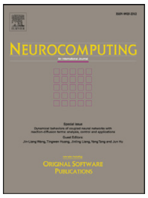
Neurocomputing 278 (2018) 99–109

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom



A hybrid Markov-based model for human mobility prediction

Yuanyuan Qiao[a,*], Zhongwei Si[a], Yanting Zhang[a], Fehmi Ben Abdesslem[b], Xinyu Zhang[a], Jie Yang[a]

[a] *Beijing Key Laboratory of Network System Architecture and Convergence, and Beijing Laboratory of Advanced Information Networks, Beijing University of Posts and Telecommunications, Beijing 100876, China*

[b] *Decisions, Networks and Analytics Laboratory, SICS Swedish ICT AB, Kista SE-164 29, Sweden*

ARTICLE INFO

Article history:
Received 1 December 2016
Revised 9 May 2017
Accepted 21 May 2017
Available online 5 September 2017

Keywords:
Non-Gaussian mobility data
Hybrid Markov-based model
Human mobility
Mobility prediction
Spatio-temporal regularity

ABSTRACT

Human mobility behavior is far from random, and its indicators follow non-Gaussian distributions. Predicting human mobility has the potential to enhance location-based services, intelligent transportation systems, urban computing, and so forth. In this paper, we focus on improving the prediction accuracy of non-Gaussian mobility data by constructing a hybrid Markov-based model, which takes the non-Gaussian and spatio-temporal characteristics of real human mobility data into account. More specifically, we (1) estimate the order of the Markov chain predictor by adapting it to the length of frequent individual mobility patterns, instead of using a fixed order, (2) consider the time distribution of mobility patterns occurrences when calculating the transition probability for the next location, and (3) employ the prediction results of users with similar trajectories if the recent context has not been previously seen. We have conducted extensive experiments on real human trajectories collected during 21 days from 3474 individuals in an urban Long Term Evolution (LTE) network, and the results demonstrate that the proposed model for non-Gaussian mobility data can help predicting people's future movements with more than 56% accuracy.

1. Introduction

Analyzing the characteristics of human mobility reveals that human trajectories are predictable. They often exhibit a high degree of temporal and spatial regularity. In order to find the basic rules governing human dynamics and build a model to predict human mobility, various studies of human mobility have been conducted in recent years. With the emergence of smartphones and location-based services, since most of location-based services require accurate or approximate position of user, predicting user's next locations shows a great potential for service providers to improve the user experience. In particular, it has become a critical enabler for a wide range of applications, such as location-based advertising, early warning systems, and city-wide traffic planning [1].

Over the previous years, different methods have been proposed in the literature, using varied types of data and aiming at predicting distinct aspects of human mobility. Real traces are crucial to train and evaluate prediction models. Nowadays, people's movements can be easily sensed with mobile phones, which are generating large volumes of mobility data, such as Call Detail Records (CDRs) [2,3], Global Positioning System (GPS) tracks [4–7], data traffic of mobile networks [8–11], and Wi-Fi access points data [12]. Recently, researchers found that data traffic from 2G/3G/4G cellular networks is extremely useful for studying human dynamics [8,9,11,13–17], and can provide people's trajectories in a large scale. Passively collecting human movement trajectories while they access the Internet with their smartphone presents many advantages: high cost efficiency, low energy consumption, covering a wide range and a large number of people. Moreover, this can be done with a fine time granularity, as people tend to surf the Internet frequently on their smartphone while commuting. Also, many apps send or receive network traffic packets, even when running in the background [18]. Collected datasets, despite having different collection methods, and covering different populations with various accuracy or time granularity, always present a similar characteristic: human trajectories described by the data all follow a non-Gaussian spatio-temporal distribution. [18–22].

Existing research works on mobility prediction are exploiting diverse types of data [1]. The most commonly used algorithms to predict mobility include machine learning algorithms (clustering techniques [4–6,8,9,23–25], Bayesian models [26–30], neural networks [2,31]), state-based techniques (Markov models [7,12,14,15,32], LeZi family [11,13], hidden Markov models [4,5,10,25,33]), and pattern matching algorithms (prediction by

[*] Corresponding author.
E-mail addresses: yyqiao@bupt.edu.cn (Y. Qiao), fehmi@sics.se (F.B. Abdesslem).

http://dx.doi.org/10.1016/j.neucom.2017.05.101
0925-2312/© 2017 Elsevier B.V. All rights reserved.

Partial Matching [34,35]). As for predicting mobility in a large population through the cellular network, Markov-based algorithms are more suitable [36] and out-perform other methods when applied to short trajectories, and when considering the temporal factor [37]. They also prove to be very appropriate for future generation mobile networks [38].

To predict the next location using large-scale non-Gaussian mobility data, this paper aims to improve the prediction accuracy of Markov-based algorithms. In order to achieve this goal, the key is to enrich the states in the underlying Markov model by considering relevant external information, increasing time and spatial complexity. In this paper, a hybrid Markov-based prediction model, which considers non-Gaussian and spatio-temporal characteristics of real human mobility data while predicting, is constructed. The prediction accuracy of a simple Markov algorithm on our real dataset, can be improved from 44.02% to 56.39% in our experiments. Overall, the contributions of this paper can be summarized as follows:

- We examine the characteristics of real mobility data extracted from user's data traffic in an LTE network, which provide properties we should consider while predicting user's future movement. The human mobility represented in the dataset shows two main characteristics that follow a non-Gaussian distribution, namely the trip distance and the radius of gyration. This means that (1) displacement within short distance is frequently seen in the dataset and (2) frequent travels occur in a limited range in individuals' daily life. In other words, people generally move within a bounded region and only occasionally travel long distance. In addition, we further study the probability of finding a user at different locations, and returning to the same location. Analysis results show that people visit some primary locations periodically with high probability, which confirms the intuitive existence of mobility patterns for each user. The periodicity of this pattern hints that temporal factors can contribute to predicting the next location of individuals.
- We propose a hybrid Markov-based model to predict users' future movements. It uses different methods consecutively to discover spatio-temporal pattern of each individual's trajectory. More specifically, the model determines the order of Markov algorithm by discovering the regular mobility patterns for each individual, and takes into account the time of the day where locations are visited. This way, non-Gaussian and spatio-temporal characteristics of users' trajectories are fully considered, which contributes a lot to getting a better prediction result.
- Markov-based algorithms fail to correctly predict future movements if the new location has never been visited by an individual. To alleviate this issue, we consider the trajectories of geo-friends: users sharing similar trajectories and mobility patterns. We then employ a user-based recommendation method with Collaborative Filtering to predict users' future movements when their own mobility pattern cannot contribute to the prediction.

The remainder of this paper is organized as follows. In Section 2, we provide a survey of the related works. Section 3 examines the non-Gaussian and spatio-temporal characteristics of users' trajectories extracted from our dataset. In Section 4 we describe the hybrid Markov-based mobility prediction model, before presenting in Section 5 the methods employed in the mobility prediction model, including the hotspot detection method, the mobility pattern discovery algorithm, the variable-order Markov prediction algorithm with temporal factors, and finally the algorithm enhancement using similarity of geo-friends' movements. The performance of the proposed mobility prediction model is then evaluated in Section 6, and conclusions are drawn in Section 7.

2. Related work

Many datasets collected from real world applications and services have been found to show non-Gaussian characteristics, such as gene networks [39], hyper-spectral images [40], or climate extremes [41], for example. As for datasets describing human mobility, despite some characteristics showing Gaussian distributions [42,43], many other characteristics show non-Gaussian distribution [44,45] and have been studied in the literature. In particular, the locations visited by people in their daily life show a non-Gaussian distribution [19], and predicting those locations is a challenging objective addressed by recent research works to enable a wide range of applications, such as location-based advertising, early warning systems, and city-wide traffic planning.

Markov models are widely used in prediction algorithms, due to their efficiency, simplicity, and low computing costs [38]. For example, a Markov chain prediction model considers the sequence of locations last visited by a user to predict the next location. The length k of that sequence of locations represents the order of the Markov chain, and we refer to this model as an order-k Markov-based model. Markov-based prediction methods fall into one of four categories: (1) order-k Markov-based methods that only use the historical locations to discover individual movement patterns [7,46,47], (2) order-k Markov-based methods that also consider external information in addition to historical trajectories [48–50], (3) hybrid Markov-based methods that are enhanced with other prediction methods [32,51,52], (4) evolved algorithms based on Markov models, such as the LeZi algorithm [11,13] or algorithms based on Hidden Markov Models [4,5,10,25,33].

