Mobile Location Prediction in Spatio-Temporal Context

Huiji Gao
School of Computing, Informatics, and Decision System Engineering
Arizona State University
Huiji.Gao@asu.edu

Jiliang Tang
School of Computing, Informatics, and Decision System Engineering
Arizona State University
Jiliang.Tang@asu.edu

Huan Liu
School of Computing, Informatics, and Decision System Engineering
Arizona State University
Huan.Liu@asu.edu

ABSTRACT

The increasing use of mobile devices and popular mobile services has led to massive availability of mobile data. Location prediction is a specific topic in mobile data mining, with its potential application in traffic planning, location-base advertisement, and user oriented coupon dispersion. Traditional location prediction methods often separately consider spatial or temporal approach. Although there have been some attempts to integrate both spatial and temporal information for location prediction, most of them suffer from the overfitting problem due to the large number of spatio-temporal trajectory patterns. Therefore, smoothing techniques are indispensable for proper training of spatio-temporal models. In this paper, we propose a novel location prediction model to capture the spatio-temporal context of user visits. It considers not only the spatial historical trajectories, but also the temporal periodic patterns. By applying smoothing techniques on both patterns, our model obtains significant improvement compared to the state-of-the-art approaches.

Keywords

Location Prediction, Spatial-temporal Mining, Mobile data

1. INTRODUCTION

The widely used mobile devices and location-based services in the world have generated a large amount of mobile data. Generally, mobile data consists of the historical information of a user's visiting sequence, which includes the detailed context of the visited locations and corresponding time stamps. The research on user's visiting behavior with mobile data, typically location prediction, can potentially benefit many areas, such as mobile advertising [1, 2] and disaster relief [5, 7].

Traditional location prediction approaches on mobile data make use of spatial trajectory patten. Zheng et al. proposed a supervised learning approach to detect people's motion modes from their historical GPS data [13]. In [14], the authors modeled various individuals' location trajectories to mine the interesting locations and travel sequences with GPS logs. In [8], Gong et al. introduced social networks to predict a user's next location by the recent location of his closest friend, which did not consider the user's own location history. Gao et al. proposed a social-historical model to study the social-historical ties of check-in behavior for location prediction [6].

Researchers have also investigated the user's temporal periodic pattern on mobile data. Thanh and Phuong proposed a Gaussian mixture model based on user's cell-residence times to learn user movement profiles and predict location [12]. Cho et al. [4] introduced the social networks to model the user's check-in behavior, and proposed a Periodic and Social Mobility Model considering the user's movement as a 2-dimensional time-independent Gaussian distribution.

In this paper, we introduce a sophisticated location prediction model to capture the spatio-temporal trajectory of user visits, which considers not only the spatial historical trajectories, but also the temporal periodic patterns.

2. METHODOLOGY

Given a series of historical visits in a previous time section, and a context of the latest visit location with the time of the next visit, the location prediction problem can be described as finding the probability

$$p(v_i = l | t_i = t, v_{i-1} = l_k), \tag{1}$$

where $v_i = l$ indicates the i-th visit at location l, $t_i = t$ indicates the i-th visit happens at time t, and $v_{i-1} = l_k$ indicates the $(i-1)$-th visit happened at location l_k. Note that the variable t here is a periodic time indicating the time stamp of the visit, such as a specific hour (e.g., 23:00pm), a day of the week (e.g., Monday), a month (e.g., January) or even a year. The candidate location l with the highest probability would be the prediction of the i-th visit location. Using Bayes' rule, the probability in Eq. (1) is equivalent to:

$$\begin{aligned} & p(v_i = l | t_i = t, v_{i-1} = l_k) \\ = & \frac{p(v_i = l, t_i = t | v_{i-1} = l_k)}{p(t_i = t)} \\ \propto & \, p(v_i = l, t_i = t | v_{i-1} = l_k) \\ = & \, p(t_i = t | v_i = l, v_{i-1} = l_k) p(v_i = l | v_{i-1} = l_k) \\ = & \, p(t_i = t | v_i = l) p(v_i = l | v_{i-1} = l_k), \end{aligned} \tag{2}$$

Note that we consider

$$p(t_i = t | v_i = l, v_{i-1} = l_k) = p(t_i = t | v_i = l). \tag{3}$$

under the assumption that the probability of the current visit time is only relevant to the current visit location, instead of considering the last visit location.

