An Improved Markov Method for Prediction of User Mobility

Yihang Cheng, Yuanyuan Qiao, and Jie Yang
Beijing Laboratory of Advanced Information Networks of BUPT, Beijing 100876, China.
Beijing Key Laboratory of Network System Architecture and Convergence of BUPT, Beijing 100876, China.
Beijing University of Posts and Telecommunications, Beijing 100876, China.
Email: {mrcheng, yyqiao}@bupt.edu.cn

***Abstract*—The developments of Information and Communication Technology (ICT) and Internet of Things (IoT) are being used to enhance quality, performance and interactivity of urban services. Benefited from the widespread adoption of mobile devices, we can collect amount of mobile data for user mobility analysis. Mining hidden information from users' mobile data is important for builders of smart city to provide better location-based service. This paper focuses on two classical domain-independent prediction models and one improved Markov model that are capable of estimating the next location. By using 27-day-long traffic data of mobile network, we extract trajectories of 4914 individuals for experiments. We find that the original Markov algorithm has a better performance in resource consumption than LZ family algorithms, but its prediction accuracy is lower than prediction accuracy of LeZi Update and Active LeZi algorithm. In order to improve the prediction accuracy of Markov and overcome drawbacks of traditional prediction algorithms, we present a new method based on Markov, which considers both temporal and spatial factors. Extensive experiments demonstrate our improved method has a better performance in location prediction. In addition, we further study the relationship between prediction accuracy and trajectory's regularity, to identify the most suitable prediction algorithm for a trajectory.**

I. Introduction

It is well known that the movement of people shows a high degree of repetition since they tend to visit regular places. Study of human mobility gains mainstream popularity in recent years, which helps to bring many location-based services for urban human. Those services, which can improve the citizens quality of life, are parts of a smart city. Nowadays, smart devices bring us the ubiquitous mobile Internet access. The movements of people could be sensed and easily collected by mobile phone, which generates large volumes of mobility data, such as Call Detail Records (CDRs) [1], Global Positioning System (GPS) tracks [2], and data traffic from 2G/3G/4G data networks [3]. Location prediction, which has been applied in many fields, is a hot topic in recent years. Since most of location-based services require accurate or approximate position of user, predicting user's next location could allow service providers to provide services for user in advance, which may help to improve the quality of services and user experience.

For a smart city, builders use digital technologies and information analysis algorithms to enhance services performance, engaging more effectively and actively with its citizens. In the area of location prediction, several algorithms have been proposed, including Markov models [4], [5], LZ family algorithms [6], [7], Bayesian networks [8], text compression-based techniques [9] and neural networks [10]. As we know, algorithms focuses on prediction are based on known properties learned from the training data. And many prediction algorithms have desirable prediction performance if there is enough memory and time for prediction. However, many services installed in mobile devices have limits on memory and time. So prediction algorithms with high resource consumption are of little practical value. It becomes important to find a prediction algorithm with high prediction accuracy and low resource consumption. Among those methods, Markov and LZ family algorithms get more attentions because of their low realization complexity and high prediction accuracy [11]. Depending on the diverse sources of experimental data sets, those two kinds of methods have different prediction performance. However, comparing with Markov algorithm, LZ family algorithms need to maintain a prediction tree in the process of prediction, which leads to a higher resource consumption [7]. As many applications used on mobile devices have restrictions on resource consumption, using Markov-based algorithm to predict users' future locations becomes a better choice.

The purpose of this paper is to improve the prediction accuracy by modifying the original Markov algorithm. Most of the prediction algorithms aforementioned only take into account trajectory's spatial factor during the process of model building and location prediction. However, they miss many other available factors of trajectory.In this paper, by considering the temporal factor and improving the original Markov algorithm, we propose an improved Markov algorithm. Comparing with prediction accuracy of original Markov, prediction accuracy of the new algorithm has a 6% increases. The key contributions and some interesting findings of our paper are as follows:

- We propose a new location prediction method that based on Markov algorithm, and consider both spatial and temporal factors while predicting. By comparing with Markov, this method can significant improve the prediction accuracy. Besides, it overcomes many drawbacks of traditional algorithms. For example, it can make the prediction even when the mobility pattern has not occurred in the history trajectory.
- In order to choose the most suitable prediction algorithm

