users' living habits, the randomness of user mobility of different time segments is exploited in terms of users' place recurrence entropy. The bigger the entropy value is, the more uncertain a user appears at a specific place in this time period. If a person has high entropy values all day long, he/she is considered to be a ranger and wanders in the city randomly the whole day. In contrast, a person appearing at a limited number of places and having low entropy values is more likely to be a worker with a regular daily routine.

The idea of user clustering on place recurrence entropy has been proposed in a recent work [40] by our research team. It uses the k-means method to cluster entropy vectors $E = [e_0, \ldots e_i]$ of all users. The variable e_i represents the entropy value of each segment $i (0 \leq i \leq 23)$, which is 1 h long with the first hour and last hour of the day indexed by 0 and 23, respectively. The vector of each user indicates the user's living habits, and the centroid of each cluster represents the typical features of each group. We define the entropy value of each time segment e_i as

$$e_i = -\sum_{k=1}^{n} p_i(c_k) \log_b p_i(c_k)$$

where n is the number of places a user has visited in the ith time segment of all days, b is a constant, c_k represents a different place, and $p_i(c_k)$ is the probability of the user staying at place c_k in the ith time segment that is given as

$$p_i(c_k) = \frac{T_i(c_k)}{T_i}$$

where $T_i(c_k)$ is the total time duration of the user staying in location c_k in time segment i, and T_i is the total time duration of the user staying in time segment i.

In Section VI, we discuss how prediction accuracies of spatio-temporal prediction and next-place prediction depend on users' living habits.

TABLE I
USER TRAVEL SEQUENCE AS A VECTOR v_i

Time(O)	00:00	...	8:00	8:30	9:00	...	23:00	23:30
Place(S)	1	...	1	0	3	...	1	1

O and S denote the observable states and hidden states of HMM, respectively.

C. HMM-Based Spatio-Temporal Prediction

Intuitively, users' current places correlate with time. For example, a user is more likely to be at home in the evening and at workplace during the day. Spatio-temporal prediction investigates the correlation between time and places. To do so, we first divide day i into 48 time slots of 30 min and convert the travel sequence of the day into a vector v_i, as shown in Table I. Each time slot contains an index (starting from 0) corresponding to a place where the user has spent the most time. Zero represents the user being at non-POI for that time slot. A user's historical travel sequence in the observed n days can then be represented by a set of vectors $V = \{v_1, v_2, \ldots, v_n\}$. Afterward, we will depict the definition along with the implementation of the spatio-temporal prediction model based on HMM.

1) Model Definition: By leveraging HMM to model user mobility, the spatio-temporal mobility model is defined as follows.

Hidden states S: These are defined by POIs in users' travel sequences, as well as another hidden state representing all the non-POIs. They are composed of the elements of the second row in Table I. We denote the ith element of the hidden states S as s_i.

Observable states O: These are the 48 time slots of each day and correspond to the first row in Table I. o_i denotes the ith element of the observable states O.

Vector π: Each element of the vector represents the probability a user appears at a given hidden state $P(s_i)$.

Transition matrix A: It represents the transition probabilities of different hidden states, i.e., $A_{\text{ij}} = P(s_{j:t+1}|s_{i:t})$.

Confusion matrix B: It represents the probabilities of time slots in which a user is at different hidden states, i.e., $B_{\text{ij}} = P(o_j|s_i)$.

Fig. 3 shows an example of an instantiated HMM for spatio-temporal mobility modeling that has five time slots. The hidden states are represented by four POIs.

By using Bayes' formula, the $\{A, B, \pi\}$ modeling mobility pattern of each user is generated from his/her historical travel

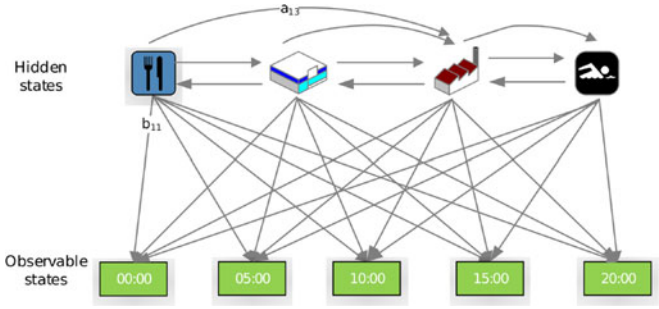


Fig. 3. HMM representing POIs and their relationships with time slots of a day.

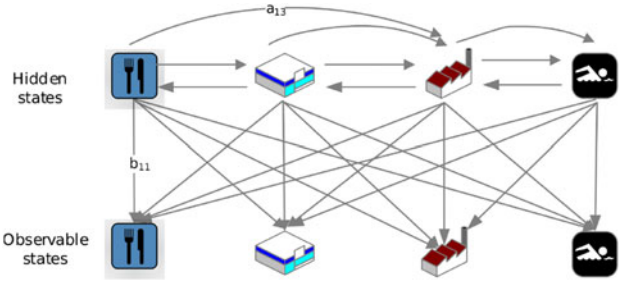


Fig. 4. HMM representing doubly stochastic process of the transition between places.

sequence vectors V. The vector π is generated as follows:

$$P(s_i) = \frac{N(s_i, V)}{|V|}$$

where $N(s_i, V)$ denotes the number of times element s_i occurs in V, and $|V|$ is the length of vectors V. Each element of the transition matrix A is derived from

$$P(s_{j:t+1}|s_{i:t}) = \frac{N(s_{i:t}s_{j:t+1}, V)}{N(s_{i:t}, V)}$$

where $s_{i:t}s_{j:t+1}$ stands for s_i and s_j being in time slots t and $t+1$, respectively. The confusion matrix B is similarly derived from

$$P(o_j|s_i) = \frac{N(s_{i:o_j}, V)}{N(s_i, V)}$$

where $s_{i:o_j}$ means the corresponding hidden state in time slot (observable state) o_j is s_i.