Eq. (2) describes the process of computing the probability of next visit at location l. $p(v_i = l|v_{i-1} = l_k)$ is a prior probability that represents the probability of next visit at location l given the last visit at l_k, without considering any temporal information. The higher the probability is, the more probable the next visit would happen at location l. $p(t_i = t|v_i = l)$ is a posterior probability indicating the probability of the i-th visit happening at time t, observing that the i-th visit location is l. The higher the probability is, the more probable the visit at location l would happen at time t. These two probabilities restrain each other and complement each other in the form of spatio-temporal context, the candidate location with the highest probability is the one that most reasonably happens at time t, after the latest visit at location l.

2.1 Spatial Prior

In Eq. (2), $p(v_i = l|v_{i-1} = l_k)$ is the probability of the current visit at location l observing that the last visit happened at location l_k. It corresponds to the probability of the next location given the current spatial context without considering any temporal information. Previous study has investigated the spatial trajectories of user movement. In [3], a logistic regression model was proposed and found that the strongest predictor is the visiting frequency of the historical visits made by the user. Song et al. reported that the Order-k Markov model considers the short-term effect of historical check-ins, which is considered as a state-of-the-art prediction algorithm for location prediction [10]. The author also suggest that a "fallback" Markov model, which is a combination of order-1 and order-2 Markov model, would most likely result in good prediction performance in mobile data. Recently, Gao et al. [6] explored the social-historical ties on location-based social networks, they observed that the text documents and location trajectories share a set of common features, and introduced the Hierarchical Pitman-Yor (HPY) language model [11] for location prediction. Their results show that by smoothing the trajectories pattern of various lengths, the HPY model gives better performance than the most frequent model and order-k Markov model. Therefore, in this paper, we adopt the HPY model to compute the $p(v_i = l|v_{i-1} = l_k)$, named as HPY spatial prior.

2.2 Temporal Constraints

To compute the probability $p(t_i = t|v_i = l)$ in Eq. (2), which indicates the probability of the next visit happening at time $t_i = t$ given the next visit happening at location l, we decompose the temporal information $t_i = t$ into two parts, $h_i = h$ and $d_i = d$. Where $h_i = h$ indicates the hourly information of the visit, h is the hour of the day when the visit happens, i.e., 10:am, 2:00pm. And the $d_i = d$ indicates the "daily" information of the visit, here we define the "daily" information as the day of the week, d is the day of the week when the visit happens, i.e., Monday, Sunday. Without loss of generality, we consider these two temporal patterns independent, therefore,

$$\begin{aligned} & p(t_i = t|v_i = l) \\ = & p(h_i = h, d_i = d|v_i = l) \\ = & p(h_i = h|v_i = l)p(d_i = d|v_i = l), \end{aligned} \tag{4}$$

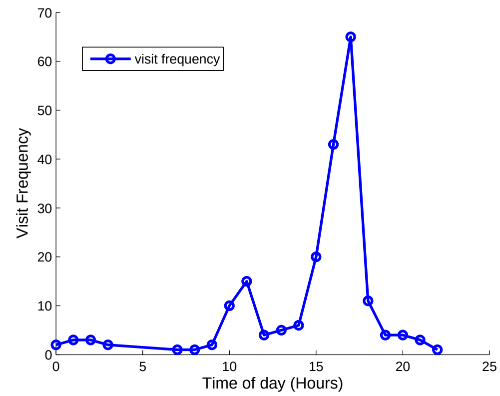

Figure 1: The visit frequency at various hour of day

where $p(h_i = h|v_i = l)$ indicates the probability of the visit happening at hour h given the observation that the visit actually happened at location l. Similarly, $p(d_i = d|v_i = l)$ indicates the probability of the visit happening at day of the week d given the observation that the visit actually happened at location l.

