2016 International Conference on Applied Mechanics, Mechanical and Materials Engineering (AMMME 2016)
ISBN: 978-1-60595-409-7

Temporal Markov Chain Location Prediction

Chang SU, De-zheng WU, Xian-zhong XIE* and Qi WU

College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China

*Corresponding author

Keywords: Location prediction, Markov chain.

Abstract. With the rapid development of social mobile networks, it is highly desirable to be able to accurately predict the next check-in location. Most methods of social mobile networks are mainly based on vast user check-in historic records to analyze the feature of users' behaviors. The Markov Model is a frequently used method in location prediction problem. However, due to the diversity and sparsity of the social mobile network check-in data, tradition Markov Model location prediction, cannot take full advantages of potential information existed in plentiful check-in data. Then, we propose a prediction algorithm, temporal Markov Model (TMM) based on time characteristic. This algorithm relies on Markov Model to make a primary prediction of user check-in position. Then, utilizing the temporal character modifies the forecast result. We have conducted extensive experiments in two real datasets, and the results demonstrate the superiority of TMM over traditional Markov Models.

Introduction

With the rapid development of the Internet, mobile social network has become a widely used platform for user communications and sharing information. Moreover, the popularization of intelligent terminal and the generalization of GPS accelerate the fusion of social network and location based service. LBSN allows user sign in real world, it also support user to upload their check-in data and share with their friends. LBSN effectively reduces the distance between the virtual world and the real world, that's why LBSN get the favors of many users. Using check-in data analyze hobby of users can provide convenience for user experiencing life and making friends, it also offers applications in intelligent transportation, urban computing, etc. Meanwhile, it has huge commercial value in terms of online advertising, mobile recommendation and other commercial applications. Investigations of the location data have more and more become the focus in the recent research, such as location naming, location recommendation etc.

Location prediction research aimed to mine user check-in behavior through user's historical record to predict the user's future position. At present, there are many classic location prediction methods, these models are mostly based on frequent patterns and association rules, and some research work has showed that frequency of human check-ins follows power-law distribution. A large number of check-in behaviors are concentrated on a certain locations [1, 2]. Some researchers use this characteristic. They use the frequency of users check-in at a location as an important indicator and propose MFC models to predict the next location that user will check-in. However, due to the diversity of mobile social network check-in information, relying solely on the frequency of check-in as the basis to predict is insufficient.

The study of human movement patterns shows [1] that people's actions are repetitive. People repeat visits several fixed locations at a relatively fixed time every day. In real life, due to work, family and other reasons people have relatively fixed daily schedule, often at the same time to the same location to perform daily activities.

In this paper, we firstly analyze the temporal characteristics of the users check-in behavior in LSBN, then proposed a location prediction algorithm based on temporal characteristic. The algorithm uses Markov model to predict the position of the modeling problem, and then use the user's check-in temporal characteristics to correct predicted results to get a better prediction.

Related Work

Some researchers had put forward some different solutions by mining users' mobility patterns. Most researchers on this field have figured out the extreme importance of historical check-ins in location prediction [3, 4, 5], so their proposed prediction algorithms have heavily depended on regular or repetitive mobility patterns. For example, in [3], Chang et al proposed a logistic regression model and they found that users always like to check-in at the place they often visited before. In other words, the check-in frequency of the historical check-ins made by the users makes a strong influence. In [6], Song et al proposed an Order-k Markov Model considers the short term effect of historical check-ins, which is reported as a state-of-the-art prediction algorithm for location prediction. In [4], Cho et al used the temporal pattern of users' check-in data in their periodic social based model to do the location prediction. They found that the social relationship make a small but significant effect. In [7], Gao et al. first introduce the HPY language model for modeling the user's historical check-in sequences of LBSNs for each user to capture power-law distribution and short term effect of check-in behavior. Then, they propose a social-historical model (SHM) that enables them to study the importance of social-historical ties in affecting user's check-in behavior. In [8], Qiang Liu et al firstly study some existing prediction models and found the limitations of these models. For example, Factorizing Personalized Markov Chain (FPMC) was constructed based on a strong independence assumption among different factors. Tensor Factorization (TF) would meet the cold start problem. In this paper, we focus on analyzing the influence of the temporal features on users' mobility patterns, and combined the temporal features with Markov Model to solve location prediction problem.