2) Prediction: The previous section has defined the HMM-based spatio-temporal mobility model. In this section, we describe how to use such a mobility model for mobility prediction with historical travel sequence vectors.

Prediction is to discover the most probable place (hidden state) at each time slot (observable state) in the next day for a given HMM $\lambda = \{A, B, \pi\}$; this is an HMM decoding problem. When predicting all the places to be visited in the day d, the HMM $\lambda = \{A, B, \pi\}$ is updated periodically at the end of day $d-1$ given the historical travel sequence vectors from day 1 to day $d-1$.

3) Computation Complexity: The computational complexity of the model is considered from two aspects: a) building the HMM model by using Bayes' formula; and b) prediction based on the Viterbi algorithm.

The complexity of building the HMM model for each individual is $O(n_s^2 + n_s \times n_o)$, where n_s is the number of hidden states, and n_o denotes the number of observable states. Note that this process is not iterative and thus incurs a rather low computational complexity.

Based on the Viterbi algorithm, the complexity for prediction is just $O(n_s^2)$. As shown in the experiments described in Section VI-B, each user has a limited number of POIs, and thus the required computation for this step is low.

In addition, the Viterbi algorithm can be approximately expressed in MapReduce [41]. So, the proposed model can be parallelized and is thus scalable and applicable to a large dataset.

D. HMM-Based Next-Place Prediction

In this section, we propose a novel next-place prediction method based on HMM. It first explores the spatial movement pattern, upon which next-place prediction is carried out.

1) Spatial Movement Pattern Discovery: To quantify the spatial movement pattern, we first apply HMM to establish the likelihood $prob$ of subsequences in travel sequences with length k varied from 2 to 10, as our group has discovered that the length of frequent subsequences is mostly no more than 10 [42]. For two subsequences of the same length, the one with a higher likelihood $prob$ means a higher occurrence rate. Fig. 4 shows the architecture of an instance of HMM for spatial movement mobility modeling.

Hidden states S: These are defined by the POIs in a user's travel sequence, as well as another hidden state representing the non-POI.

Observable states O: These are defined as the same as hidden states.

The vector π, transition matrix A, and confusion matrix B are obtained through the process of parameter learning, which is defined as follows.

a) Generating an $P \times Q$ matrix for parameter learning from travel sequence $\mathrm{TS} = \mathrm{ts}_1, \mathrm{ts}_2, \ldots, \mathrm{ts}_P$. The number of rows P equals the length of the travel sequence TS, and the number of columns Q is ten representing the maximum allowable length (k) of subsequences. Elements in row i correspond to the substring $(\mathrm{ts}_{\mathrm{rem}(i,P)}, \mathrm{ts}_{\mathrm{rem}((i+1),P)}, \ldots, \mathrm{ts}_{\mathrm{rem}((i+9),P)})$ of TS. The $\mathrm{rem}(i, P)$ returns the remainder of the division of i by P.

b) Deriving parameter $\{\pi, A, B\}$. The Baum–Welch algorithm takes the matrix produced in step a) as input, and obtains the optimal parameters $\{\pi, A, B\}$ upon convergence.

Then, given the defined HMM $\lambda = \{A, B, \pi\}$, the likelihood $prob$ of subsequences $\mathrm{ts}_i, \mathrm{ts}_{i+1}, \ldots, \mathrm{ts}_{i+k-1} (0 < i \leq P, 2 \leq k \leq 10)$ can be effectively calculated by the forward algorithm.

2) Prediction: In the beginning, a hash table $seqPattern < key, value >$ is created to cache all probable next-places and the corresponding $prob$ following the previous $k-1$ $(2 \leq k \leq 10)$

TABLE II
EXAMPLE OF THE key–$value$ PAIRS IN $seqPattern$

key (first k-1 places of a subsequence)	$value$ (List{(kth place, $prob$ of the subsequence)})
(1,2,3,4)	(1,0.39); **(5,0.90)**;
(1,2,3,4,5,3,4,2)	**(3,0.85)**; (5,0.75);

places. An example of the key–$value$ pairs is illustrated in Table II. It uses the first $k-1$ elements of a subsequence as hash key. The hash $value$ is a linked list, whose node is composed of the kth element and the $prob$ of the subsequence.

Then, the predicted next-place $\widehat{ts}$ is derived by looking up the $seqPattern$ based on the previous place sequence $[\text{ts}_1, \text{ts}_2, \ldots, \text{ts}_l](1 \leq l \leq 9)$. If available, the length of the initial previous place sequence is up to 9, which equals to the maximum length of sequences in the hash key. In case that the previous place sequence does not exist in $seqPattern$, the length decreases. The pseudocode is given in Algorithm 1, detailed below.

a) If the length-(l) sequence $[\text{ts}_1, \text{ts}_2, \ldots, \text{ts}_l]$ is found in the hash table, the probable next-place will be obtained from the returned linked list. Otherwise, a length-($l-1$) sequence will be generated repeatedly by removing the first element to perform lookup operation until a linked list is returned.

b) If multiple possible places are contained in the returned linked list, the place with the highest $prob$ is selected as the predicted next-place $\widehat{ts}$.

3) Computation Complexity: The computational complexity of the model is considered from three aspects:

a) deriving parameters $\{\pi, A, B\}$ by using the Baum–Welch algorithm;
b) inferring the likelihood of subsequences;
c) prediction.

The former two aspects occur in the spatial movement pattern discovery.

In deriving parameters $\{\pi, A, B\}$ by using the Baum–Welch algorithm, the complexity of each iteration is $O(QPn_s^2)$, where n_s is the number of hidden states. The likelihood of subsequences is inferred by using the forward algorithm with the complexity of $O(kn_s^2)$. Hence, to enhance efficiency, the spatial movement pattern discovery step should not be performed after each user's activity but rather be launched offline and periodically repeated, e.g., once a day. It is worth mentioning that the Baum–Welch algorithm and forward algorithm can be implemented in MapReduce [41]. So, the proposed mode can be parallelized to handle large-scale datasets.