Over a decade ago, classical Markov models of low order (generally, 1 or 2) have already been used to predict future movements [7,46,47]. It has been found that low-order Markov-based predictors performed as well as, or better than the more complex and more space-consuming compression-based predictors such as Prediction by Partial Matching (PPM), or Sampled Pattern Matching Algorithm (SPM) [46]. Markov models of higher order, combined with external information about people's schedule, have also been used to improve the prediction accuracy. An order-k Markov-based predictor with fall-back was proposed in [48]. The predictor proposed by the authors falls back to a lower order of k if a certain order-k predictor was unsuccessful. In addition, they also incorporate external information, such as the event notes on Microsoft Outlook and Google Calendar, into the context of an $O(k)$ Markov predictor, enriching the states in the underlying Markov model. In other works, the distances of trajectories, travel times, driving habits, and social context have been incorporated into corresponding models to compute the probabilities of future locations in [49,50,53,54]. Recently, it has been found that different predictors should be applied according to the application scenario, to the mobility characteristics of the data, and to individuals [55]. To this aim, hybrid methods have been proposed, which uses different prediction methods consecutively [51,52,56,57], for different datasets [32], for each individuals [14,55], or even for different times of the day [58].

Authors in [51] present a prediction model in LTE networks that combines two complementary algorithms: the global profiles-based and the local profiles-based algorithm. The former is implemented in the enhanced Node B and the home enhanced Node B, and the latter works at the user terminal level. By using a real vehicle passage record dataset, the authors in [32] present the Global, Personal, and Regional Markov Model for a dataset with different granularity to tackle the problem of predicting next locations for vehicles' trajectories. In [58], the authors proposed a prediction methods to capture the relationship between movement patterns in different time periods, which are based on the important observation that movement patterns often change over time. In addition, our previous studies suggest applying different prediction methods to users with distinct spatio-temporal characteristics. Spatio-temporal based prediction methods are more suitable for those who visit very limited locations during the day and follow a regular pattern during the week. For those who spend much time commuting and follow a fairly regular pattern during the week, the next location prediction is more efficient [14,55]. Based on previous studies, in this paper, we focus on improving Markov-based prediction algorithms by addressing the following drawbacks of many existing prediction methods:

- Time independence: Markov-based models often disregard the time of the day. However, the temporal factor is an important feature that commonly affects people's mobility, since people tend to visit the same location at the same time of the day: going to work in the morning and going back home at night, for instance.
- High memory cost: for every pattern representing a sequence of locations visited by an individual, Markov-based algorithms typically store the frequencies of all the next possible locations. Storing all those values requires a large amount of memory, especially when the length of location sequences is high, increasing the number of possible patterns. Instead, considering the spatio-temporal periodicity of people's movement can reduce the number of possible patterns and hence the amount of memory needed by Markov-based algorithms to predict the next location.
- New sequence of locations: Markov-based algorithm usually fail to predict the next location when an individual is visiting a sequence of locations for the first time. In this case, considering the next locations visited by other individuals who also visited the same sequence of locations can help predicting the next location.

In addition, in our paper [59], we find that users' mobility patterns have strong spatio-temporal correlation property, i.e., mobility patterns have their own typical occurrence time depending on the pattern's context. It drives us fully consider user's mobility pattern and its occurrence time while proposing the prediction model. In general, compared to previous works, we propose a hybrid Markov-based algorithm to address the issues mentioned above, for predicting human mobility using non-Gaussian data consisting of three basic attributes: the ID of users, their location, and the corresponding timestamps. Firstly, characteristics of real mobility data extracted from data traffic of LTE network are examined, providing the factors we should consider when improving the prediction accuracy of Markov-based methods. Then, instead of using a Markov chain of fixed order, the frequent mobility sequences of locations are used to define the order. Secondly, the probability of each Markov state is calculated by considering the time of the day. Finally, when a prediction cannot be computed, we estimate the probability of the next location by looking at the locations visited by other users visiting similar locations. In summary, without introducing other external information or datasets to enrich the states, the proposed model improves the Markov-based model by discovering and exploring the characteristics of user's trajectory.

3. Mobility data characteristics

In this section, we examine the characteristics of our dataset. First, we provide a brief description of the LTE network architecture and how the data is collected from the network. Then the dataset is analyzed to provide an overview of mobility features, demonstrating how predicting users' future locations could be done by considering the non-Gaussian and spatio-temporal characteristics of the dataset.

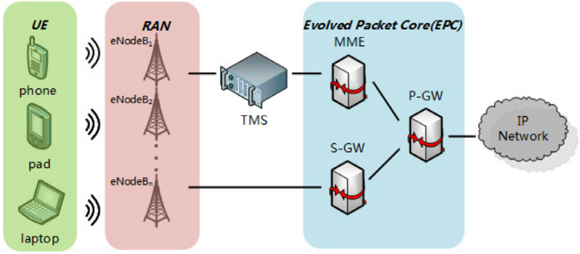


Fig. 1. LTE network architecture.

3.1. Data collection

The dataset is collected from the LTE network of a large Chinese city from October 10th 2013 to October 31st 2013, capturing the mobility patterns of 3474 individuals for 21 days. A high-level view of the LTE mobile network with our data traffic capture device is shown in Fig. 1. There are three major components in the LTE mobile network, namely the User Equipment (UE), the Radio-Access Network (RAN), and the Evolved Packet Core (EPC).

The UE is any device used directly by an end-user to communicate on the network, such as a smartphone, a laptop computer equipped with a mobile broadband adapter, or any other communication device with a SIM card.

The RAN establishes the connection between the UE and the EPC. It uses a flat architecture with multiple eNodeBs (evolved NodeBs). An eNodeB is a hardware equipment connected to the mobile phone network and communicating directly with mobile handsets (UEs).

The EPC is a packet-only core network. It serves as the equivalent of GPRS networks via the Mobility Management Entity (MME), Serving Gateway (SGW) and Packet Data Network (PDN) Gateway (PGW) sub-components.

The dataset used in this study is composed of LTE control-plane packets, which are collected by our custom Traffic Monitoring System (TMS), deployed between eNodeBs and the MME, as shown in Fig. 1. When a user actively uses the LTE network, the locations of the corresponding eNodeBs associated with their UE are logged and later considered as the user's location. If the user does not use the LTE network, their UE in state EMM-REGISTERED initiates the tracking area updating procedure by sending a Tracking Area Update (TAU) request every 12 min. This means that we record the location of active users with a small time granularity, and the location of inactive users every 12 min.

The collected dataset comprises a sequence of time-stamped records, each of which mainly contains the user unique identifier, the associated eNodeB unique identifier, and the online time. For security reasons, users' confidential identifiers are replaced by a hashed number, which is used to mark subscribers without infringing their privacy nor affecting our study.

3.2. Non-Gaussian characteristic analysis

We use three widely accepted indicators to describe large-scale human mobility: the trip distance r, the radius of gyration r_g, and the number of visited locations over time $S(t)$. The radius of gyration captures how far the subscribers move instead of the actual distance they travel. Visiting the same sequence of locations in a circle continuously does not increase the radius of gyration, but a long distance travel on a straight line does. The radius of gyration is defined as:

$$r_g = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\vec{r_i} - \vec{r_{cm}})^2}, \tag{1}$$

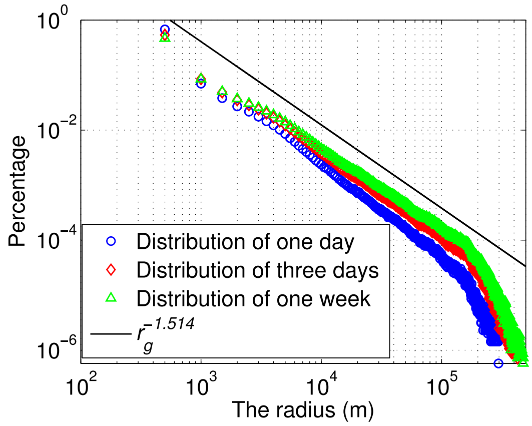


Fig. 2. The $P(r_g)$ distribution of the radius of gyration r_g for users.

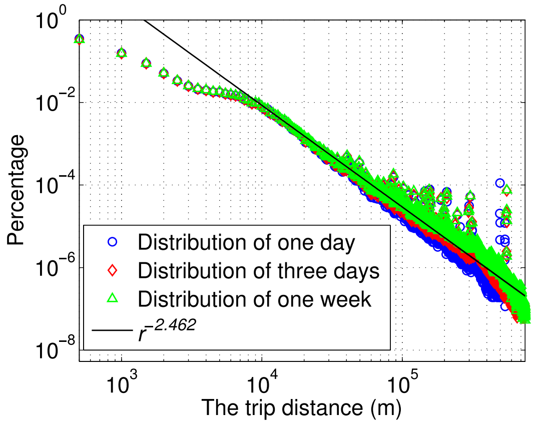


Fig. 3. The trip distance distribution $P(r)$.