978-3-901882-85-2 © 2016 IFIP

for a trajectory, we study the relationship between trajectory regularity and prediction accuracy. Through the experiments, we find that the more regular the trajectory is, the higher the prediction accuracy is. Although the Improved Markov which considers temporal factor performs better than original Markov, but the gap between the two algorithms' prediction accuracy is narrowing with the decreasing of trajectory regularity. This may be because that Improved Markov considers too much unnecessary mobility information for trajectories with low regularity. When we predict future locations for trajectories with low regularity, using the original Markov algorithm may be a better choice.
- Our work focuses on the mobile Internet data, instead of GPS or CDR data. By using nearly 19TB traffic data bills that cover 3 million users, the data can describe users' mobility more comprehensive. Different from GPS and CDR data, passively collecting human movement trajectories while he/she is accessing to mobile Internet has lots of advantages: high cost efficiency, low energy consumption, covering a wide range and a large number of people, and with fine time granularity.

The remainder of this paper is structured as follows. We introduce related works in the field of user location prediction in Section II. In Section III, three kinds of prediction algorithms, including Markov, LZ family algorithms and Improved Markov algorithm, are introduced. Experimental results will be shown in Section IV. Finally, conclusion and future work are presented in Section V.

II. Related work

Recent studies have found that trips mostly consist of regular travels, such as commuting to work or grocery shopping, following the daily circadian rhythm [12]–[14]. So it's possible to predict users' future movements according to their history trajectories. Researchers have found that the theoretical maximum predictability can be higher than 85%, when measuring the uncertainties of movements using entropy [15] and considering both the frequencies and temporal correlations of individual trajectories [16].

Previous studies have proposed many different methods to forecast users' future locations, including Markov [4], [5], Bayesian networks [8], Hidden Markov Models [17] and LZ family algorithms [6], [7] and so on. Most of those methods only consider spatial factor to predict future movements, missing other hidden information of trajectories. For example, based on a Global System for Mobile Communication (GSM) data set which contains 95 different users, researchers in [6] focused on LZ family algorithms (LZ, LeZi Update and Active LeZi) to estimate the next location. However, studies have found that time significantly impact randomness, size and probability distribution of people's movements, and the prediction accuracy can be increased by considering the temporal factor [18], [19]. Authors in [4] improved the prediction accuracy to 59.2% from 46.8% by applying a time-based Markov algorithm.

In addition, there are more and more location-based services that are applied on mobile devices [6]. Since there are restrictions on resource consumption on the smart phone, many traditional prediction methods may sacrifice prediction accuracy for a higher prediction speed. Researchers in [20] present a method which uses the past trajectory of the object and combines it with movement rules discovered in the moving objects database. The method can be performed offline which can improve the prediction speed. In a previous study [7], we examined the prediction accuracy and resource consumption of Markov and LZ family algorithms. We found that although Active LeZi outperformed Markov at prediction accuracy, but Markov performed better at prediction speed.

- Independent of time: For location prediction, both of Markov and LZ family algorithms are independent of time. But we consider the temporal factor, which has already been taken into account in [4], is an important feature that may affect user's mobility.
- Many unpredictable situations: Markov and LZ algorithm can't make the prediction if a mobility pattern has not occurred in the history trajectory [7].
- High computer resources: The LZ family algorithms take too much time and computer resources on prediction, because they need to maintain prediction trees during prediction [7].

In order to overcome above drawbacks, this paper aims at proposing a new location prediction method with low complexity and high prediction accuracy, which could be apply to environment with limited resources, such as applications of mobile phone. Both spatial and temporal factors are considered by this new method. As the original Markov and LZ algorithms don't consider many hidden information of trajectories, they can not make a prediction for patterns that haven't appeared in the history [4], [6]. However, in this situation, the new method can also make location prediction because that it considers the trajectories more comprehensive. What's more, in order to find the predictable of trajectories and the most suitable prediction method for a trajectory, we future research the relationship between trajectory regularity and prediction accuracy.

III. Methodology

In this section, our Improved Markov algorithm, which is modified from Markov algorithm, is presented in detail.

A. Time-based Markov Algorithm

Here we propose a Time-based Markov algorithm to overcome the first drawback (i.e., Independent of time) of Markov and LZ methods. User's history trajectory consists of time-location pairs like (l, t), which means user visit location l at time t. The k-order Time-based Markov can be tailored as below:

Step 1: Build user's everyday trajectory T_{day_i} (the trajectory of i-th day, where $i = 1, 2, \dots$). Then, divide T_{day_i} into 24 time intervals, each of which lasts one hour long and contains a series of locations like "$(l_1, t_1), (l_2, t_2), (l_1, t_3), (l_1, t_4)$".