As for the process of prediction, the complexity of the prediction process by applying hash table is only $O(l)$. The prediction task with such a low complexity can thus be performed online.

Algorithm 1: Pext-Place Prediction

Require: $[\text{ts}_1, \text{ts}_2, \ldots, \text{ts}_l]$ sequence of previous l places, $seqPattern < key, value >$ hash table of subsequence pattern
Ensure: $\widehat{ts}$ predicted next-place
1: **function** NextPlacePrediction $[\text{ts}_1, \text{ts}_2, \ldots, \text{ts}_l]$, $seqPattern$
2: $len = l$; $placeSeq = [\text{ts}_1, \text{ts}_2, \ldots, \text{ts}_l]$;
3: **while** len **do**
4: key=$placeSeq$
5: **if** key in $seqPattern$ **then**
6: Linkedlist $[possiblePlace, prob] = seqPattern$. get (key);
7: $\widehat{ts}$ = SelectPlaceofMaxProb(Linkedlist $[possiblePlace, prob]$);
8: **break;**
9: **else**
10: $len = len - 1$; $placeSeq = [\text{ts}_{l-\text{len}+1}, \ldots, \text{ts}_{l-1}, \text{ts}_l]$;
11: **end if**
12: **end while**
13: **return** $\widehat{ts}$
14: **end function**

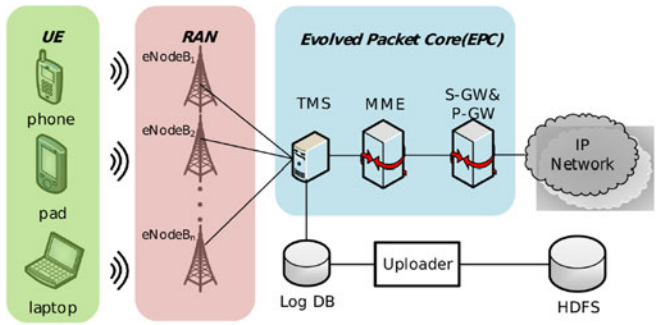


Fig. 5. LTE mobile network with data capture devices.

VI. EXPERIMENTAL EVALUATION

In this section, we give an overview of the dataset before describing the results of POI identification. Then, we exhibit the entropy profiles of clustered groups. In the end, the effectiveness of the next-place predictor and spatio-temporal predictor, together with their performance on different user groups, is presented.

A. Dataset

We use the real mobile data collected by the commercially deployed traffic monitoring system (TMS), which has been developed by our research team. As shown in Fig. 5, the TMS is deployed between evolved Node Bs (eNodeBs) and mobility management entity of an LTE mobile network. It analyzes LTE control-plane traffic generated by user equipment, such as mobile phone and tablet. A sequence of time-stamped records of signaling procedures [43], which contain current service

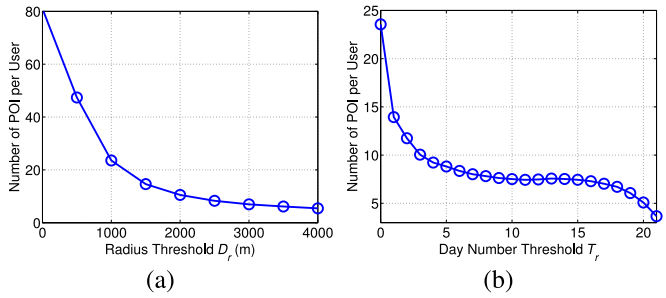


Fig. 6. Average number of POIs per user versus (a) D_r w.r.t $T_r = 0$ and (b) T_r w.r.t $D_r = 1000$ m.

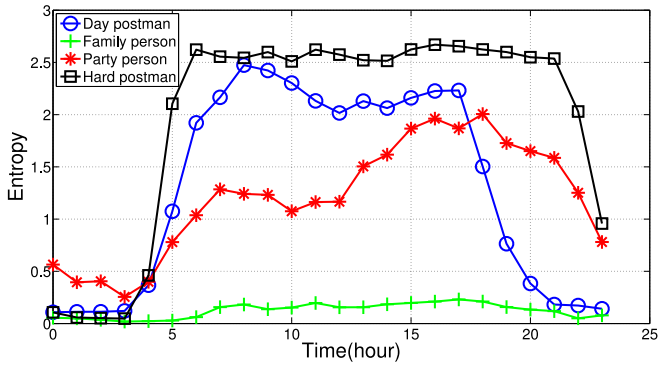


Fig. 7. Entropy profiles of four different groups of users. The *x*-axis represents the 24 time slots of a day while the *y*-axis reports the entropy values of the centroid of each group.

eNodeB ID, signaling procedure code, user's anonymized ID, etc., are produced. The data are stored in a log database and periodically uploaded by an uploader to Hadoop distributed file system [1].

By leveraging self-developed MapReduce programs that run on a distributed computing platform [1], we extract users' trajectories from all mobility related records, including not only path switch and handover procedures during data transmission but also normal and 12-min periodic tracking area update procedures in the idle period without data network activity [43]. Owing to the dominance of cellular data traffic as well as the denser locations extracted from control-plane traffic than GTP-U (GPRS Tunneling Protocol for User Planes) traffic, our dataset provides more sufficient location information than previous works from the view of cellular networks [3], [44].

Twenty-two days, from October 10, 2013 to October 31, 2013, of data consisting of over 3000 mobile phone users moving around a city with 1613 eNodeBs were collected. Overall, we have collected 37 570 167 records of mobility-related signaling procedures. Specifically, the dataset from October 10 to October 24 is treated as the *training set* to build models and the rest is the *testing set*. These records consist of a broad range of users' outdoor movements, including life routines (like going home and going to work), some entertainment activities (such as shopping, sightseeing, dining, or hiking), and so on. These voluminous records can thus be convincingly used to validate movement prediction methods under the big data scenario with users sharing different living habits.