Temporal Markov Model

For a specific user in social mobile network, let L be a set of his check-in positions and C is a sequence of his check-in historic record, $C = \{c_1, c_2, ..., c_n\}$.

We take positions as states of Markov Process, then state space $S=L$. The location prediction problem of the social mobile network could be formally defined as follows:

Computing the probability of user check-in at position l $P_u(c_{n+1} = l \mid C)$ through known check-in sequence. We utilize the Markov Chain Model to build location prediction model.

Computing the Initial Probability Q of the System

First, we need compute initial probability of the state (position). In general, the probability is obtained by Maximum Likelihood Estimation. The initial state distribution is represented by $Q_u(l)$, counting the number of user u check-in at different positions. Markov initial probability distribution is as follow:

$$Q_u(l) = \frac{|c_i \in C : c_i = l|}{|C|} \tag{1}$$

Where $|c_i \in C : c_i = l|$ represents the number of user u has signed at position l. $|C|$ represent the total number user u has signed.

Computing Probability Transfer Matrix P

In prediction of location, state transfer is: user has signed at place i and the next place he/she signed is j. The probability of user transfer from position i to position j is $P_{i,j}$ as:

$$P_{i,j} = \frac{|c_i \in C : c_i = i \wedge c_{i+1} = j|}{|C|} \tag{2}$$

So, if the number of location is N, then the form of probability transfer matrix is as

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \tag{3}$$

According to Markov Prediction Model, compute probability of user sign at position *l* in the next time.

$$P_u^M(c_{n+1} = l \mid C) = Q_u(l) * P \tag{4}$$

Location Prediction Based on Temporal Feature

We have already analyzed in section 3.2, the behavior of user sign have regularity. Obviously, time feature can improve the accuracy rating in user check-in problem. Then, utilize time feature to optimize Markov Prediction Model and its result.

Take day as timestamp, computing the probability of user sign at position *j* is as:

$$P_u^T(l) = \frac{| c_i \in C : c_i.t = tod \wedge c_i = l |}{| C_u |} \tag{5}$$

Where *tod* represents a day of week $tod \in [1, 2, ..., 7]$.

On the basis of the periodic of day, calculate the probability of the user sign at position *l* in different week. Combine with result of the previous Markov Model. By means of weighted methods obtain the probability of the user sign at position *l*.

$$P_u^M(c_{n+1} = l \mid C) = \alpha P_u^M(l) + (1-\alpha) P_u^T(l) \tag{6}$$

Where α is the representation of correction factor and $\alpha \in [0,1]$.

Experimental Results

In this section, we conduct empirical experiments to demonstrate the effectiveness of our methods on next location prediction. First introduce the dataset, baseline methods and evaluation metrics of our experiments. Finally, analyze the influence of the correction factor to experimental result.

Experimental Settings

We evaluate our method in Brightkite dataset to verify the effectiveness:

Brightkite is a location-based dataset online social network dataset which is similar to Gowalla. It contains check-in information about more than two years since 2008. We choose a total of 4491143 check-in records and 58228 users.

The first 60% elements of check-in dataset are selected for training.

We compare TMM with several classic methods for location prediction:

- MFC: MFC is composed by Chang et al. [3] in 2011. It's a logistic regression model. The main idea of this model is regarded the frequency of user sign in the location as an important indicator of location prediction, namely the number of users sign in a certain position is large, the more possibility of his return to this position to sign again in the future, therefore the probability of target users at various points in the sign in to the site in the user's personal history in the sign in record of frequency.
- MM: MM represents standard Markov Model. MM is a degraded TMM, that is we do not consider temporal character.