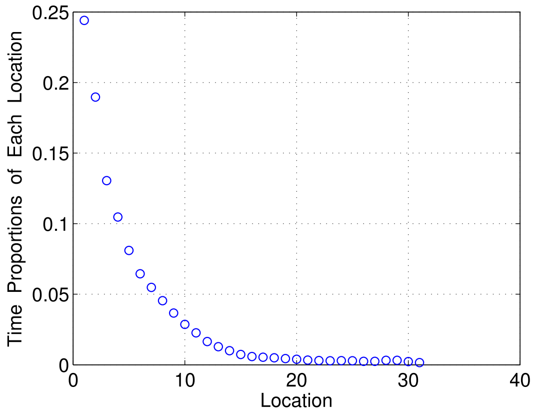


Fig. 4. A Zipf distribution shows the probability of finding a user at different locations that are ranked on the basis of their visit frequencies.

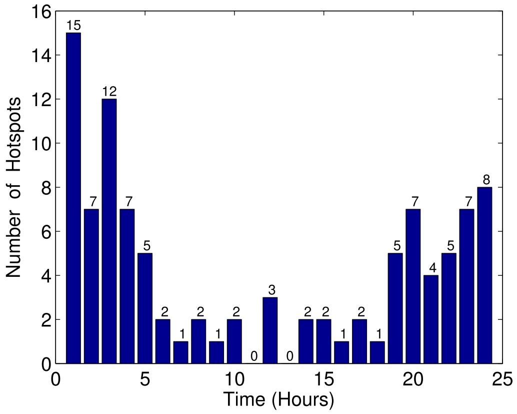


Fig. 5. Histogram of lifetime duration of hotspots.

where $\vec{r_i}$ presents the $i = 1, 2, \ldots, n$ -*th* locations recorded for a given user describing their trajectory. $\vec{r_{cm}} = \sum_{i=1}^{n} \vec{r_i}$ is the center of mass point of the users' trajectory.

For our dataset, we found that the $P(r_g)$ distribution of the radius of gyration r_g for users follows a power law distribution, i.e., $P(r_g) \sim r_g^{-\beta}$, with $\beta \approx 1.514$, as shown in Fig. 2 [18]. In addition, the trip distance distribution $P(r)$, which quantifies the relative probability of finding a displacement of length r in a short time, follows a power law distribution, i.e., $P(r) \sim r^{-\beta}$, with $\beta \approx 2.462$, as shown in Fig. 3 [18]. According to the non-Gaussian characteristics of users' mobility data, we can conclude that, for most of the users, the trip distance between two continuous observed displacements is very small, and the movement range is limited, which imply that people frequently move in a very small area.

In order to better understand movement patterns in a spatio-temporal perspective, we further analyze the dataset by answering the following questions. "How many locations are visited by users", "Since most users visit a limited number of locations, how long does it take for users to visit all these locations?", "How often do users return to the visited locations". Answers to such questions will provide the characteristics we should consider when improving the mobility prediction model.

3.3. Spatial characteristic analysis

In this subsection, we focus on the frequency of visits for each user. We rank each location a user visits according to how often the location is visited. For instance, a location with rank $L = 1$ indicates the most visited location of the selected user. For each user, we create the list of locations where they appear in the ascending order of the rank. Fig. 4 shows a Zipf distribution of the probability distribution of visit frequency of locations ranked L. Note that people spend roughly 40% of their time in their top two preferred locations. This clearly shows that the number of primary locations each user visits is limited and even when users move between multiple locations, they can be found in their favorite locations with high probability. We also keep in mind that if an inactive user moves to another place and then moves back within less than 12 min, this displacement cannot be observed.

We find that, for each individual, the locations users visit are non-homogeneous, in both time and space. We would like to further examine how frequently different locations are visited by the population in the city. Places with large population, such as big shopping malls, residential areas, traffic hubs, or places for group activities, are referred as "city hotspots" and are significant to the city. Identifying hotspots requires selecting an appropriate threshold for a parameter measuring the popularity of the location. Traditional hotspot detection methods set a unique threshold for this parameter, but this approach suffers from a lack of flexibility and adaptability. Here we will employ a parameter-free method that allows us to control the effect of the threshold's selection, and select the threshold according to the actual situation of human's mobility [60]. The detailed implementation for this method can be found in [59]. From our dataset, 101 hotspots in the city are identified.

However, hotspots always change with time, which shows the movements of population in different regions in the city, as shown in Fig. 5. We can essentially distinguish three groups: the permanent (from 19 up to 24 h), intermittent (from 1 up to 5 h) and intermediary (all the others) hotspots. These hotspots are crucial

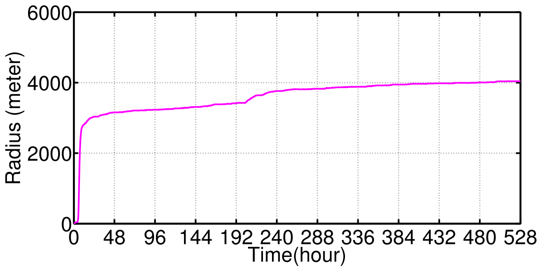


Fig. 6. The increasing of users' radius of gyration r_g with time in 21-day period.

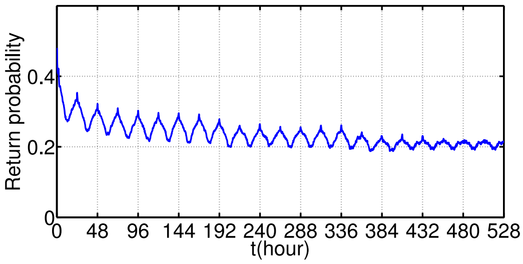


Fig. 7. Probability distribution of time to returning to the same location.

for a range of technology and policy decisions in areas such as telecommunications and transportation infrastructure deployment. In our experiment, in order to focus on the popular places in the city, we use the identified hotspots to construct user's trajectory.

3.4. Spatio-temporal characteristic analysis

In order to understand the temporal periodicity of user's movement, we draw users' radius of gyration against the time t. It is expected that the longer the duration t, the larger the radius of gyration $r_g(t)$. However, Fig. 6 indicates that there is a boundary of the movement area for people [59]. The radius of gyration will reach almost close to the boundary, as it rises rapidly within 24 h and very slowly after that.

We further investigate the reason for the quick saturation of the radius of gyration by measuring the 'return probability' [61] for each user, defined as the probability that a user returns after t hours to the same position. Fig. 7 shows the distribution has relative peaks at 24th, 48th, 72nd h [59]. It indicates the periodic nature of human mobility within a 24-h period and tendency of returning to the same location periodically. The analyses of user movement verify that human trajectories do not show complete randomness. On the contrary, it often exhibits a high degree of temporal and spatial periodicity. That is, each individual may be characterized by a significant probability to return to a few highly frequently visited locations and dwell a longer duration in those locations.

In summary, we find that people move regularly by day with a limit range and the regularity can last for a long period of time, which imply that spatial and temporal factors can contribute a lot to location prediction of users. The above observations drive us to improve the prediction accuracy of mobility prediction model by fully considering the non-Gaussian and spatio-temporal characteristics of real mobility dataset.

4. Hybrid Markov-based prediction model

In this section, we propose a hybrid Markov-based model for mobility prediction, as illustrated in Fig. 8. In the pre-processing module, after collecting the data traffic from the cellular network, and in order to generate the trajectories of users, three-tuples $<$ user ID, Base Station ID, Time $>$ are extracted from the data. Here, the visited locations of users are the locations of the corresponding base stations, which can be distinguished by their unique ID. Then, we identify the hotspots in the city, and use hotspots as locations in users' trajectories. New three-tuples $<$ user ID, Hotspot ID, Time $>$ are now the new input of the mobility prediction algorithm.