B. POI Identification

As described in Section V-A, the POI identification process takes user trajectory *Traj*, radius threshold D_r, and day number threshold T_r as input. To extract POIs correctly, we select suitable thresholds of D_r and T_r by investigating how the average number of POIs changes as a function of the threshold itself.

We first study the effect of D_r with fixed $T_r = 0$. Fig. 6(a) reports that the average number of clusters decreases as D_r (in meters) increases. To find the optimal threshold, a "knee" is found in the curve, where a significant change in the slope of the graph is observed [16], as the knee signifies the radius just before the number of locations begins to converge to the number of places. The threshold of D_r is set to 1000 m.

Next, we tune T_r with fixed $D_r = 1000$ m. Fig. 6(b) shows the average number of POIs declines rapidly in two ranges ($T_r = 0$ to $T_r = 7$ and $T_r = 16$ to $T_r = 20$), and two knees are observed in the curve. As T_r increases from 0 to 7, the curve drops sharply and the number of POIs converges to a stable level. Then, the second decline from 16 to 20 is for excessive deletes. For example, a woman regularly goes to the gym on Monday other than every day. If the threshold is set at the second knee, the actual POIs of the gym will be removed, and the remaining POIs are places visited almost every day. Hence, we set the threshold value of T_r to 7. In this way, all possible POIs, which are visited more than 31.8% of the overall 22 days, are identified.

The average number of POIs is 7.5 among all users. We denote C_1 as the place that the user stays for the longest time, C_2 as the place with the second longest duration, and so on. On average, users spend 51.35% of the whole observation period at C_1, 14.86% at C_2, and 7% at C_3. The resident time of the top two places (usually home and workplace) accounts for 66.22%. Similar to [45], users spend most at a limited number of places (the findings of [45] indicate that users spend 56% of their time at C_1, 14% at C_2, and 7% at C_3.). The experimental result shows that the mobility trajectories applied in this paper have similar characteristics with those from others' works. Moreover, our method of identifying POIs is effective in analyzing user trajectories extracted from the cellular data network.

C. User Clustering

According to users' place recurrence entropy of different time segments, test users are clustered into four groups empirically to distinguish users' living habits. The labels shown in Fig. 7 intuitively describe the properties of the four groups. The properties and the proportion of users per cluster are as follows.

1) Day postman (27%): During the day, he/she moves around the city and spends time at various places. At about 6 p.m., he/she commutes back home and the entropy value decreases.
2) Family person (34%): All day, his/her entropy is low in that he/she visits few fixed places. He/she is inferred to lead a regular life pattern in his/her daily life.

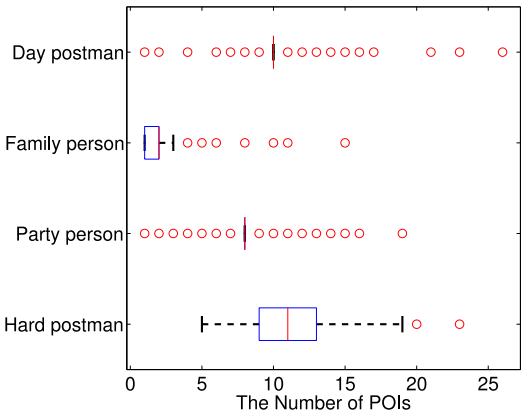


Fig. 8. Distribution of the numbers of POIs w.r.t. different user groups.

3) Party person (17%): Relatively, he/she spends afternoon and night hours at various places.
4) Hard postman (20%): He/she moves around the city and spends time at various places from early morning until late at night.

Fig. 8 exhibits the distribution of the number of POIs with regard to different user groups. Users with distinct living habits show different distribution patterns. A "Family person" leading a regular life pattern has few POIs. In contrast, individuals with the label of "Hard postman" possess the most POIs, ranging from 9 to 13. The users in the group of "Day postman," who are active in the morning, have relatively more POIs than individuals in the group of "Party person" moving around in the afternoon and night hours.

D. Performance of Prediction Models

In this section, we introduce the performance metrics for evaluating the two prediction models. We then provide a comparative assessment of our proposed models with existing location prediction methods. Finally, we study the effects of the two models on different user groups.

1) Evaluation Metrics and Baselines: To quantitatively evaluate the two prediction models, we consider the metric *prediction accuracy*:

a) HMM-based spatio-temporal mobility prediction model (HMM-ST): The metric *prediction accuracy* represents the ratio between the number of correct predictions and the number of all attempted predictions in the *testing set* of a user. Correctness is defined as follows: If we predict user i will be at place c at time T, the prediction is considered correct if the user is at c at any time during the interval $[T-\theta, T+\theta]$, where θ is the error margin.
b) HMM-based next-place prediction model (HMM-NEXT): the metric *prediction accuracy* stands for the proportion of correct predictions to all attempted predictions in the *testing set* of a user.

Comparison methods. The HMM-ST and HMM-NEXT are, respectively, compared with the following existing state-of-the-art predictors:

TABLE III
TIME CONSUMPTION OF MODELS IN PREDICTION

Type	Prediction Model	Time Consumption
Spatio-temporal prediction	NextPlace (**baseline**)	$O(P)$ P is the length of a user's all historical travel sequence
Spatio-temporal prediction	HMM-ST	$O(n_s^2)$ n_s is the number of hidden states (POIs)
Next-place prediction	$O(2)$-Markov (**baseline**)	$O(l)$ l is the length of a user's sequence of previous places applied to prediction
Next-place prediction	HMM-NEXT	$O(l)$

a) NextPlace [10]: This model predicts spatio-temporal behavior based on the nonlinear time series analysis of the arrival and residence times of users. The time series is embedded in an m-dimensional space. Performance evaluation indicates that the model with dimension $m = 3$ performs better.
b) Order-2 Markov model with fallback ($O(2)$-Markov) [14]: The trajectory of each individual is modeled as a Markov chain of order 2 when conducting next-place prediction. Moreover, the fallback technique, which uses the result of the Order-1 model when it encounters an unknown context, is employed to the normal Order-2 model.