Step 2: For each time interval, select the most frequently location, which appeared more times in that time interval, as the interval's representative location. For example, we select l_1 as the representative location of "$(l_1, t_1), (l_2, t_2), (l_1, t_3), (l_1, t_4)$". Then we can rebuild everyday's trajectory with 24 selected locations.
Step 3: Select the last k locations of trajectory as a mobility pattern L_k, where $1 \leq k \leq 24$, and select a location l_x that we grant as the next interval's location. L_k and l_x make up the prediction pattern $L_x = L_k + l_x$.
Step 4: Caculate the number of times N_{L_k} and N_{L_x} that pattern L_k and L_x have occurred in the history. And if there are n days' trajectories and location l_x have occurred m times at the next time interval in the history, where $n \geq 1$ and $0 \leq m \leq n$. We can caculate the probability that the next location of user is l_x by the formula bellow:

$$p(l_x) = 2^{\frac{m}{n}} \times \frac{N_{L_x}}{N_{L_k}} \tag{1}$$

Comparing with Markov algorithm, the Time-based Markov considers time parameter, in the form of a weighted factor $2^{\frac{m}{n}}$, as a influence factor of the prediction probability. Results of contrast experiments will be given in Section IV.

B. Improved Markov Algorithm

Time-based Markov algorithm overcome the "Independent of time" drawback of Markov by considering the temporal factor. However, there are still many drawbacks that make performance of Markov bad. For example, when we want to use a 3-order Markov or 3-order Time-based Markov algorithm to predict the next location of user with history trajectory "$l_1, l_2, l_3, l_1, l_2, l_4, l_1, l_2, l_3, l_3, l_1, l_2$", we find that all the assumed next location l_x's appeared probability are 0. It's because that any L_x with $L_k = l_3 l_1 l_2$ hasn't occurred in the history. But for a 2-order Markov or 2-order Time-based Markov predictor, the probability can be caculated because that the L_x with $L_k = l_1 l_2$ has occurred in the history. Using 2-order Markov, the prediction probability of "l_1, l_2, l_3, l_4" are "$\frac{2}{3}$, 0, 0, $\frac{1}{3}$" respectively. In view of this, we propose our Improved Markov algorithm by counting the low order probability. The k-order Improved Markov method is presented in detail as below:
Step 1 and ***Step 2*** are the same as Time-based Markov.
Step 3 Select the next location as l_x and let the prediction probability $p(l_x) = 0$. Let $i = k$.
Step 3: Select the last i locations of trajectory as a mobility pattern L_i, where $1 \leq i \leq 24$. L_i and l_x make up the prediction pattern L_x.
Step 4: Caculate the times N_{L_i} and N_{L_x} that pattern L_i and L_x have occurred in the history. Then, add the ratio $\frac{N_{L_x}}{N_{L_i}}$ to $p(l_x)$ with a weighted factor $g(i) = 2^{i-1}$. So the $p(l_x)$ changes to $p(l_x) + \frac{g(i) N_{L_x}}{N_{L_i}}$. If the $i > 0$, let $i = i - 1$. And return to *Step 3* to caculate the prediction probability again.
Step 5: If there are n days' trajectories and location l_x have occurred m times at the next time interval in the history, where $n \geq 1$ and $0 \leq m \leq n$. We can caculate the probability that the next location of user is l_x by the formula bellow:

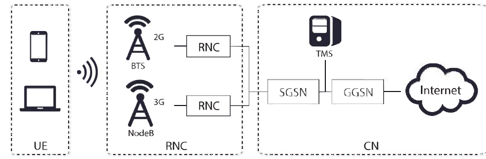


Fig. 1: Architecture of 2G/3G networks and the deployment of TMS.

$$p(l_x) = 2^{\frac{m}{n}} \times \sum_{i=1}^{k} \frac{g(i) N_{L_x}}{N_{L_i}} \tag{2}$$

Here, $g(i)$ is employed to assign bigger weight for a longer prediction pattern.

We can see that Improved Markov considers more information than the Markov and Time-based Markov algorithm. Based on our Improved Markov algorithm, we can give a more accurate prediction for user's future location at next time interval than original Markov algorithm. Experimental results will be shown at the next section.