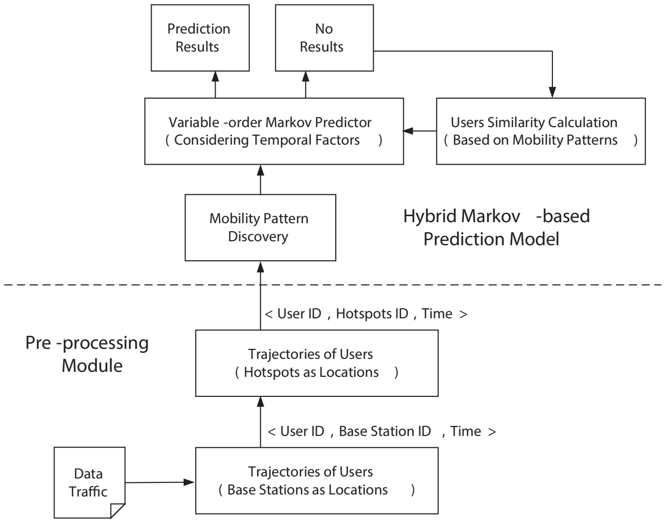

Prediction Results
No Results
Variable -order Markov Predictor (Considering Temporal Factors)
Users Similarity Calculation (Based on Mobility Patterns)
Mobility Pattern Discovery
Hybrid Markov -based Prediction Model
< User ID , Hotspots ID , Time >
Pre -processing Module
Trajectories of Users (Hotspots as Locations)
< User ID , Base Station ID , Time >
Data Traffic
Trajectories of Users (Base Stations as Locations)

Fig. 8. Pre-processing module and hybrid Markov-based prediction model.

In order to fully consider the characteristics of our real dataset, and overcome the shortcomings of existing Markov-based prediction algorithms, the proposed hybrid Markov-based prediction model for human mobility contains three parts: Mobility Pattern Discovery, Variable-order Markov Predictor, and Users Similarity Calculation. At the first stage, we discover individuals' mobility patterns, and the frequent mobility sequences that are used next to estimate the order of the Markov predictor for each individual in the second stage. Then at the second stage, a temporal factor is taken into consideration when predicting individual's future location with the variable-order Markov predictor. At the last stage, for the next locations that could not be predicted by individuals' own historical mobility patterns, we get the prediction results from individuals' geo-friends, who exhibit similar mobility patterns. The specific methods used in the pre-processing module and in the mobility prediction model are illustrated in the next section.

5. Methods applied in the hybrid Markov-based model

In this section, all the methods used in the pre-processing module and in the mobility prediction model are presented in the order of prediction process, including the mobility pattern discovery algorithm, the variable-order Markov prediction algorithm considering temporal factors, and the mobility pattern based on users' similarity.

5.1. Mobility pattern discovery

5.1.1. Algorithm for discovering the mobility pattern

In order to discover mobility patterns from the mobility trajectories, the following formal definitions are given:

Definition 1. (The length of mobility pattern) The length of mobility pattern $p = < h_1, h_2, \ldots, h_n >$ is n when the pattern p contains n hotspots. We denote the length of mobility pattern p by $len(p)$ and the mobility pattern as length-(n) pattern.

Table 1
Main notations of mobility pattern discovery algorithm.

Notation	Description
δ	Minimum support threshold
T	Set of mobility trajectories
H	Set of hotspots
C_K	Set of length-(k) candidate mobility patterns
P_K	Set of length-(k) mobility patterns
P	Set of all mobility patterns

For example, if mobility pattern $p =< a, b, c, d >$, $len(p) = 4$ and mobility pattern p is length-4 pattern.

Definition 2. (Sub-pattern) A mobility pattern p = $< a_1, a_2, \ldots, a_n >$ is a super-pattern of another mobility pattern $q =< b_1, b_2, \ldots, b_m >$, written as $q \subset p$, if pattern p contains pattern q. And then, q is called a sub-pattern of p.

For example, if mobility pattern $p =< a, b, c, d >$ and mobility pattern $q =< a, b, c >$, mobility pattern p is a super-pattern of pattern q, and q is a sub-pattern of p.

Definition 3. (Candidate mobility pattern) A mobility pattern $p =< a_1, a_2, \ldots, a_n >$ is a candidate mobility pattern, if its sub-pattern $q =< a_1, a_2, \ldots, a_{n-1} >$ is discovered as a mobility pattern.

For example, if $q =< a, b, c >$ is a mobility pattern, $p =< a, b, c, d >$ is a candidate mobility pattern.

Definition 4. (Support value) Let $T = \{t_1, t_2, \ldots, t_N\}$ be a trajectory set that contains N trajectories. The support value of pattern p is defined as

$$supp(P) = \frac{|\{t_i | P \subset t_i \, and \, 1 \le i \le N\}|}{N}. \qquad (2)$$

For example, if a trajectory set contains 10 trajectories and 6 trajectories contain the mobility pattern *p*, $supp(p)$ equals 0.6.

Definition 5. (Mobility pattern) Given a minimum support threshold, δ, a candidate mobility pattern p is defined as a mobility pattern if and only if p has support value satisfying: $supp(P) \ge \delta$.

We modify the Apriori algorithm in order to mine the frequent mobility sequences in the trajectories [59]. We refer to the modified version of the Apriori algorithm as the mobility pattern discovery algorithm. Contrary to the traditional frequent item discovery algorithm, here the frequent mobility sequences we discovered contain consecutive locations and the maximum mobility patterns are selected as the final result. The main idea of this algorithm is to discover a continuous trajectory, for which the support value is larger than δ. We first calculate each hotspot's support value and the set of mobility patterns with length-1 are generated. Then the mobility patterns with length-k are generated through mobility patterns with length-($k - 1$). The iteration is ended when the set of length-k is ϕ.

The main notations used in our method are listed in Table 1 and the pseudo-code of our algorithm is shown in Algorithm 1.

5.2. Variable-order Markov prediction algorithm considering temporal factors

The order-k Markov predictor assumes that the next location depends only on the last k locations. The history's length l is called the order of the Markov predictor. Some Markov predictors fix, in advance of the model creation, the value of l, presetting it in a constant k in order to reduce the size and complexity of the prediction model. These predictors are termed fixed length Markov predictors of order k [62]. However, we advocate to set the order of Markov predictors according to the mobility pattern of each individual.

Algorithm 1 Mobility pattern discovery.

```
Input: Support threshold: $\delta$
       Set of mobility trajectories: $T$
       Set of hotspots: $H$
Output: Set of mobility patterns: $P$
procedure (mobilityPatternDiscovery) ($\delta$, $T$, $H$)
    $k = 1$
    $C_k = \{h | h \in H\}$
    $P_k = \{h | h \in H \wedge supp(h) > \delta\}$
    $P = \{\}$
    Repeat
        $k = k + 1$
        for all mobility pattern $p_{k-1} \in P_{k-1}$ do
            for all frequent pattern $p_1 \in P_1$ do
                $C_k = \{c_k | c_k = p_{k-1} \cup p_1\}$
            end for
        end for
        for all trajectory $t \in T$ do
            $C_t = subset(C_k, t)$
            for all candidate $c \in C_t$ do
                $count(c) = count(c) + 1$
            end for
        end for
        $P_k = \{c \mid c \in C_k \wedge sup(c) > \delta\}$
        $P = \cup P_k$
    until $P_k = \phi$
Return $P$
end procedure
```

In this section, we present a variable-order Markov predictor considering temporal factors. In this predictor, the length-(n) mobility patterns we have discovered for each user are the current n movements. Also, the mobility patterns' length is different for each user. For a next location of a user we want to predict, we first seek for frequent patterns with the maximum length. If the prediction cannot be made with the maximum mobility pattern, we look for the maximum mobility patterns' sub-pattern, which indicates that the predictor's order is variable for each prediction with different users.

In addition, the temporal characteristic of mobility pattern indicates the time regularity of mobility pattern, which should be considered in our model. Here, the occurrence time distribution of mobility pattern $P(p_t)$ is measured by the following formula:

$$P(p_t) = \frac{\text{frequency of mobility pattern } p \text{ occurring in time slot } t}{\text{total frequency}}. \qquad (3)$$

Here, the time period is divided into hours, and each time slot equals to one hour. Let X_t be a random variable, and x_t a history hotspot in a specific trajectory that appears in time slot t. Then $P(X_t = x_t | \ldots)$ denotes the probability that X_t takes the value x_t. We calculate the score for each x_t by the following equation:

$$P(X_{t+1}{=}x_{t+1}){=}P(X_{t+1} = x_{t+1} | X_{t-k+1} = x_{t-k+1}, \ldots, X_t = x_t) \times P(p_t), \qquad (4)$$

where p_t starts with $< x_{t-k+1}, \ldots, x_t >$. In the final prediction, we set x_{t+1} with the maximum probability $P(X_{t+1} = x_{t+1})$.

5.3. Users similarity calculation algorithm

Notice that if the next location has never occurred in the history mobility patterns, the variable-order Markov prediction based

Table 2
Main notations of users similarity calculation algorithm.

Notation	Description
$sim(P, Q)$	The similarity value of two mobility patterns P and Q
$sim(u\|u')$	The relative similarity of u to u'
$LCS(P, Q)$	The longest common sequence of two mobility patterns P and Q
$lenLCS(P, Q)$	The length of two mobility patterns' longest common sequence

on mobility pattern makes no predictions. Under these circumstances, the mobility pattern of his geo-friend is taken into consideration. A geo-friend shares similar mobility patterns with the current user. When the user's own history mobility pattern cannot predict his next location, we use the mobility pattern of his geo-friends to predict his next location, which can improve the prediction accuracy.