To the best of our knowledge, NextPlace's performance has been demonstrated to be the best in the literature in terms of spatio-temporal mobility prediction. Moreover, extensive experiments [14], [36], [40] indicate that $O(2)$-Markov performs better than other complex predictors in terms of predicting the next-place. Hence, the two most effective models, NextPlace and $O(2)$-Markov, are selected as baselines in comparative evaluation: the HMM-ST and HMM-NEXT will be, respectively, compared against NextPlace and $O(2)$-Markov.

Moreover, our proposed models and the baseline models are comparable in terms of time consumption when conducting prediction, as shown in Table III. They all have the advantage of short time-consumption and satisfy the requirement of real-time prediction.

2) Comparative Evaluation:

a) Comparison between HMM-ST and NextPlace: Taking a user's travel sequence as input, both the two predictors predict where the user will be in the 48 slots of the $i+1$ day at the end of the ith day. Specifically, the group of Family person is chosen for the comparative evaluation, as the main idea behind the NextPlace is that human behavior is strongly determined by daily patterns. Also, owing to the fact that NextPlace performs better for a bigger error margin θ[10], we set the error margin θ as 15 min, which is half of each time slot. Fig. 9(a) shows that the HMM-ST performs better. Besides, nearly 65% of users are excessively predicted to be at non-POI [10] in the whole day by NextPlace, which does not conform to the fact. Furthermore, Fig. 9(b) shows the performance of NextPlace in terms of *prediction precision* applied in [10] where NextPlace was

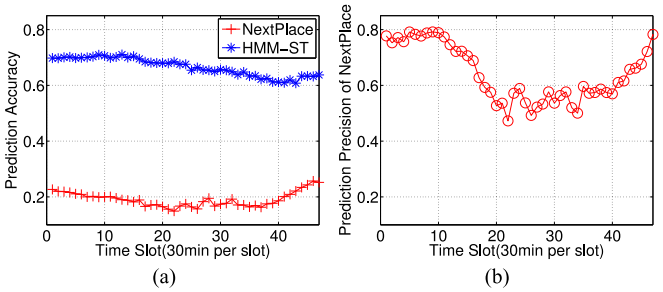


Fig. 9. Comparative evaluation of spatio-temporal mobility prediction model: (a) average *prediction accuracy* of each time slot of the two mobility predictors and (b) *prediction precision* of NextPlace of different time slots in a day.

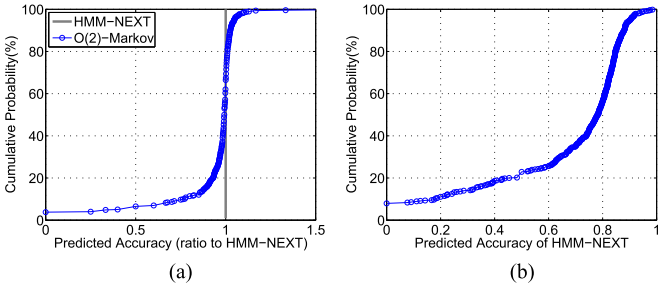


Fig. 10. Comparative evaluation of next-place prediction model: (a) cumulative probability of *prediction accuracy* of $O(2)$-Markov predictor relative to HMM-NEXT, and (b) *prediction accuracy* distribution of HMM-NEXT.

proposed. Different from *prediction accuracy* representing the ratio between the number of correct predictions and the number of all attempted predictions, *prediction precision* is the ratio between the number of correct predictions and the number of predictions forecasting the user to be at a POI. It reaches as high as 80% and is about 50% in some time slots. In NextPlace, predicting users to be at non-POI during the whole day can be caused by changing mobility pattern or deviation from fixed mobility pattern. Thus, NextPlace as an effective predictor performs outstandingly for users with fixed daily pattern. HMM has a stronger learning ability and HMM-ST can be applied to a wider range of users.

The fact that HMM-based spatio-temporal predictor has a better performance can be explained as follows. NextPlace merely focuses on the arrival and residence time at POIs, and it does not consider the transition probabilities among POIs as well as the different probabilities of user appearing at various POIs. Especially, when several places satisfy the prediction, NextPlace randomly chooses one among them [10].

b) Comparison between HMM-NEXT and O(2)-Markov: The two predictors both use a user's travel sequence as input and predict most the probable place to be visited next. Fig. 10(a) shows the relative *prediction accuracy* of $O(2)$-Markov as compared to HMM-NEXT. The relative *prediction accuracy* less than one indicates HMM-NEXT works better than $O(2)$-Markov for these users, who account for almost 70% of users. The *prediction accuracy* of 10% of users achieved by HMM-NEXT is twice that of $O(2)$-Markov. Moreover, though the remaining 30% of users have a lower accuracy using HMM-NEXT, the relative ratio is mostly less than 1.1. Overall, HMM-based predictor performs better than Markov models in terms of *prediction accuracy*. Fig. 10(b) illustrates that for almost 60% of users, HMM-NEXT achieves an accuracy of over 80%.

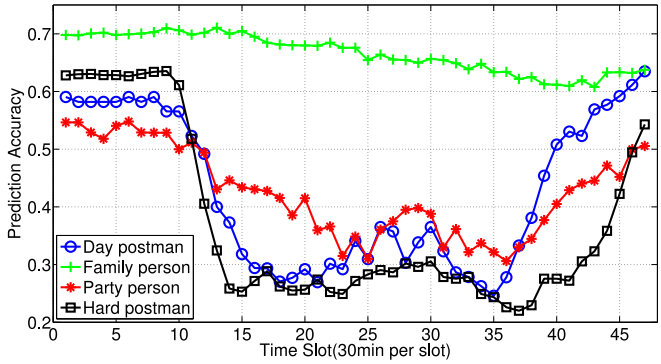


Fig. 11. Average accuracy of each time slot in the whole day w.r.t the four user groups.