For the visit location l that has happened at a specific hour h (day d) in the training set, it is easy to compute the posterior probability $p(h_i = h|v_i = l)$ and $p(d_i = d|v_i = l)$. However, most of the locations in the training set do not have such information. For a specific user, the challenge of leveraging his temporal information is to estimate the probability of his visit at l happening at time h, while l has never been visited at time h by the user before. In figure 1, we plot the distribution of a user's visits at a specific location in 24 hours, the figure is generated from the training data set (user id: 013, place id: 3). The figure presents the Gaussian-like distribution, which corroborates the findings in [12] and [4]. We also find similar trends of the distribution over the day of the week. Therefore, to estimate the probability of a location at a time $p(t_i = t|v_i = l)$ without prior time knowledge, we assume the user's visits of a specific location follow Gaussian distribution over time, and leverage it to smooth the probability of visits at non-appear time. In the case of $p(h_i = h|v_i = l)$, for each place by a specific user, we assume its visiting frequency is a Gaussian distribution over 24 hours, i.e.,

$$p(h_i = h|l) = \mathcal{N}_l(h_i|\mu_h, \sigma_h^2), \tag{5}$$

where μ_h and σ_h^2 are the mean and variance of Gaussian distribution. Assume our data set is i.i.d, the joint probability of the training data set, given μ_h and σ_h, is in the form:

$$\begin{aligned} & p(\mathbf{h}|\mu_h, \sigma_h^2, l) \\ = & \prod_{i=1}^{N_l} \mathcal{N}_l(h_i|\mu_h, \sigma_h^2) \\ = & \prod_{i=1}^{N_l} \frac{1}{(2\pi\sigma_h^2)^{1/2}} \exp\{-\frac{1}{2\sigma_h^2}(h_i - \mu_h)^2\}, \end{aligned} \tag{6}$$

where N_l is the total number of visits at location l of this user. By maximizing the log likelihood above, we obtain the mean μ_h and σ_h^2, we use the unbiased estimation of σ_h^2 in

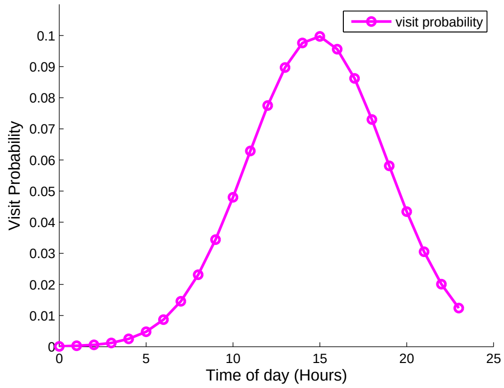

Figure 2: The smoothed visit probability at various hour of day

our experiment,

$$\sigma_h^2 = \frac{1}{N_l - 1}\sum_{i=1}^{N_l}(h_i - \mu_h)^2. \tag{7}$$

Figure 2 plots the estimated Gaussian distribution of user (user id: 013) visiting a location (place id: 3) based on figure 1. The plot shows that our estimation is able to capture the major trend of the user's periodic visiting pattern, therefore can be used to estimate the $p(h_i = h|v_i = l)$. Similarly, we assume the user's visit frequency is also a Gaussian distribution over the day of the week, with its mean μ_d and variance σ_d^2.

Therefore, Eq. (2) can be written as

$$\begin{aligned} & p(v_i = l|t_i = t, v_{i-1} = l_k) \\ &= p(v_i = l|v_{i-1} = l_k)p(h_i = h|v_i = l)p(d_i = d|v_i = l) \\ &= p(v_i = l|v_{i-1} = l_k)\mathcal{N}_l(h|\mu_h, \sigma_h^2)\mathcal{N}_l(d|\mu_d, \sigma_d^2). \end{aligned} \tag{8}$$

where $p(v_i = l|v_{i-1} = l_k)$ is the HPY prior. We define this model as "HPY Prior Hour-Day Model" model (HPHD), which will be used as our proposed method to predict the next location.

3. EXPERIMENTS

We use the mobile data set provided by Nokia Mobile Data Challenge which contains 80 users over one year of time [9]. Following the challenge instructions, we use setA as our training data, and setC as the testing data. SetA contains the location data of those 80 users except the 50 last days, while setC contains the 50 last days of location data. Since there are no ground truth to setC, we divide the training data setA into two parts (training and testing) to evaluate our approach. The testing part of setA is exactly the same as the toy data, which is provided by the challenge organizer that contains 3373 unknown location visits from the 80 users. For each user, we generate his training set from setA which contains all the visits before the visits in testing data. We ensure that the training set and testing data set generated from setA are strictly separated by time. We use prediction accuracy to evaluate our approach, which is the ratio of correctly predicted visits over the total number of predictions (3373).