Similarity based on mobility patterns is an indicator of two users' similarity in space, which could be quantified by applying the core idea of Collaborative Filtering (CF) algorithm, i.e., looking for users who share the same rating patterns with the active user (the user whom the prediction is for). CF, a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating), usually contains user-based and item-based recommendation. In this paper, we employ a user-based recommendation method to predict user mobility when the user's own mobility pattern cannot contribute to his/her prediction. Here we employ a similarity calculation algorithm based on user's mobility patterns to find the his/her geo-friends [63]. Some important definitions of the mobility pattern based similarity calculation algorithm are introduced in Table 2.

In addition, given two users u and u', function $\psi_{u,u'} : P_u \to P_{u'}$ is used to map a maximal mobility pattern of u to the most similar maximal mobility pattern in $P_{u'}$. Specially, for each $P_i \in P$, $\psi_{u,u'}(P_i) = \max_{Q_j \in Q} sim(P_i, Q_j)$.

Our main idea is to exploit the intuition that if user u is similar to user u', then any pattern of u will correspond to a similar pattern of user u'. Therefore, the similarity calculation consists of two steps. In the first step, each pair of sequence patterns from the given two maximal patterns sets respectively are compared and the result is called the pattern similarity between them. In the second step, the calculated pattern similarity values are combined in a specific way as the final value of user similarity. The specific steps are shown below.

5.3.1. Calculating similarity of mobility patterns

The mobility patterns in a pattern set contain duplicated information. If we compare two users' mobility patterns using the original pattern sets, some behavior will be used more than once. Therefore, maximal pattern sets are employed in computing user similarity. The similarity between two maximal mobility patterns is calculated based on the intuition that the more similar they are, the longer common pattern they share. The Longest Common Sequences (LCS) refer to their longest common patterns. The similarity between P and Q is calculated as follows:

$$sim(P, Q) = \frac{2 \times lenLCS(P, Q)}{len(P) + len(Q)}. \tag{5}$$

5.3.2. Calculating similarity of two users

For each user, we compute his relative similarity to others. The relative similarity of u to u', denoted by $sim(u|u')$, is calculated as the average weighted value of all the pattern similarity values of the identified most similar pattern pairs:

$$sim(u|u') = \frac{\sum_{P_i \in P} sim(P_i, \psi_{u,u'}(P_i)) \times \omega(P_i, \psi_{u,u'}(P_i))}{\sum_{P_i \in P} \omega(P_i, \psi_{u,u'}(P_i))}. \tag{6}$$

The weight function can be defined in different ways according to the various requirements of applications for user similarity. For instance, a user may be considered to be more similar to another as long as they share more common movements in some applications, while other applications may require two similar users to share more behaviors, the source data of which are only possessed by them. In this paper, we adopt the first interpretation. Thus

$$\omega(P, Q) = \frac{supp_u(P) + supp_{u'}(Q)}{2}. \tag{7}$$

We can also compute the relative similarity of u' to u, i.e., $sim(u'|u)$. As the relation of similarity should be symmetric, we calculate the average of the two relative similarity values as the similarity between users u and u':

$$sim(u, u') = \frac{sim(u|u') + sim(u'|u)}{2}. \tag{8}$$

The pseudo-code of mobility pattern based similarity calculation algorithm is shown in Algorithm 2.

Algorithm 2 User similarity calculation algorithm.

```
Input: Set of mobility patterns of user $u$: $P_u$
       Set of mobility patterns of user $u'$: $P_{u'}$
Output: Similarity of two users: $sim(u, u')$
procedure userSimilarityCalculation $(P_u, P_{u'})$
for all mobility pattern $p_u$ in $P_u$ do
    for all mobility pattern $p_{u'}$ in $P_{u'}$ do
        $sim(P_u, P_{u'}) = \frac{2 \times lenLCS(p_u, p_{u'})}{len(p_u) + len(p_{u'})}$
    end for
    $\psi_{u,u'}(p_u) = \max sim(p_u, p_{u'})$
end for
$sim(u|u') = \frac{\sum sim(p_u, \psi_{u,u'}(p_u)) \times \omega(p_u, \psi_{u,u'}(P_u))}{\sum \omega(p_u, \psi_{u,u'}(P_u))}$
for all mobility pattern $p_{u'}$ in $P_{u'}$ do
    for all mobility pattern $p_u$ in $P_u$ do
        $sim(P_{u'}, P_u) = \frac{2 \times lenLCS(p_{u'}, p_u)}{len(p_{u'}) + len(p_u)}$
    end for
    $\psi_{u',u}(p_{u'}) = \max sim(p_{u'}, p_u)$
end for
$sim(u'|u) = \frac{\sum sim(p_{u'}, \psi_{u',u}(p_{u'})) \times \omega(p_{u'}, \psi_{u',u}(P_{u'}))}{\sum \omega(p_{u'}, \psi_{u',u}(P_{u'}))}$
$sim(u, u') = \frac{sim(u|u') + sim(u'|u)}{2}$
end procedure
```

6. Performance evaluation of hybrid Markov-based model

The mobility prediction problem can be formalized as a contextual prediction problem where the future movements are assumed to depend only on the user context, which is characterized by a real dataset in this paper. We assume the existence of a gradually populated/trained knowledge base and try to compare the movement pattern of a certain object with stored information in order to predict its future locations. The assumption is based on the repetitive nature of human mobility: similar contexts might imply similar movements in the future. Therefore, simple and effective prediction algorithm can achieve high prediction accuracy by fully considering the temporal and spatial regularity in people's history trajectories. In this section, by using users' trajectories extracted from real data traffic of a 4G network, we evaluate the proposed model by analyzing the results of each applied method, and comparing the prediction accuracy at each stage of the model. In addition, the proposed model can also be further evaluated by comparing the prediction accuracy with other prediction algorithms.

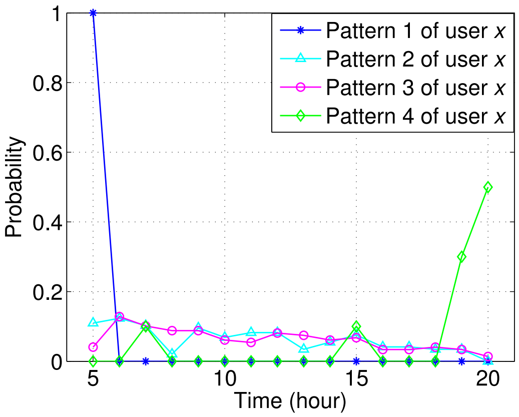


Fig. 9. Distribution of occurrence time for individual mobility patterns.

6.1. Verify the spatio-temporal characteristics of discovered mobility pattern

To improve the prediction accuracy of Markov-based algorithm, one of our key ideas is applying the variable-order Markov predictor. In this paper, the order and state transferring probability of variable-order Markov predictor are determined by the length of most frequent mobility pattern and occurrence time distribution of mobility pattern for current individual respectively. In order to examine the characteristics of mobility patterns from individuals' point of view, based on the obtained results from the first stage of proposed hybrid Markov-based Prediction Model, we draw the top four mobility patterns for user *X* per hour, as shown in Fig. 9. Pattern 1 always occurs in the morning and Pattern 4 usually occurs in the evening. Moreover, Pattern 2 and Pattern 3 both occur during the daytime.

Individual mobility pattern is a general description of human mobility. The individual mobility patterns provide significant information about mobile phone users' behaviors. We can conclude from the analysis above that they have strong spatio-temporal correlations. This means that each individual mobility pattern has its own most probable occurrence time. We can implement the mobility pattern discovery algorithm for each individual, to discover each individual's mobility patterns and their most probable occurrence time. Then, we can predict each individual's position at a specific time instead of just predicting the next hop of one user, which can improve the prediction accuracy when predicting user's future movements.

6.2. Examine the effectiveness of users similarity calculation

In order to predict user's future movement even when the new location has never been visited by him/her, we use his/her geo-friend's prediction result. In this paper, a CF based method is employed to calculate the similarity of users' trajectories based on users' mobility patterns. Here, reasonable questions to ask are, does every user have a geo-friend? How similar two geo-friends are? To answer above questions, in this section, we study the mobility characteristics of crowds in the city, and analyze the results of Users Similarity Calculation Algorithm with real dataset.