There are two reasons why our approach gets better results than the Markov model without considering algorithm complexity: i) Different from the order-2 Markov model that merely considers the current and previous places, our approach takes as many as nine latest places into account. The longer sequence of latest places leads to better prediction [14]. Moreover, when the sequence is not found in the hash table, the length of the sequence decreases, which indeed helps improve the predictor's performance; and ii) HMM uses doubly stochastic process to describe the transition between places and shows significant advantage over the simple Markov model [38].

3) Performance on Different User Groups: a) Efficiency of HMM-ST w.r.t different user groups: Fig. 11 depicts the average *prediction accuracy* of every time slot for each group. By combining the analysis of Figs. 11 and 7, the *prediction accuracy* shows a negative correlation with entropy profiles. Family person with the lowest entropy in the whole day has the highest *prediction accuracy*. The *prediction accuracy* of Hard postman is as low as 30% from 7:00 a.m. to 8:00 p.m., which results from high entropy values at these time segments. Thus, high entropy increases the uncertainty of a typical user's whereabouts and decreases the mobility prediction accuracy.

b) Efficiency of HMM-NEXT w.r.t different user groups: Considering the performance distinction between the Family person and others in terms of spatio-temporal prediction, the Hard postman, Day postman, and Party person are treated as one group named "others" in the evaluation of HMM-NEXT. As shown in Fig. 12(a), over 90% of "others" have an accuracy of more than 60%, and 50% of them have an accuracy of over 80%. However, users with the *prediction accuracy* exceeding 60% merely account for less than 30% of Family person. For further details, we measure the relationship between *prediction accuracy* and the length of travel sequences in the training set (training set size), since the diversity in users' living habits leads to the disparity in frequencies of place transition. Fig. 12(b) presents the distribution of *prediction accuracy* of HMM-NEXT w.r.t.

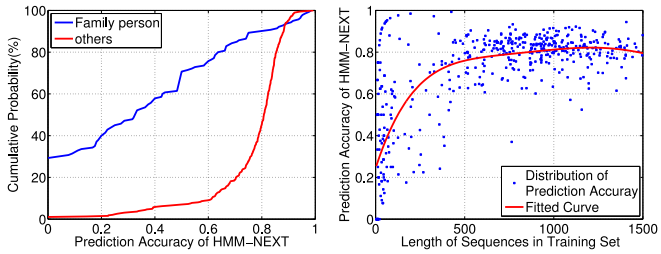


Fig. 12. Performance of HMM-NEXT: (a) *prediction accuracy* of HMM-NEXT with different user groups, and (b) *prediction accuracy* of HMM-NEXT versus the length of travel sequences in training set.

the length of travel sequences and its fitted curve. It indicates that the length is directly linked to *prediction accuracy*, especially when the length is very short. It infers that the size of the training set affects how much information is included for the next-place prediction. Users with limited size of movement history are more likely to produce false prediction. For over 80% of "others," the training set size is more than 500. However, the travel sequences of Family person with a length mostly ranging from 0 to 100 fail to stabilize the movement patterns, thus resulting in a low accuracy of next-place prediction.

We summarize our main findings as follows.

1) The prediction accuracy of the spatio-temporal mobility predictor shows a negative correlation with users' entropy profiles and varies among distinct user groups with different living habits.

2) The spatio-temporal predictor performs better than the next-place prediction model for users leading regular lives and with short traces.

3) For users who move randomly, next-place prediction provides a more effective way to conduct user mobility analysis and produces more convincing prediction results. Overall, considering entropy profile and trace length of users, the spatio-temporal prediction and next-place prediction can be used alternatively as an effective method for user mobility prediction.

VII. Conclusion

Summary of contribution: In this paper, we have investigated the effect of living habits on the models of spatio-temporal prediction and next-place prediction. HMM can be employed to model user mobility from the two prediction perspectives. By investigating the movements of large-scale users collected from LTE control-plane traffic of a whole city in Southern China for 22 days, we have compared the HMM-based predictors with existing models and verified the feasibility of the proposed predictors. In addition, the living habits of users were analyzed based on entropy, according to which users were clustered into distinct groups. We have applied the two predictors to these groups and discovered many meaningful observations. The accuracies of spatio-temporal mobility prediction depend on users' entropy profiles and vary among distinct user groups. Users leading regular lives and with short traces are better modeled from the spatio-temporal perspective. Next-place prediction provides an efficient way to model user mobility for ones leading highly mobile lives. It also suggests that the factors of entropy profile and the length of user traces should be taken into consideration in future user mobility modeling so that an optimal mobility prediction method can be applied to an individual to reach reliable prediction results in practical applications. In particular, it is essential to give individuals control over whether the mobility prediction model can be applied to their daily lives in view of privacy concerns.

Limitations: The proposed methodology clusters users into several groups according to users' living habits by analyzing entropy values of different time segments. Our future efforts aim at extracting users' mobility pattern to classify users and characterize personal life routines. Then, we will further explore human-centered prediction algorithms with high prediction accuracy. Moreover, we will apply datasets of different sizes to further validate the performance and efficiency of the proposed models.

References

[1] J. Liu, F. Liu, and N. Ansari, "Monitoring and analyzing big traffic data of a large-scale cellular network with hadoop," *IEEE Network*, vol. 28, no. 4, pp. 32–39, Jul./Aug. 2014.

[2] J. Liu and N. Ansari, "Identifying website communities in mobile internet based on affinity measurement," *Comput. Commun.*, vol. 41, pp. 22–30, 2014.

[3] Y. Zhang, "User mobility from the view of cellular data networks," in *Proc. IEEE INFOCOM*, 2014, pp. 1348–1356.

[4] J. Yang, Y. Qiao, X. Zhang, H. He, F. Liu, and G. Cheng, "Characterizing user behavior in mobile internet," *IEEE Trans. Emerg. Topics Comput.*, vol. 3, no. 1, pp. 95–106, Mar. 2015.