IV. EXPERIMENTAL RESULTS

In this section, we first briefly introduce how the experimental data is collected from mobile Internet. Secondly, we evaluate performance of the three kinds of prediction algorithms which are introduced in Section III from the prediction accuracy. At last, the relationship between prediction accuracy and trajectory's regularity is studied. All the experiments are performed under the specified conditions of the data set.

A. Data Collection

By the help of our self-developed Traffic Monitoring System (TMS, which is a kind of network probe that can offer real-time monitoring without any overhead), we collect data traffic, which is extracted as flow records, from real mobile networks. Here, we define 'flow' as bidirectional data transmission at the usual 5-tuple source IP, destination IP, source port, destination port, and transport protocol within a certain period of 64 seconds. As shown in Fig.1, there are three major components in 2G/3G networks: User Equipment (UE), Radio Access Network (RAN) and Core Network (CN). UE is a terminal equipment that user uses to connect with communication networks, such as cellphones, laptop computers, electronic paper books, or other devices that access to the Internet through cellular networks. RAN, which consists of transceiver stations (Base Transceiver Station (BTS) or Node-B), establishes the connection between UE and CN. It receives data traffic from UE (or CN) and then sends the traffic to CN (or UE). In 2G/3G networks, the CN consists of two kinds of node: Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). The SGSN establishes a tunnel with a GGSN to provide connectivity to Internet.

TABLE I: The average prediction accuracy of four prediction algorithms.

Prediction algorithm	Prediction accuracy
Markov	0.294
LZ	0.299
LeZi Update	0.358
Active LeZi	0.361

We deploy TMSs between RAN and CN to collect mobile Internet traffic and store it as flow record. Both the uplink and downlink Internet Protocol (IP) packets are collected. And the packets will be decoded to text bills that contain the connected transceiver stations' IP, user identification (ID) and time stamp.

In this paper, we use the data set collected from a Chinese 2G/3G service provider from July 25, 2015 to August 20, 2015. The data set covers over 3 million users in northern China, contains nearly 19 TB of traffic bills and covers more than 20 thousand cellular towers. Note that some users generate few records or don't move frequently, we select trajectories of 4914 users who connect the mobile networks more than 1,500 times as our experimental data.

B. Prediction Performance

In this part, by applying original Markov algorithm, LZ family algorithms and the Improved Markov algorithm, we first evaluate the performance of different prediction algorithms. Then, we test the performance of each algorithm on different kinds of user groups with distinct regularity of trajectory.

1) Prediction Accuracy: As we know, probability of correct predictions is the most common statistical metric for evaluating performance of location prediction algorithms. There are two kinds of prediction accuracy: individual's prediction accuracy, and all users' prediction accuracy. For individual, the probability of correct predictions is the value that the number of right predictions divided by the total number of locations. Since our users' history trajectories are extracted from a 27-day-long data set and everyday's trajectory consists of 24 locations. So each user's history trajectory has 648 locations. For all users, we define the accuracy of each predictor for each location to be the fraction of users for which the algorithm correctly identified the next location. Then, we obtain the average prediction accuracy for each predictor at each location.

Firstly, we run each user's trace independently with original Markov, LZ, LeZi Update and Active LeZi algorithms, using the average prediction accuracy to see algorithms' performance. As shown in Table I, LeZi Update's performance is similar with Active LeZi's, and higher than the other two. To compare the prediction performance of different algorithms and quantify the difference, we subtract one algorithm's prediction accuracy by another algorithm's prediction accuracy for all users. As shown in Fig.2, it can be seen from the fitting line that LeZi Update's accuracy value is 6.07% higher than Markov's accuracy value on average. It may be because of that LeZi Update and Active LeZi consider more information when predicting.

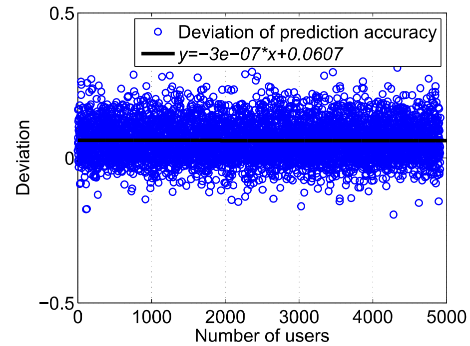


Fig. 2: The distribution of prediction accuracy's deviation between LeZi Update and Markov for all users.

