*Accepted to International Journal of Data Mining, Modelling and Management (Special Issue on Spatial Information Modelling, Management and Mining).*

# Using trajectories for collaborative filtering based POI recommendation

Haosheng Huang* and Georg Gartner

Institute of Geoinformation and Cartography, Vienna University of Technology, Karlsplatz 13, 1040 Vienna, Austria
E-mail: haosheng.huang@gmail.com
E-mail: georg.gartner@tuwien.ac.at
*Corresponding author

**Abstract:** Current mobile guides often suffer from the following problems: a long knowledge acquisition process of recommending relevant Points of Interest (POIs), the lack of social navigation support, and the challenge of making implicit user-generated content (e.g., trajectories) useful. Collaborative filtering (CF) is a promising solution for these problems.

This article employs CF to mine GPS trajectories for providing Amazon-like POI recommendations. Three CF methods are designed: simple_CF, freq_CF (considering visit frequencies of POIs), and freq_seq_CF (considering both user's preferences and spatio-temporal behaviour). With these, services like "after visiting …, people similar to you often went to …" can be provided.

The methods are evaluated with two GPS datasets. The results show that the CF methods can provide more accurate predictions than simple location-based methods. Also considering visit frequencies (popularity) of POIs and spatio-temporal motion behaviour (mainly the ways in which POIs are visited) in CF can improve the predictive performance.

**Keywords:** collaborative filtering, user similarity, spatio-temporal motion behaviour, trajectory, POI recommendation, user modelling, social navigation, mobile guide, Location Based Services, user-generated content.

## 1. Introduction

Technical advances in mobile devices and mobile communication have led to the introduction of Location Based Services (LBS). Mobile guide is the largest group of LBS applications (Raper et al., 2007). One of the key goals of mobile guides is to provide users with relevant information/services for satisfying their need, e.g., recommending some Points of Interest (POIs) to visit.

Currently, POI recommendation in mobile guides often relies on knowledge about POIs (domain model), knowledge about users (user model), and an adaptation engine. However, building these models and the adaptation engine usually has to undergo a long process of knowledge acquisition, which is very time-consuming and impractical for many LBS applications. Additionally, in daily life, we often employ social navigation strategy, i.e., using cues from "the behaviour [experiences/opinions] of other people" to manage our activities (e.g., choosing where to visit) (Höök, 2003). Social navigation can enable users to gain more efficient and more satisfying answers to their problems (Wexelblat, 1999). However, little work has been done on incorporating social navigation into LBS. What's more, recently, there is a trend towards incorporating Web2.0's "participation" notion into LBS. More and more user-generated content (UGC, especially implicit UGC, e.g., trajectories) is created. Harnessing UGC to provide smart services in LBS has become increasingly popular.

Collaborative filtering (CF, known as Amazon-like recommendation) is a promising solution for the above problems. It aggregates opinions (i.e., UGC) of similar users in similar contexts to help the current user efficiently identify interesting information (Resnick and Varian, 1997). Therefore, by incorporating CF into mobile guides, relevant information (POIs) matching a user's need can be identified (by aggregating opinions of similar users).

This article investigates methods of incorporating CF into mobile guides to provide personalised POI recommendations. Specifically, we aim at applying CF methods on the highly available GPS trajectories to enhance visitors with Amazon-like POI recommendations, i.e., "after visiting POI A, other people similar to you often went to POI B". Three CF methods are proposed and implemented: simple_CF, freq_CF (considering visit frequencies of POIs) and freq_seq_CF (considering both users' preferences and motion behaviour).

The rest of this article is structured as follows. In section 2, we outline related work. Section 3 incorporates CF into mobile guides, and develops three different CF methods. Section 4 experimentally evaluates the proposed methods with two GPS datasets. Finally, we draw conclusions and present future work in section 5.

## 2. Related work

The article concerns how mobile guides can be improved in the era of Web 2.0. It integrates several mainstream trends and concepts, such as recommendation in mobile guides, Web 2.0, CF and trajectories. We summarise the related work on these aspects.