[5] Y.-B. Lin and P.-K. Huang, "Prefetching for mobile web album," *Wireless Commun. Mobile Comput.*, vol. 16, no. 1, pp. 18–28, 2016.

[6] R. Becker *et al.*, "Human mobility characterization from cellular network data," *Commun. ACM*, vol. 56, no. 1, pp. 74–82, 2013.

[7] Y. Qiao, L. J. Chen, Yihang, and N. Kato, "A mobility analytical framework for big mobile data in densely populated area," *IEEE T. Veh. Technol.*, to be published.

[8] J. Zheng, S. Liu, and L. M. Ni, "An unsupervised framework for sensing individual and cluster behavior patterns from human mobile data," in *Proc. 2012 ACM Conf. Ubiquitous Comput.*, 2012, pp. 153–162.

[9] Y. Ye, Y. Zheng, Y. Chen, J. Feng, and X. Xie, "Mining individual life pattern based on location history," in *Proc. 10th Int. Conf. Mobile Data Manage.: Syst., Services Middleware*, 2009, pp. 1–10.

[10] S. Scellato, M. Musolesi, C. Mascolo, V. Latora, and A. T. Campbell, "Nextplace: A spatio-temporal prediction framework for pervasive systems," in *Proc. 9th Int. Conf. Pervasive Comput.*, 2011, pp. 152–169.

[11] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[12] T. Anagnostopoulos, C. Anagnostopoulos, and S. Hadjiefthymiades, "Efficient location prediction in mobile cellular networks," *Int. J. Wireless Inform. Networks*, vol. 19, no. 2, pp. 97–111, 2012.

[13] W. Dong, N. Duffield, Z. Ge, S. Lee, and J. Pang, "Modeling cellular user mobility using a leap graph," in *Proc. 14th Int. Conf. Passive Active Meas.*, 2013, pp. 53–62.

[14] L. Song, D. Kotz, R. Jain, and X. He, "Evaluating next-cell predictors with extensive Wi-Fi mobility data," *IEEE Trans. Mobile Comput.*, no. 12, pp. 1633–1649, Dec. 2006.

[15] M. Kim, D. Kotz, and S. Kim, "Extracting a mobility model from real user traces," in *Proc. MobiCom*, 2006, vol. 6, pp. 1–13.

[16] D. Ashbrook and T. Starner, "Using GPS to learn significant locations and predict movement across multiple users," *Pers. Ubiquitous Comput.*, vol. 7, no. 5, pp. 275–286, 2003.

[17] C. Zhou, D. Frankowski, P. Ludford, S. Shekhar, and L. Terveen, "Discovering personally meaningful places: An interactive clustering approach," *ACM Trans. Inform. Syst.*, vol. 25, no. 3, 2007, Art. no. 12.

[18] J. Alvarez-Lozano, J. A. García-Macías, and E. Chávez, "Crowd location forecasting at points of interest," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 18, no. 4, pp. 191–204, 2015.
[19] J. H. Kang, W. Welbourne, B. Stewart, and G. Borriello, "Extracting places from traces of locations," in *Proc. 2nd ACM Int. Workshop Wireless Mobile Appl. Services WLAN Hotspots*, 2004, pp. 110–118.
[20] S. Isaacman *et al.*, "Identifying important places in people's lives from cellular network data," in *Proc. 9th Int. Conf. Pervasive Comput.*, 2011, pp. 133–151.
[21] J. Liu, S. Zhang, N. Kato, H. Ujikawa, and K. Suzuki, "Device-to-device communications for enhancing quality of experience in software defined multi-tier LTE-A networks," *IEEE Network*, vol. 29, no. 4, pp. 46–52, Jul./Aug. 2015.
[22] H. Nishiyama, M. Ito, and N. Kato, "Relay-by-smartphone: Realizing multihop device-to-device communications," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 56–65, Apr. 2014.
[23] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Proc. 4th Annu. ACM/IEEE Int. Conf. Mobile Comput. Networking*, 1998, pp. 85–97.
[24] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Commun. Mobile Comput.*, vol. 2, no. 5, pp. 483–502, 2002.
[25] J. Ariyakhajorn, P. Wannawilai, and C. Sathitwiriyawong, "A comparative study of random waypoint and Gauss-Markov mobility models in the performance evaluation of manet," in *Proc. Int. Symp. Commun. Inform. Technol.*, 2006, pp. 894–899.
[26] J. Zheng, S. Liu, and L. M. Ni, "Effective routine behavior pattern discovery from sparse mobile phone data via collaborative filtering," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2013, pp. 29–37.
[27] J. Zheng and L. M. Ni, "Effective mobile context pattern discovery via adapted hierarchical Dirichlet processes," in *Proc. IEEE 15th Int. Conf. Mobile Data Manage.*, 2014, vol. 1, pp. 146–155.
[28] M. A. Bayir, M. Demirbas, and N. Eagle, "Mobility profiler: A framework for discovering mobility profiles of cell phone users," *Pervasive Mobile Comput.*, vol. 6, no. 4, pp. 435–454, 2010.
[29] Q. Lv, Z. Mei, Y. Qiao, Y. Zhong, and Z. Lei, "Hidden Markov model based user mobility analysis in LTE network," in *Proc. IEEE Wireless Pers. Multimedia Commun.*, 2014, pp. 379–384.
[30] Y. Chon, H. Shin, E. Talipov, and H. Cha, "Evaluating mobility models for temporal prediction with high-granularity mobility data," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2012, pp. 206–212.
[31] P. Baumann, W. Kleiminger, and S. Santini, "How long are you staying?: Predicting residence time from human mobility traces," in *Proc. 19th Annu. Int. Conf. Mobile Comput. Networking*, 2013, pp. 231–234.
[32] G. Smith, R. Wieser, J. Goulding, and D. Barrack, "A refined limit on the predictability of human mobility," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2014, pp. 88–94.
[33] J. McInerney, S. Stein, A. Rogers, and N. R. Jennings, "Breaking the habit: Measuring and predicting departures from routine in individual human mobility," *Pervasive Mobile Comput.*, vol. 9, no. 6, pp. 808–822, 2013.
[34] H. Si, Y. Wang, J. Yuan, and X. Shan, "Mobility prediction in cellular network using hidden Markov model," in *Proc. 7th IEEE Consum. Commun. Networking Conf.*, 2010, pp. 1–5.
[35] W. Mathew, R. Raposo, and B. Martins, "Predicting future locations with hidden Markov models," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 911–918.
[36] X. Lu, E. Wetter, N. Bharti, A. J. Tatem, and L. Bengtsson, "Approaching the limit of predictability in human mobility," *Sci. Rep.*, vol. 3, 2013, Art. no. 2923.
[37] T. G. Dietterich, "Machine learning for sequential data: A review," in *Joint IAPRInt. Workshops Statist. Techn. Pattern Recognit. Struct. Syntactic Pattern Recognit.*, 2002, pp. 15–30.
[38] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
[39] D. Shah and T. Zaman, "Community detection in networks: The leader-follower algorithm," arXiv:1011.0774, 2010.
[40] Y. Qiao, J. Yang, H. He, Y. Cheng, and Z. Ma, "User location prediction with energy efficiency model in the long term-evolution network," *Int. J. Commun. Syst.*, vol. 29, pp. 2169–2187, 2015.
[41] J. Lin and C. Dyer, "Data-intensive text processing with mapreduce," *Synthesis Lectures Human Lang. Technol.*, vol. 3, no. 1, pp. 1–177, 2010.
[42] J. Yang, X. Zhang, Y. Qiao, Z. Fadlullah, and N. Kato, "Global and individual mobility pattern discovery based on hotspots," in *Proc. IEEE Int. Conf. Commun.*, 2015, pp. 5577–5582.
[43] *GPRS enhancements for E-UTRAN access (Release 12)*, 3GPP, TS 23.401, Mar. 2013.
[44] S. Isaacman *et al.*, "Human mobility modeling at metropolitan scales," in *Proc. 10th Int. Conf. Mobile Syst., Appl., Services*, 2012, pp. 239–252.
[45] P. Baumann, W. Kleiminger, and S. Santini, "The influence of temporal and spatial features on the performance of next-place prediction algorithms," in *Proc. ACM Conf. Ubiquitous Comput.*, 2013, pp. 449–458.