After applying our mobility pattern discovery algorithm to real dataset, Fig. 10 shows the top five mobility patterns of crowds for each hour. The mobility pattern of crowds changing with time indicates a large crowds' traffic among the specific locations at a specific time, which can contribute to the discovery of common mobility routes and users sharing similar movement patterns. We can conclude that there is a strong correlation between the occurrence time and daily user activities. The probability of these five most popular patterns' occurrence times is higher in the morning, which indicates that these common mobility patterns tend to occur at this time of the day.

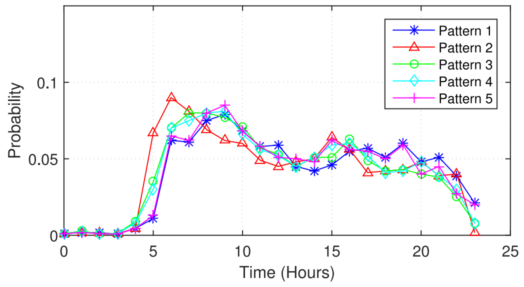


Fig. 10. Distribution of occurrence time for mobility pattern of crowds.

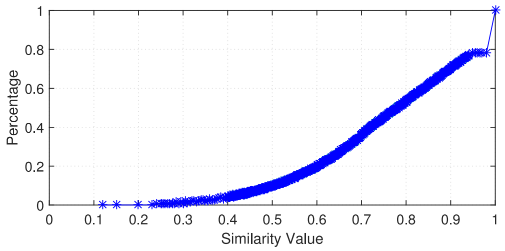


Fig. 11. CDF of geographic friend similarity.

Mobility pattern of crowds and the distribution of their time of occurrence suggests that humans tend to share common mobility patterns and that the occurrence time distributions tend to be similar with each other. This indicates that two users with high spatial similarity are capable of describing each other's mobility pattern.

Then, we apply the Users Similarity Calculation Algorithm to our dataset, and try to understand how well geo-friends can tell a user's mobility patterns. Firstly, we compute the similarity in space between each user. And then for each user we select his/her geo-friend, who shares the maximum number of common mobility pattern. As shown in Fig. 11, we find that among all user's geo-friends, nearly 50% of their similarity value can achieve 0.8 and 80% can achieve 0.6 , which indicates that the most similar geo-friend of a user can tell a lot about his/her mobility. Therefore, when a user's history mobility pattern cannot predict user's future movement, his geo-friend's mobility pattern can be taken into consideration.

6.3. Evaluate different stages of hybrid Markov-based model

The predictors are implemented according to the description of Section 4. The model returns a wrong prediction result if the predictor estimates a wrong next place. The most common statistical metric for assessing the capability of Markov to predict user's future location is the proportion of correct predictions. Here we define the accuracy of each predictor for each location to be the fraction of users for which the predictor correctly identified the next move. The prediction accuracy can be evaluated as follows:

$$prediction\,accuracy = \frac{\text{number of correct prediction}}{\text{number of all predictions}}. \tag{9}$$

Then we get the average prediction accuracy for each predictor at each location. Note that, Markov models will inevitably encounter situations where they are unable to make a prediction. If the predictor returns "no prediction", it is counted as an incorrect prediction.

In our experiment, the dataset is divided into two parts. Since the data covers three weeks (21 days), the data of the first two weeks is used to discover the mobility patterns and the data in the third week is used to make prediction.

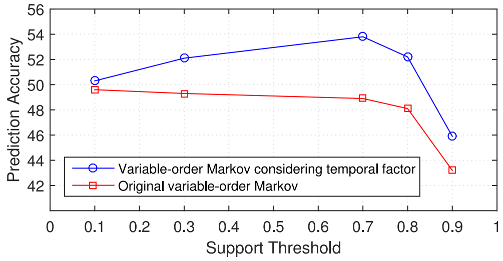


Fig. 12. The prediction accuracy varies with support threshold δ.

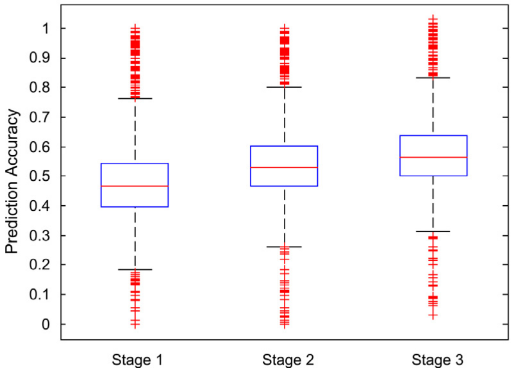


Fig. 13. Box plot of prediction accuracy for three stages of our hybrid Markov-based prediction model in three stages.

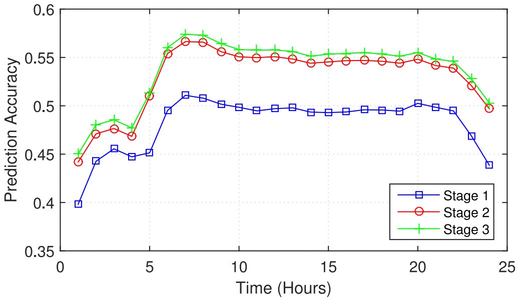


Fig. 14. Prediction accuracy of three stages of hybrid Markov-based prediction model evolving with time.

In our mobility pattern discovery algorithm, the support threshold δ may affect the number and the maximum length of the mobility patterns, which may affect the prediction accuracy finally. Therefore, we study the influence of δ over the overall prediction accuracy firstly. Fig. 12 shows the prediction accuracy of variable-order Markov changing with different value of the support threshold δ. From Fig. 12 we can see that the prediction accuracies of two algorithms both achieve the highest score when the support threshold δ equals to 0.7. Therefore, we set support threshold δ to 0.7, under which condition the prediction accuracy can achieve 48.81% (original variable-order Markov) and 53.42% (modified variable-order Markov considering temporal factor).

Fig. 13 illustrates the prediction accuracy under three stages of hybrid Markov-based prediction model, i.e., stage 1: original variable-order Markov, stage 2: variable-order Markov considering temporal factor, and stage 3: variable-order Markov considering temporal factor and geo-friend. By employing the mobility pattern of geo-friends, our hybrid Markov-based prediction model can reach 56.39% prediction accuracy. For the same dataset, the original Markov has 44.02% prediction accuracy.

In addition, we further examine the prediction accuracy at different time period in a day. Fig. 14 illustrates the distribution of prediction accuracy varying with time. From the figure we can conclude that the prediction accuracy reaches the highest during 6:00 a.m. and 7:00 a.m. It indicates that it is easier to predict human mobility in the morning, when a large number of people follow the similar commute patterns (as shown in Fig. 10).

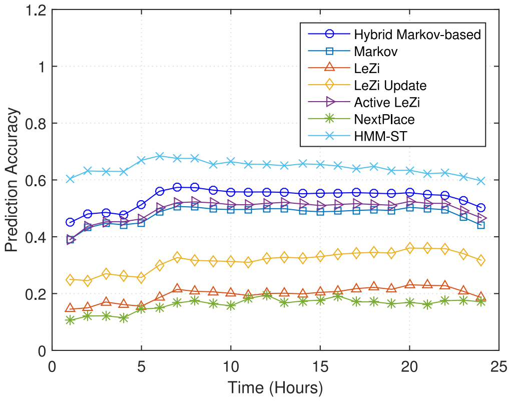


Fig. 15. Comparison of prediction accuracy between several prediction algorithms.

6.4. Compare the prediction accuracy between different prediction algorithms

In our previous studies, we applied varied many methods to predict users' future movements [14,15,18,55]. Prediction algorithms have their own characteristics, and the prediction results are different when applying the algorithm to different datasets, crowds, and individuals, even to the same dataset in different time period. This is because although human trajectories have a high degree of temporal and spatial regularity, their behaviors are diverse. Therefore, we aim at improving the prediction accuracy by discovering the mobility characteristics of users. Comparing the prediction accuracy of the proposed Hybrid Markov Prediction Model with other prediction algorithms, i.e., HMM-based (Hidden Markov Model-based) [15], NextPlace [15], variable-order Markov, and LZ family (LeZi family) [55], as shown in Fig. 15, we can clearly see that, HMM outperforms others. However, HMM has high time and spatial complexity, and tends to consume a lot of resource in production environment, which greatly restrict its practical application. The proposed Hybrid Markov-based Prediction Model in this paper has the best prediction accuracy among other evolved algorithms based on Markov models. As a result, we can confirm that our model is efficient, which drives us to further improve the prediction algorithm by considering the mobility characteristics of human trajectories.