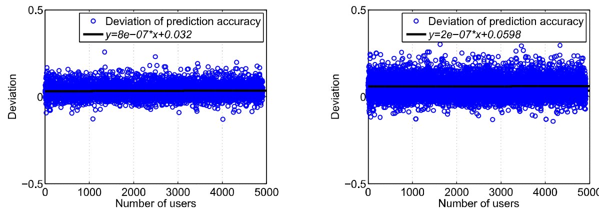
TABLE II: The prediction time and average resource consumption (calculated by the method in [7]) of a sample user's 1500 times of prediction.

Prediction algorithm	Prediction time	Resource consumption
Markov	1.39s	24.55
Improved Markov	3.97s	177.68
LeZi Update	16.19s	766.44
Active LeZi	32.04s	1097.18

In order to explore the influence of temporal feature on location prediction, we perform experiments with the Time-based Markov and the Improved Markov algorithm on all users. Results are shown in Fig.3. From the figure, we can see that prediction accuracy value of Time-based Markov is 3.2% higher than original Markov's. And Improved Markov's accuracy value can get a nearly 6% increase, which is similar with LeZi Update's and Active LeZi's. In addition, we further study the distribution of prediction accuracy for the three prediction algorithm, as shown in Fig.4. Improved Markov algorithm outperforms others that 60.6% of users achieved 40% accuracy, and in the case of Markov and Time-based Markov algorithms, the percentage of users is 38.6% and 46.6%, respectively. So we can have a conclusion that Improved Markov can significantly improve the prediction accuracy value. In addition, by comparing the prediction time and resource consumption of algorithms, we find the Improved Markov has a better performance than LZ family algorithms. The result is shown in Table II.

2) Prediction Performance Among Different Trajectory Classes: As we know, there are many feathers that influence the accuracy value of prediction algorithms, such as users' mobile abilities [21], time interval of prediction [4], trajectory's regularity [22] and so on. In Fig.5, we draw all users' accuracy value of each movement based on 27 days' trajectories. It can be found that prediction results during the nighttime are better than results during the daytime. It's because that trajectories' regularities during the daytime are worse than them during the nighttime when users are rarely move.

In order to describe the diversity of trajectory's regularity,



(a) Deviation of prediction accuracy between Markov and Time-based Markov

(b) Deviation of prediction accuracy between Markov and Improved Markov

Fig. 3: The distribution of prediction accuracy's deviation between two Markov-based algorithms for all users.

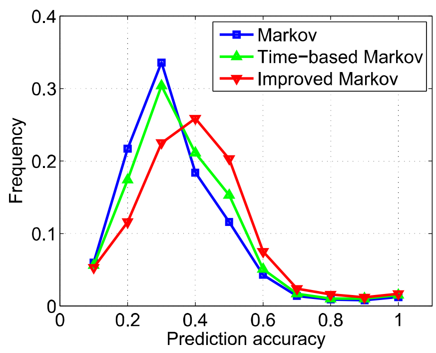


Fig. 4: The distribution of prediction accuracy of three kinds of Markov algorithms for users.

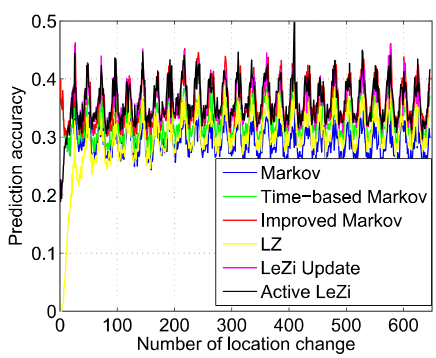
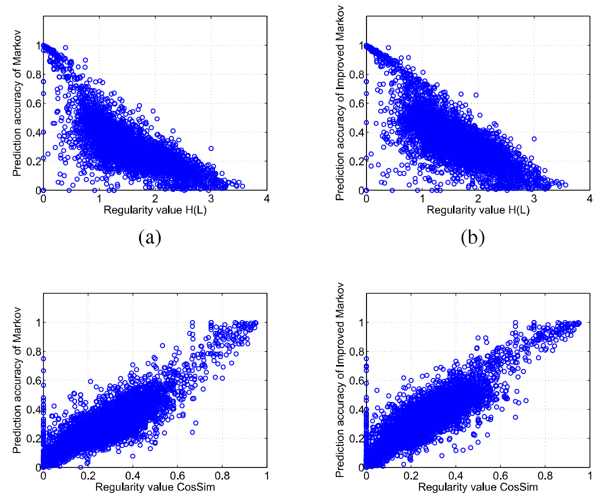