Qiujian Lv is currently working toward the Ph.D. degree in information and communication engineering from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China.

Her research interests include traffic measurement and classification, mobile Internet traffic analysis, cloud computing, and data mining.

Yuanyuan Qiao (M'15) received the B.E. degree in electronic information engineering from Xidian University, Xi'an, China, in 2009 and the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2014.

She is currently a Lecturer at the School of Information and Communication Engineering, BUPT. Her research interests include traffic measurement and classification, mobile Internet traffic analysis, and big data analytics.

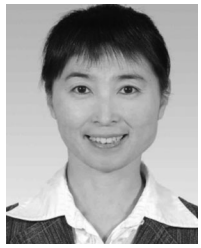
Nirwan Ansari (F'09) received the Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 1988, the BSEE (summa cum laude with a perfect GPA) from New Jersey Institute of Technology, Newark, NJ, USA, in 1982, and the MSEE degree from the University of Michigan, Ann Arbor, MI, USA, in 1983.

He is with the Advanced Networking Laboratory, Department of Electrical and Computer Engineering, New Jersey Institute of Technology, and Distinguished Professor in the Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, USA. He has also been a Visiting (Chair) Professor at several universities. He is the author, with T. Han, of *Green Mobile Networks: A Networking Perspective* (Wiley-IEEE, 2016), and coauthored two other books. He has also (co-)authored more than 500 technical publications and more than 200 published in widely cited journals/magazines. He was a Guest Editor for a number of special issues, covering various emerging topics in communications and networking. He has served on the Editorial/Advisory Board of more than 10 journals. His research interests include green communications and networking, cloud computing, multimedia communications, and various aspects of broadband networks.

Dr. Ansari was elected to serve on the IEEE Communications Society (ComSoc) Board of Governors as a Member-At-Large, has chaired ComSoc technical committees, and has been actively organizing numerous IEEE International Conferences/Symposia/Workshops. He has frequently been delivering keynote addresses, distinguished lectures, tutorials, and invited talks. Some of his recognitions include several Excellence in Teaching Awards, a few best paper awards, the NCE Excellence in Research Award, the ComSoc Ad Hoc and Sensor Networks Technical Committee Outstanding Service Recognition Award, the NJ Inventors Hall of Fame Inventor of the Year Award, the Thomas Alva Edison Patent Award, Purdue University Outstanding Electrical and Computer Engineer Award, and designation as a COMSOC Distinguished Lecturer. He holds more than 30 U.S. patents.

Jun Liu (M'14) received the B.E. degree in information engineering, and the Ph.D degree in signal and information processing both from the Department of Information Engineering, Beijing University of Posts and Telecommunications (BUPT), Biejing, China, in 1998 and 2003, respectively.

He is currently the Director of the Center for Data Science, BUPT. His research interests include network traffic monitoring, telecom big data analysis, and streaming data algorithm.



Jie Yang received the B.E. degree in information engineering, and the M.E. and Ph.D degrees in signal and information processing from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 1993, 1999, and 2007, respectively.

She is currently a Professor and the Deputy Dean of the School of Information and Communication Engineering, BUPT. Her research interests include broadband network traffic monitoring, user behavior analysis, big data analysis in Internet and Telecom, etc. She has published several papers in international magazines and conferences, including the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE TRANSACTION ON WIRELESS COMMUNICATIONS and the IEEE TRANSACTION ON PARALLEL AND DISTRIBUTED SYSTEMS. Dr. Yang was the Vice Program Committee Cochair of the IEEE International Conference on Network Infrastructure and Digital Content 2014 and 2012.