3.1 Baseline Models

We employ 9 baseline models to evaluate our proposed method, with detailed descriptions below:

3.1.1 HPY Prior Model (HP)

The HP model simply considers the HPY prior only to predict the next location,

$$p^{HP}(v_i = l|t_i = t, v_{i-1} = l_k) = p(v_i = l|v_{i-1} = l_k), \tag{9}$$

3.1.2 HPY Prior Hourly Model (HPH)

The HPH model considers HPY prior and hourly information for location prediction,

$$\begin{aligned} & p^{HPH}(v_i = l|t_i = t, v_{i-1} = l_k) \\ &= p(v_i = l|v_{i-1} = l_k)\mathcal{N}_l(h|\mu_h, \sigma_h^2), \end{aligned} \tag{10}$$

3.1.3 HPY Prior Daily Model (HPD)

The HPD model considers HPY prior and daily information for location prediction,

$$\begin{aligned} & p^{HPD}(v_i = l|t_i = t, v_{i-1} = l_k) \\ &= p(v_i = l|v_{i-1} = l_k)\mathcal{N}_l(d|\mu_d, \sigma_d^2), \end{aligned} \tag{11}$$

3.1.4 Most Frequent Visit Model (MFV)

The MFV model assigns the probability of next visit v_i at location l as the probability of l appearing in the user's visiting history,

$$p^{MFV}(v_i = l|t_i = t, v_{i-1} = l_k) = \frac{|\{v_r|v_r\epsilon\mathcal{V}, v_r = l\}|}{|\{v_r|v_r\epsilon\mathcal{V}\}|}, \tag{12}$$

where $\mathcal{V} = \{v_1, v_2, ..., v_n\}$ is the set of check-in history.

3.1.5 Order-1 Markov Model (OMM)

The order-1 Markov model considers the latest visited place as context, and searches for frequent patterns to predict the next location. The probability of the next visit v_i at location l with order-1 Markov model is defined as:

$$\begin{aligned} & p^{OMM}(v_i = l|t_i = t, v_{i-1} = l_k) \\ &= \frac{|v_r|v_r\epsilon\mathcal{V}, v_r = l, v_{r-1} = l_k|}{|v_r|v_r\epsilon\mathcal{V}, v_{r-1} = l_k|}, \end{aligned} \tag{13}$$

note that the MFV is actually Order-0 Markov model.

3.1.6 Fallback Markov Model (FMM)

As suggested in [10], we introduce the "fallback" Markov model, which uses the results of the order-1 Markov model when it makes a prediction, or the MFV predictor if the order-1 Markov predictor has no prediction.

3.1.7 Most Frequent Hourly Model (MFH)

We choose the most frequent hour model (MFH) as another baseline considering the temporal patterns of the visits. Let $t_i = h$ denote that the time at the i-th check-in is h, where $h \in \mathcal{H} = \{1, 2, ..., 24\}$ is a discrete set of 24 hours. MFH model assigns the probability of next check-in v_i at location l at time h as the probability of the location l occurring at time h in the previous check-in history,

$$\begin{aligned} & p^{MFH}(v_i = l|h_i = h, v_{i-1} = l_k) \\ &= \frac{|v_r|v_r\epsilon\mathcal{V}, v_r = l, h_r = h|}{|v_r|v_r\epsilon\mathcal{V}, h_r = h|}, \end{aligned} \tag{14}$$

Table 1: Location Prediction Results

	Models	Correct No.	Accuracy
Spatial-based	MFV	1148	0.3402
	OMM	1466	0.4345
	FMM	1583	0.4692
	HP	1610	0.4772
Temporal-based	MFH	1462	0.4333
	MFD	1156	0.3426
	FHD	1538	0.4558
Spatio-temporal	HPH	1680	0.4979
	HPD	1583	0.4692
	HPHD	**1705**	**0.5053**

3.1.8 Most Frequent Daily Model (MFD)

Similar to the MFH model, we define the most frequent daily model as

$$\begin{aligned} & p^{MFD}(v_i = l | d_i = d, v_{i-1} = l_k) \\ = & \frac{|v_r | v_r \epsilon \mathcal{V}, v_r = l, d_r = d|}{|v_r | v_r \epsilon \mathcal{V}, d_r = d|}, \end{aligned} \tag{15}$$

3.1.9 Most Frequent Hour-Day Model (MFHD)

We define the "fallback" hour-day model, as a combination of MFH and MFD.