Fig. 5: The average accuracy of each algorithm for each location over the users' histories.

we use entropy $H(L)$ and cosine similarity $CosSim(L)$ as the metrics. The bigger the entropy value is, the more locations the user visits and the less regular the trajectory is. And the bigger the cosine similarity value is, the more regular the trajectory is. The two metrics are defined as follows:

$$H(L) = -\sum_{i=1}^{n} p(l_i) log_2 p(l_i) \tag{3}$$



(a) (b) (c) (d)

Fig. 6: The average prediction accuracy of Markov and Improved Markov for users with diverse trajectory's regularity value (represented by $H(L)$ and $CosSim(L)$).

$$CosSim(L) = \frac{1}{m} \sum_{i=1}^{n} cos(L_i, L_{i+1}) \tag{4}$$

where $n \geq 1$ is the number of different locations that user has visited, $p(l_i) \in (0, 1)$ is the probability for user staying in a certain place l_i, m is the number of everyday' trajectory L_i (in this paper, m equals to 27) and $cos(L_i, L_{i+1}) \in (-1, 1)$ is the cosine of vector angle. In Fig.6, the influences of trajectory's regularity (represented by $H(L)$ and $CosSim(L)$) on the probability distribution are presented. We can see that different algorithms' distributions of prediction accuracy have the same tendency. The prediction accuracy is higher with a lower entropy value and a higher cosine similarity value. The results indicate that the more regular the trajectory is, the higher the prediction accuracy is.

In addition, we examine Improved Markov and Markov by focusing on all users to discover the relationship between trajectory regularity and prediction accuracy. Because the Improved Markov considers more trajectory's information than Markov, its performance may be worse when the trajectory's regularity is low (which shows that Improved Markov may achieve a bad prediction performance when considers many unnecessary information). As shown in Fig.7, when the cosine similarity value is big (or entropy value is small), the prediction accuracy of Improved Markov is significant higher than Markov's. However, with the decrease of regularity, the gap between the two algorithms' prediction accuracy is narrowing as shown by the trend line (the ratio of prediction accuracy is increasing with the increasing of regularity value). We can suppose that the effect of temporal factor of trajectory becomes little important when trajectory's regularity is low.

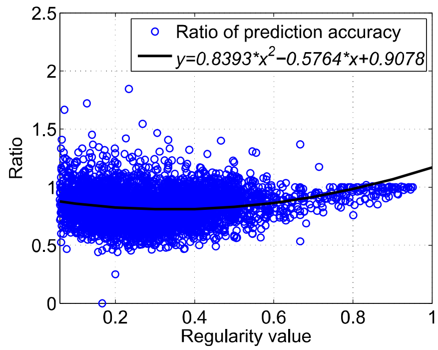


Fig. 7: The distribution of prediction accuracy's ratio between Markov and the Improved Markov.

V. Conclusion

It has been suggested that a smart city uses information technologies to make more efficient use of physical infrastructure, to enhance quality, performance and interactivity of urban services for citizens and so on. In this paper, we use real data traffic collected from 2G/3G mobile network in northern China to do our experiments. It indicates that we can capture the basic characteristics of human movements. We used two kinds of classical domain-independent prediction algorithms: Markov algorithm and LZ family algorithms to predict the next location. By using users' real trajectories extracted from our data set, we evaluated the prediction performance of Markov algorithm and LZ family algorithms. We find that Markov algorithm achieve lower accuracy value than LeZi Update and Active LeZi.

In order to improve the prediction accuracy and overcome drawbacks of the two kinds of prediction algorithms, Improved Markov algorithm, which considers the temporal feature and pattern weighting, is applied. Comparing with original Markov algorithm, the average prediction accuracy of Improved Markov algorithm gets a nearly 6% increase without building the complex prediction tree. At last, we performed experiments to find the relationship between prediction accuracy and trajectory's regularity. Result indicates that, the more regular trajectory is, the higher the prediction accuracy is. What's more, we find that although our Improved Markov performs better than the original Markov method in general. But, when the trajectory's regularity value is very low, the gap between two methods' prediction accuracy is narrowing.

For the future work, considering that different prediction algorithms may adapt to different kinds of users, we will further reveal the relationship between prediction accuracy of algorithms and user mobility.