Due to Markov Model and Location Prediction Algorithm work not well in spatial check-in data. For this reason, we delete those users who check-in number less than one hundred. On Brightkite, it's sorted by time. We employ accuracy as evaluation metrics. There are several different with previous accuracy. If positions are showed in TOP-K, we assume that's a successful prediction. We employ the Accuracy@k as the indicator of the algorithm, where Accuracy@k is the proportion of k locations occupies total location. Using Accuracy@k as standard, compare MMT with MFC based on Bayesian Model, standard Markov Model.

Analysis of Experimental Results

Respectively, we test our algorithm and baseline methods. The performance comparison on Brightkite evaluated by Accuracy@k is illustrated in Figure 1. Obviously, we can observe that, TMM outperforms baseline methods. With the increase of k, all methods prediction accuracy is improved. But, when k hit about 40, respectively, prediction accuracy of every method is nearly stable in two datasets. Moreover, the TMM improvements comparing with MM and MFC are 15.8% and 21.3% respectively.

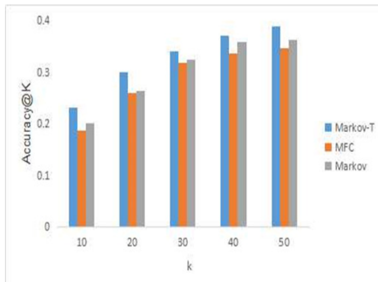

Figure 1. Performance comparison on Brightkite evaluated by accuracy.

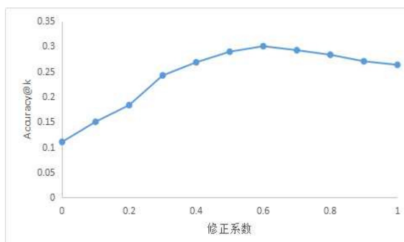

Figure 2. Performance of different correction factor on Brightkite.

On Brightkite dataset, Figure 2 show that when $\alpha = 0.6$ hit the highest prediction accuracy. This is because the user of Brightkite transfers more frequently. Especially, when $\alpha = 0$, the predicted results are completely affected by the temporal characteristics. When $\alpha = 1$, the TMM degrade to MM.

Conclusion

In this paper, we have proposed a novel Markov Model through analyzing regularity of user check-in behaviors in mobile social network, i.e. TMM.

In TMM, we consider the time character and capture regularity of user check-in behavior to modify the Markov Model. The experimental results show that TMM outperforms the classical Markov Model and MFC in two different datasets. Moreover, TMM is simple and effective to predict user check-in behavior.

Acknowledgements

We would like to show our deepest gratitude to Professor John E. Hopcroft, who has provided us with valuable guidance. This work was supported by the National Nature Science Foundation of China under contract number 61271259, the Chongqing Nature Science Foundation under contract number CSTC2016jcyjA0398, CTSC2011jjA40006, the Research Project of Chongqing Education Commission under contract number KJ120501.

References

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

[2] Lin, J., Xiang, G., Hong, J. I., & Sadeh, N. 2010. Modeling people's place naming preferences in location sharing. In Proceedings of the 12th ACM international conference on Ubiquitous computing (pp. 75-84). ACM.

[3] Chang, J., and Sun, E. 2011. Location 3: How users share and respond to location-based data on social networking sites. ICWSM.

[4] Cho, E.; Myers, S.; and Leskovec, J. 2011. Friendship and mobility: user movement in location-based social networks. In KDD, 1082–1090. ACM.

[5] Anagnostopoulos, A.; Kumar, R.; and Mahdian, M. 2008.Influence and correlation in social networks. In KDD, 7–15.ACM.

[6] Song, L; Kotz, D, Jain, R.; and He, X. 2004. Evaluating location predictors with extensive wi-fi mobility data. In INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies, volume 2, 1414–1424. IEEE.

[7] Gao H, Tang J, Liu H. 2012. gSCorr: modeling geo-social correlations for new check-ins on location-based social networks. In Proceedings of CIKM' 12, 1582-1586. ACM.

[8] Liu, Q., Wu, S., Wang, L., & Tan, T. 2016. Predicting the Next Location: A Recurrent Model with Spatial and Temporal Contexts.