$$\begin{aligned} & p^{MFHD}(v_i = l | h_i = h, v_{i-1} = l_k) \\ = & p^{MFH}(v_i = l | h_i = h, v_{i-1} = l_k) * \\ & p^{MFD}(v_i = l | d_i = d, v_{i-1} = l_k). \end{aligned} \tag{16}$$

3.2 Evaluation Results

We evaluate our proposed model ***HPHD*** by comparing it with the nine baseline models, the results are shown in Table 1. The models in spatio-temporal family (HPH, HPD, HPHD) that consider both spatial and temporal context obtain significant improvement over the spatial-based or temporal-based models. The best prediction in spatio-temporal context (HPHD) has 6% relative improvement over the best prediction with spatial-based model (FMM), and 12% relative improvement over the best prediction with temporal-based model (FHD). Specifically, the HPHD model has the best performance among all the models since it considers not only the spatial trajectories, but also the hourly and daily patterns of the user visits.

Furthermore, we notice that Figure 1 can be decomposed into two Gaussian distributions with respective means and variances. It can be explained that a user would like to go to the same place in two specific time periods during a day, such as the restaurant for lunch and dinner, the home and work place, and so on. Therefore we propose an alternative HPHD version, named as AHPHD, which detects two peaks and assigns the desired probability into the Gaussian distribution, to which it belongs to. This approach is able to achieve good performance in our experiment (around 1% relative improvement over HPHD), but currently is not stable enough and is very sensitive to the detected peak position. In the Nokia Mobile Data Challenge, we try this approach as one of our 5 run submissions.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a location prediction model based on spatio-temporal context of user visiting history, which models a user's visiting behavior in the context of spatial historical trajectories and the temporal periodic patterns. By applying smoothing techniques to the model training, our model obtains significant improvement compared to many state-of-the-art approaches. As a result, in the Mobile Data Challenge, we adopt FHD, HP, HPH, HPHD and AHPHD as our five run submissions. In the future, we will consider using social network information together with spatio-temporal patterns of visiting trajectories to achieve better performance of location prediction.

5. ACKNOWLEDGMENTS

This work is supported, in part, by ONR (N000141010091).

6. REFERENCES

[1] S. Barnes and E. Scornavacca. Mobile marketing: the role of permission and acceptance. *IJMC*, 2(2):128–139, 2004.

[2] P. Barwise and C. Strong. Permission-based mobile advertising. *Journal of interactive Marketing*, 16(1):14–24, 2002.

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

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

[5] H. Gao, G. Barbier, and R. Goolsby. Harnessing the crowdsourcing power of social media for disaster relief. *Intelligent Systems, IEEE*, 26(3):10–14, 2011.

[6] H. Gao, J. Tang, and H. Liu. Exploring social-historical ties on location-based social networks. In *ICWSM*, 2012.

[7] H. Gao, X. Wang, G. Barbier, and H. Liu. Promoting coordination for disaster relief–from crowdsourcing to coordination. *Social Computing, Behavioral-Cultural Modeling and Prediction*, pages 197–204, 2011.

[8] Y. Gong, Y. Li, D. Jin, L. Su, and L. Zeng. A location prediction scheme based on social correlation. In *VTC Spring*, pages 1–5. IEEE, 2011.

[9] J. K. Laurila, D. Gatica-Perez, I. Aad, J. Blom, O. Bornet, T.-M.-T. Do, O. Dousse, J. Eberle, and M. Miettinen. The mobile data challenge: Big data for mobile computing research. 2012.

[10] L. Song, D. Kotz, R. Jain, and X. He. Evaluating location predictors with extensive wi-fi mobility data. In *INFOCOM 2004*, volume 2, pages 1414–1424. IEEE, 2004.

[11] Y. Teh. A bayesian interpretation of interpolated kneser-ney. 2006.

[12] N. Thanh and T. Phuong. A gaussian mixture model for mobile location prediction. In *2007 IEEE International Conference on Research, Innovation and Vision for the Future*, pages 152–157. IEEE, 2007.

[13] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W. Ma. Understanding mobility based on gps data. In *UbiComp*, pages 312–321. ACM, 2008.

[14] Y. Zheng, L. Zhang, X. Xie, and W. Ma. Mining interesting locations and travel sequences from gps trajectories. In *WWW*, pages 791–800. ACM, 2009.