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China (61671078), the Fundamental Research Funds for the Central Universities (2015RC11) , Director Foundation Project (2015BKL-NSAC-ZJ-01), and 111 Project of China (B08004).

References

[1] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.

[2] F. Giannotti, M. Nanni, F. Pinelli, and D. Pedreschi, "Trajectory pattern mining," in *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2007, pp. 330–339.

[3] F. Simini, M. C. González, A. Maritan, and A.-L. Barabási, "A universal model for mobility and migration patterns," *Nature*, vol. 484, no. 7392, pp. 96–100, 2012.

[4] H. He, Y. Qiao, S. Gao, J. Yang, and 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.

[5] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez, "Next place prediction using mobility markov chains," in *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*. ACM, 2012, pp. 1–6.

[6] A. Rodriguez-Carrion, C. Garcia-Rubio, C. Campo, A. Cortés-Martín, E. Garcia-Lozano, and P. Noriega-Vivas, "Study of lz-based location prediction and its application to transportation recommender systems," *Sensors*, vol. 12, no. 6, pp. 7496–7517, 2012.

[7] Y. Qiao, J. Yang, H. He, Y. Cheng, and Z. Ma, "User location prediction with energy efficiency model in the long term-evolution network," *International Journal of Communication Systems*, vol. 40, no. 7, pp. 1267–1287(21), 2015.

[8] Z. Ma, P. K. Rana, J. Taghia, M. Flierl, and A. Leijon, "Bayesian estimation of dirichlet mixture model with variational inference," *Pattern Recognition*, vol. 47, no. 9, pp. 3143–3157, 2014.

[9] K. Gopalratnam and D. J. Cook, "Online sequential prediction via incremental parsing: The active lezi algorithm," *IEEE Intelligent Systems*, vol. 22, no. 1, pp. 52–58, 2007.

[10] J. Petzold, F. Bagci, W. Trumler, and T. Ungerer, "Next location prediction within a smart office building," *Cognitive Science Research Paper-University of Sussex CSRP*, vol. 577, p. 69, 2005.

[11] L. Song, D. Kotz, R. Jain, and X. He, "Evaluating next-cell predictors with extensive wi-fi mobility data," *IEEE Transactions on Mobile Computing*, vol. 5, no. 12, pp. 1633–1649, 2006.

[12] V. Belik, T. Geisel, and D. Brockmann, "Natural human mobility patterns and spatial spread of infectious diseases," *Physical Review X*, vol. 1, no. 1, p. 011001, 2011.

[13] C. M. Schneider, V. Belik, T. Couronné, Z. Smoreda, and M. C. González, "Unravelling daily human mobility motifs," *Journal of The Royal Society Interface*, vol. 10, no. 84, p. 20130246, 2013.

[14] J. Yang, X. Zhang, Y. Qiao, Z. Fadlullah, and N. Kato, "Global and individual mobility pattern discovery based on hotspots," in *2015 IEEE International Conference on Communications (ICC)*. IEEE, 2015, pp. 5577–5582.

[15] 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.

[16] X. Lu, E. Wetter, N. Bharti, A. J. Tatem, and L. Bengtsson, "Approaching the limit of predictability in human mobility," *Scientific reports*, vol. 3, no. 2923, pp. 1–9, Oct. 2013.

[17] H. Si, Y. Wang, J. Yuan, and 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.

[18] S. Gatmir-Motahari, H. Zang, and P. Reuther, "Time-clustering-based place prediction for wireless subscribers," *IEEE/ACM Transactions on Networking (TON)*, vol. 21, no. 5, pp. 1436–1446, 2013.

[19] S. Scellato, M. Musolesi, C. Mascolo, V. Latora, and A. T. Campbell, "Nextplace: a spatio-temporal prediction framework for pervasive systems," in *International Conference on Pervasive Computing*. Springer, 2011, pp. 152–169.

[20] M. Morzy, "Prediction of moving object location based on frequent trajectories," in *International Symposium on Computer and Information Sciences*. Springer, 2006, pp. 583–592.

[21] D. Brockmann, L. Hufnagel, and T. Geisel, "The scaling laws of human travel," *Nature*, vol. 439, no. 7075, pp. 462–465, 2006.

[22] Y. Qiao, Y. Cheng, J. Yang, and J. Liu, "A mobility analytical framework for big mobile data in densely populated area," *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2016.