### *2.1 POI recommendation in mobile guides*

For providing POI recommendations, current mobile guides often rely on knowledge about POIs (domain model, DM), knowledge (preferences and need) about users (user model, UM) and an adaptation engine. The engine measures the appropriateness of the objects (DM) for satisfying a particular user's need (UM), and returns relevant objects. Building these models and the adaptation engine usually has to undergo a long underlying learning (knowledge acquisition) process, which is often very time-consuming and impractical for many LBS applications. More importantly, current mobile guides are unable to effectively provide users with relevant services/information in unseen situations or situations with little previous knowledge, which are very common in real world applications.

In contrast with the above approaches, we adopt a CF approach. CF is the most popular recommendation technique (Adomavicius and Tuzhilin, 2005). It is well-known through its Amazon-like recommendation: "users who bought … [i.e., like you] also bought …". CF aggregates opinions of similar users to help individuals efficiently identify interesting information (Resnick and Varian, 1997). Opinions from other users may reflect their perceptions on the fitness/appropriateness of a particular item (information) for the context. If most of the similar users like/visit that particular item, it can be considered as a matching item for the current user. Therefore, CF can be viewed as a real-time learning process of building DM and UM, and an automatic engine for identifying relevant information. Moreover, as CF solely relies on user feedback (explicit or implicit) and requires no previous domain knowledge, mobile guides employing CF will be able to effectively provide POI recommendations in dynamic environments and unseen situations. To sum up, CF is a novel method for providing POI recommendations in mobile guides.

### *2.2 CF in LBS*

CF is often applied in web-based applications, such as movie recommendation, and product recommendation. There are studies applying CF in LBS, such as restaurant recommendation (Horozov, Narasimhan, and Vasudevan, 2006), event recommendation (de Spindler et al., 2006; Li et al., 2009), and POI recommendation for tourism (van Setten, Pokraev, and Koolwaaij, 2004). However, many of them rely on explicit ratings from users. Explicit ratings require users' active involvement, which will bring some burden to users, and interrupt normal patterns of users' action (Nichols, 1997).

This article aims at using the highly available GPS trajectories to enhance visitors with Amazon-like POI recommendations. With this, users do not need to do anything other than using the system. The system logs users' moving tracks to unobtrusively infer their preferences and makes POI recommendations. As trajectories and explicit ratings have very different characteristics, current CF methods designed for explicit ratings should be adapted and improved.

### *2.3 Recommendation in the era of Web 2.0*

We are in the era of Web 2.0. Web 2.0 websites allow users to do more than just retrieve information. Users can also actively contribute their own data (i.e., UGC) to the web. Currently, with the impetus of Web 2.0 applications, such as Facebook, Flickr, Twitter and Foursquare, huge amount of UGCs are being created every hour, even every second. These UGCs may be considered as users' personal opinion expression, and are very useful for making recommendation.

There are studies harnessing UGCs for recommendation. For example, Phelan et al. (2011) use a content-based technique and real-time twitter data to recommend news. Based on Twitter lists, Nasirifard and Hayes (2011) assist Twitter users to select which followers would best be able to propagate the message to a relevant community-oriented audience. Choudhury et al. (2010) extract moving tracks from photo postings on Flickr, and aggregate these tracks to recommend itineraries according to the user's time and destination constraints.

Unlike the above methods relying on users' postings, this article harnesses another kind of highly available user-generated data - GPS trajectories - to provide personalised POI recommendations in mobile guides.

### *2.4 Mining trajectories*

With the increasing ubiquity of GPS-enabled devices, more and more people start to record their travel/sports experience with GPS loggers, and then upload, visualise and browse their GPS data on a web map. Therefore, large spatio-temporal datasets (e.g., trajectories) are highly available. Recently, mining these kinds of user-generated GPS data is receiving considerable attention.