7. Conclusions

In this paper, we aim at improving the prediction accuracy of Markov based prediction algorithms by considering non-Gaussian and spatio-temporal characteristics of a real dataset. In our experiments, by using real human trajectories extracted from data traffic of an LTE network, we analyze the characteristics of user's mobility with trajectories of 3474 people during 21 days. The interesting findings include: the indicators of user's mobility (trip distance, and radius of gyration) have non-Gaussian characteristics, people are frequently visiting a limited number of places, people are generally moving periodically within a bounded region, but are occasionally traveling long distance, and individuals tend to share similar mobility patterns in the city. The analysis results reveal that a strong temporal and spatial regularity exists in people's daily activities, which drives us to improve the prediction accuracy by focusing on discovering spatio-temporal regularity of people's movements.

Based on the analysis results, we proposed a hybrid Markov-based prediction model that contains three stages: mobility pattern discovery, variable-order Markov predictor considering temporal factors, and mobility pattern based users similarity calculation. In the first stage, a modified Apriori algorithm is applied to discover the frequent mobility patterns in user's trajectories, the length of which provides the value of order for the Markov predictor in the next stage. In the second stage, we apply a modified variable-order Markov algorithm to make the prediction. In this stage, the transition probability of each state is calculated by considering the occurrence probability and the occurrence time distribution of mobility patterns. In the third stage, in order to get the prediction result even when the current location has not occurred in previous trajectories, we consider the trajectories of "geo-friends" into account. We employ the Collaborative Filtering algorithm to find "geo-friends", who share similar mobility patterns with the current user in daily life. Experiment results show that prediction accuracy increases to 53.42% from 48.81% if we add the occurrence time distribution into the variable-order Markov. In addition, 56.39% prediction accuracy could be achieved by considering the mobility pattern of "geo-friends" when a context has not been previously seen. The proposed model outperforms other evolved algorithms based on Markov models we have applied in our previous studies.

For future work, some other special situations such as weekend and activity changes should also be considered in our model. In addition, extensive experimental evaluations should be conducted to compare our model with other classical prediction algorithms on different datasets. The quantitative and qualitative comparison in terms of prediction accuracy, time complexity, energy and resource consumption, is essential, which will guide us to apply the model to the most suitable production environment.

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China (61671078), Director Funds of Beijing Key Laboratory of Network System Architecture and Convergence (2017BKL-NSAC-ZJ-06), 111 Project of China (B08004), and EU FP7 IRSES MobileCloud Project (Grant No. 612212).

References

[1] A.R. Carrión, Contributions to the understanding of human mobility and its impact on the improvement of lightweight mobility prediction algorithms, Universidad Carlos III de Madrid, 2016.

[2] S. Akoush, A. Sameh, Mobile user movement prediction using bayesian learning for neural networks, in: Proceedings of the 2007 International Conference on Wireless Communications and Mobile Computing, ACM, 2007, pp. 191–196.

[3] S. Hoteit, S. Secci, S. Sobolevsky, C. Ratti, G. Pujolle, Estimating human trajectories and hotspots through mobile phone data, Comput. Netw. 64 (2014) 296–307.

[4] J.A. Alvarez-Garcia, J.A. Ortega, L. Gonzalez-Abril, F. Velasco, Trip destination prediction based on past gps log using a hidden markov model, Expert Syst. Appl. 37 (12) (2010) 8166–8171.

[5] W. Mathew, R. Raposo, B. Martins, Predicting future locations with hidden markov models, in: Proceedings of the 2012 ACM Conference on Ubiquitous Computing, ACM, 2012, pp. 911–918.

[6] J.J.-C. Ying, W.-C. Lee, V.S. Tseng, Mining geographic-temporal-semantic patterns in trajectories for location prediction, ACM Trans. Intell. Syst. Technol. 5 (1) (2013) 2.

[7] D. Ashbrook, T. Starner, Using gps to learn significant locations and predict movement across multiple users, Personal Ubiquitous Comput. 7 (5) (2003) 275–286.

[8] T. Anagnostopoulos, C. Anagnostopoulos, S. Hadjiefthymiades, Efficient location prediction in mobile cellular networks, Int. J. Wirel. Inf. Netw. 19 (2) (2012) 97–111.

[9] K. Laasonen, Clustering and prediction of mobile user routes from cellular data, in: European Conference on Principles of Data Mining and Knowledge Discovery, Springer, 2005, pp. 569–576.

[10] H. Si, Y. Wang, J. Yuan, X. Shan, Mobility prediction in cellular network using hidden markov model, in: 2010 7th IEEE Consumer Communications and Networking Conference, IEEE, 2010, pp. 1–5.

[11] A. Rodriguez-Carrion, C. Garcia-Rubio, C. Campo, Performance evaluation of lz-based location prediction algorithms in cellular networks, IEEE Commun. Lett. 14 (8) (2010) 707–709.

[12] A.J. Nicholson, B.D. Noble, Breadcrumbs: forecasting mobile connectivity, in: Proceedings of the 14th ACM International Conference on Mobile Computing and Networking, ACM, 2008, pp. 46–57.

[13] A. Rodriguez-Carrion, C. Garcia-Rubio, C. Campo, A. Cortés-Martín, E. Garcia-Lozano, P. Noriega-Vivas, Study of lz-based location prediction and its application to transportation recommender systems, Sensors 12 (6) (2012) 7496–7517.

[14] H. He, Y. Qiao, S. Gao, J. Yang, J. Guo, Prediction of user mobility pattern on a network traffic analysis platform, in: Proceedings of the 10th International Workshop on Mobility in the Evolving Internet Architecture, ACM, 2015, pp. 39–44.

[15] Q. Lv, Y. Di, Y. Qiao, Z. Lei, C. Dong, Spatial and temporal mobility analysis in lte mobile network, in: 2015 IEEE Wireless Communications and Networking Conference (WCNC), IEEE, 2015, pp. 795–800.

[16] D. Naboulsi, M. Fiore, S. Ribot, R. Stanica, Mobile Traffic Analysis: A Survey, Université de Lyon, INRIA, Grenoble-Rhône-Alpes, 2015 Ph.D. thesis.

[17] X. Zhou, Z. Zhao, R. Li, Y. Zhou, J. Palicot, H. Zhang, Human mobility patterns in cellular networks, IEEE Commun. Lett. 17 (10) (2013) 1877–1880.

[18] Y. Qiao, Y. Cheng, J. Yang, J. Liu, N. Kato, A mobility analytical framework for big mobile data in densely populated area, IEEE Trans. Veh. Technol. 66 (2) (2017) 1443–1455.

[19] C.M. Schneider, V. Belik, T. Couronné, Z. Smoreda, M.C. González, Unravelling daily human mobility motifs, J. R. Soc. Interface 10 (84) (2013). Xxxviii.

[20] C. Song, T. Koren, P. Wang, A.L. Barabási, Modelling the scaling properties of human mobility, Nat. Phys. 6 (10) (2010) 818–823.

[21] M.C. Gonzalez, C.A. Hidalgo, A.L. Barabasi, Understanding individual human mobility patterns, Nature 453 (7196) (2008) 779–782.

[22] D. Brockmann, L. Hufnagel, T. Geisel, The scaling laws of human travel, Nature 439 (7075) (2006) 462–465.

[23] J. Jeong, M. Leconte, A. Proutiere. Human Mobility Prediction Using Non-parametric Bayesian Model, arXiv preprint arXiv:1507.03292.

[24] S.-B. Cho, Exploiting machine learning techniques for location recognition and prediction with smartphone logs, Neurocomputing 176 (2016) 98–106.

[25] Y.-J. Kim, S.-B. Cho, A hmm-based location prediction framework with location recognizer combining k-nearest neighbor and multiple decision trees, in: International Conference on Hybrid Artificial Intelligence Systems, Springer, 2013, pp. 618–628.

[26] Z. Ma, P.K. Rana, J. Taghia, M. Flierl, A. Leijon, Bayesian estimation of dirichlet mixture model with variational inference, Pattern Recogn. 47 (9) (2014) 3143–3157.

[27] Z. Ma, A. Leijon, Bayesian estimation of beta mixture models with variational inference, IEEE Trans. Pattern Anal. Mach. Intell. 33 (11) (2011) 2160–2173.

[28] Z. Ma, J.-H. Xue, A. Leijon, Z.-H. Tan, Z. Yang, J. Guo, Decorrelation of neutralvector variables: theory and applications, IEEE Trans. Neural Netw. Learn. Syst. PP (99) (2016) 1–15.

[29] Z. Ma, S. Chatterjee, W.B. Kleijn, J. Guo, Dirichlet mixture modeling to estimate an empirical lower bound for lsf quantization, Signal Process. 104 (2014) 291–295.

[30] Z. Ma, A. Leijon, W.B. Kleijn, Vector quantization of lsf parameters with a mixture of dirichlet distributions, IEEE Trans. Audio, Speech, Lang. Process. 21 (9) (2013) 1777–1790.

[31] J. Petzold, F. Bagci, W. Trumler, T. Ungerer, Next location prediction within a smart office building 577 (2005) 69.

[32] M. Chen, X. Yu, Y. Liu, Mining moving patterns for predicting next location, Inf. Syst. 54 (2015) 156–168.

[33] S. Qiao, D. Shen, X. Wang, N. Han, W. Zhu, A self-adaptive parameter selection trajectory prediction approach via hidden markov models, IEEE Trans. Intell. Transp. Syst. 16 (1) (2015) 284–296.

[34] S.K. Pulliyakode, S. Kalyani, A modified ppm algorithm for online sequence prediction using short data records, IEEE Commun. Lett. 19 (3) (2015) 423–426.

[35] J.G. Cleary, W.J. Teahan, Unbounded length contexts for ppm, Comput. J. 40 (2 and 3) (1997) 67–75.

[36] J.A. Torkestani, Mobility prediction in mobile wireless networks, J. Netw. Comput. Appl. 35 (5) (2012) 1633–1645.

[37] Y. Chon, H. Shin, E. Talipov, H. Cha, Evaluating mobility models for temporal prediction with high-granularity mobility data, in: IEEE International Conference on Pervasive Computing and Communications (PerCom), 2012, IEEE, 2012, pp. 206–212.

[38] N. Bui, M. Cesana, S.A. Hosseini, Q. Liao, I. Malanchini, J. Widmer. Anticipatory networking in future generation mobile networks: a survey, arXiv preprint arXiv:1606.00191.

[39] M. Žitnik, B. Zupan, Gene network inference by fusing data from diverse distributions, Bioinformatics 31 (12) (2015) i230–i239.

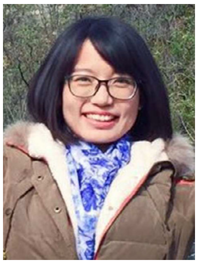
[40] H. Aghighi, J. Trinder, K. Wang, Y. Tarabalka, S. Lim, Smoothing parameter estimation for markov random field classification of non-gaussian distribution image, ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci. 2 (7) (2014) 1.

[41] C. Qian, W. Zhou, S.K. Fong, K.C. Leong, Two approaches for statistical prediction of non-gaussian climate extremes: a case study of macao hot extremes during 1912–2012, J. Clim. 28 (2) (2015) 623–636.

[42] B.K. Chaurasia, S. Verma, G. Tomar3, Gaussian profile based vehicular mobility modeling, AJMSCAHS 1 (2) (2011) 59–76.
[43] T.M.T. Do, O. Dousse, M. Miettinen, D. Gatica-Perez, A probabilistic kernel method for human mobility prediction with smartphones, Pervasive Mobile Comput. 20 (2015) 13–28.
[44] S. Zhang, Y. Zhang, J. Zhu, Residual life prediction based ondynamic weighted markov model and particle filtering, J. Intell. Manuf. (2015) 1–9, doi:10.1007/s10845-015-1127-4.
[45] M.F. Steel, M. Fuentes, Non-gaussian and nonparametric models for continuous spatial data, Handbook of Spatial Statistics (2010) 149–167.
[46] L. Song, D. Kotz, R. Jain, X. He, Evaluating next-cell predictors with extensive wi-fi mobility data, IEEE Trans. Mobile Comput. 5 (12) (2006) 1633–1649.
[47] L. Song, U. Deshpande, U.C. Kozat, D. Kotz, R. Jain, Predictability of wlan mobility and its effects on bandwidth provisioning.,, in: INFOCOM, 2006.
[48] M.H. Sun, D.M. Blough, Mobility prediction using future knowledge, in: Proceedings of the 10th ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems, ACM, 2007, pp. 235–239.
[49] B.D. Ziebart, A.L. Maas, A.K. Dey, J.A. Bagnell, Navigate like a cabbie: probabilistic reasoning from observed context-aware behavior, in: Proceedings of the 10th International Conference on Ubiquitous Computing, ACM, 2008, pp. 322–331.
[50] V. Gogate, R. Dechter, B. Bidyuk, C. Rindt, J. Marca, Modeling Transportation Routines Using Hybrid Dynamic Mixed Networks, arxiv preprint arxiv:1207.1384.
[51] D. Barth, S. Bellahsene, L. Kloul, Combining local and global profiles for mobility prediction in lte femtocells, in: Proceedings of the 15th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, ACM, 2012, pp. 333–342.
[52] S. Bellahsene, L. Kloul, A new markov-based mobility prediction algorithm for mobile networks, in: European Performance Engineering Workshop, Springer, 2010, pp. 37–50.
[53] H. Bapierre, G. Groh, S. Theiner, A variable order markov model approach for mobility prediction, Pervasive Comput. (2011) 8–16.
[54] H. Xiong, D. Zhang, D. Zhang, V. Gauthier, K. Yang, M. Becker, Mpaas: Mobility prediction as a service in telecom cloud, Inf. Syst. Front. 16 (1) (2014) 59–75.
[55] Y. Qiao, J. Yang, H. He, Y. Cheng, Z. Ma, User location prediction with energy efficiency model in the long term-evolution network, Int. J. Commun. Syst. 29 (14) (2016) 2169–2187.
[56] S. Bellahsene, L. Kloul, D. Barth, A hierarchical prediction model for two nodes-based ip mobile networks, in: Proceedings of the 12th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems, ACM, 2009, pp. 173–180.
[57] G. Gidófalvi, F. Dong, When and where next: individual mobility prediction, in: Proceedings of the First ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems, ACM, 2012, pp. 57–64.
[58] M. Chen, Y. Liu, X. Yu, Nlpmm: a next location predictor with markov modeling, in: Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer, 2014, pp. 186–197.
[59] J. Yang, X. Zhang, Y. Qiao, Z. Fadlullah, N. Kato, Global and individual mobility pattern discovery based on hotspots, in: 2015 IEEE International Conference on Communications (ICC), IEEE, 2015, pp. 5577–5582.
[60] T. Louail, M. Lenormand, O.G.C. Ros, M. Picornell, R. Herranz, E. Frias-Martinez, J.J. Ramasco, M. Barthelemy, From mobile phone data to the spatial structure of cities, Sci. Rep. 4 (2014) 5276, doi:10.1038/srep05276.
[61] Y. Zheng, Trajectory data mining: an overview, ACM Trans. Intell. Syst. Technol. (TIST) 6 (3) (2015) 29.
[62] L.T. Yang, Mobile Intelligence, Vol. 69, John Wiley & Sons, 2010.
[63] X. Chen, J. Pang, R. Xue, Constructing and comparing user mobility profiles, ACM Trans. Web (TWEB) 8 (4) (2014) 21.

Yuanyuan Qiao is a lecturer in the School of Information and Communication Engineering, BUPT. She received her B.E. degree from Xidian University in 2009, and received her Ph.D degree from BUPT in 2014. Her research focuses on traffic measurement and classification, mobile Internet traffic analysis and big data analytics.

Zhongwei Si received the Ph.D. degree from the KTH Royal Institute of Technology, Sweden, in 2013. In 2013, she joined the Beijing University of Posts and Telecommunications, where she is currently an Associate Professor. Her research interests include wireless communication, information theory, and data mining.



Yanting Zhang is a doctoral student in the School of Information and Communication Engineering, BUPT, and received her B.E. degree in communication engineering from BUPT in 2015. She is engaged in the research of human mobility analysis and big data analytics.

Fehmi Ben Abdesslem received his M.Sc and Ph.D. from the University of Paris 6 in 2008, before working as a research associate at the University of St Andrews, and at the University of Cambridge. He has then been awarded a Marie-Curie research fellowship from the European Commission (ERCIM) to join SICS, and is now a permanent Senior Research Scientist at the Decisions Networks and Analytics laboratory.

Xinyu Zhang received her B.E., and M.E. degrees from Beijing University of Posts and Telecommunications, China in 2013, and 2016 respectively. She is now a engineer of FreeWheel in Beijing.

Jie Yang received her B.E., M.E., and Ph.D. degrees from Beijing University of Posts and Telecommunications, China in 1993, 1999, and 2007 respectively. She is now a professor and deputy dean of School of Information and Communication Engineering, BUPT. Her current research interests include broadband network traffic monitoring, user behavior analysis, big data analysis in Internet and Telecom, etc. She has published several papers on international magazines and conferences including IEEE JSAC, IEEE Trans. on Wireless Communications and IEEE Trans. on Parallel and Distributed Systems. Also, she is the Vice Program Committee Co-Chairs of IEEE IC-NIDC 2014 and 2012.