There are studies focusing on mining personal location history based on individual trajectories. They focus on detecting significant locations of users, predicting users' behaviour among these locations, identifying users' spatio-temporal behaviour patterns, and recognizing users' activities on each location (Li et al., 2008). In the meantime, many other studies mine multiple users' trajectories to understand mobility-related phenomena. For example, Gonotti et al. (2007) aggregate a set of individual trajectories to identify spatio-temporal behaviour patterns, Zheng et al. (2008) infer users' transportation modes (e.g., walking or driving) based on trajectories of different users, Zheng et al. (2009a) mine interesting locations and travel sequences from multiple users' trajectories. Li et al. (2008) propose an interesting user similarity measure based on different users' trajectories; however, the measure is not integrated into the CF process.

Recently, a significant number of articles have presented work aiming to mine GPS trajectories of car drivers (e.g., taxi drivers) for route recommendation for car navigation (Letchner, Krumm, and Horvitz, 2006; Yuan et al., 2010). There are also studies mining trajectories for city visitors. For example, Yoon et al. (2011) recommend itineraries to visitors based on user-supplied queries (start, destination, and duration) and GPS trajectories from other users. However, the aspect of personalization has not been comprehensively addressed. Unlike the above studies, we aim at mining GPS trajectories to enhance visitors with Amazon-like POI recommendations, such as "after visiting …, people similar to you often went to …". Personalisation is the focus.

A very similar work is provided by Bohnert et al. (2008), which mine visitors' moving tracks to provide exhibit recommendations in a museum. However, in this article, we aim to make recommendations in outdoor scenarios (e.g., urban environment or zoo). More importantly, we will investigate whether considering visit frequencies (popularity) of POIs and spatio-temporal motion behaviour in the CF process can improve the predictive performance.

## 3. CF-based POI recommendation

Among different CF methods, neighbourhood-based CF (user-based and item-based) gain a huge popularity because of its simplicity, justifiability (easy to explain the reason behind prediction), efficiency (less computation and memory cost, suitable to mobile application environment), and abilities to provide serendipitous recommendations (Desrosiers and Karypis, 2010). As a result, in this article, user-based CF is employed for mining trajectories to provide Amazon-like POI recommendations. It includes three key stages: building user profiles, computing of user similarities, and aggregating of ratings from the N most similar users for recommendation.

### *3.1 Building user profiles*

The first stage of a CF is to build user profiles from feedback on items made over time. Specifically, a set of POIs visited by each user should be extracted from his/her trajectory. The POI set can be viewed as his/her preference profiles.

In order to identify the POIs visited by each user, we employ the SMoT (Stops and Moves of Trajectories) method developed by Alvares et al. (2007). It requires a set of candidate stops as inputs. In this article, a candidate stop corresponds to a POI. Each candidate stop is a tuple $C_i = \{POI_i, R_i, D_i\}$, where $POI_i$ is a descriptive text, $R_i$ is a topologically closed polygon in $\mathbf{R}^2$, and $R_i$ is a strictly positive real number. In a semantic level, $POI_i$ is the name or ID of the POI, $R_i$ defines the boundary geometry of the POI, and $D_i$ is the time threshold which defines a stop. If a user has stayed within a polygon (e.g., $R_j$) with a duration exceeding the time threshold $D_j$, this user can be considered to have visited the corresponding $POI_j$.

Therefore, for each user, a set of POIs visited by him/her can be extracted from his/her trajectory. We simple represent the set of visited POIs with the name (ID) of each visited POI, $POIS_u$ = *{all* $POI_k$ *| the user u has stayed within* $R_k$ *with a duration exceeding the time threshold* $D_k$*}*. All POIs in $POIS_u$ are ordered according to the time when they were visited. As mentioned before, the POI set can be viewed as a user's preference profiles.

### *3.2 Measuring user similarity*

The key in a CF is to locate other users whose opinions can be used for generating recommendations for the current user. In this article, we identify these users in terms of their preference similarities with the current user, and similarity between two users is measure by comparing the POIs they visited. Three kinds of user similarity measures are proposed.

#### *3.2.1 A simple user similarity measure (simple_USim)*

The simple_USim measures the preference similarity between two users. It compares the POIs visited by the two users, and is calculated as follows:

$$simple\_USim(a,b) = \frac{|POIS_{a,b}|}{\sqrt{|POIS_a| * |POIS_b|}}$$

$POIS_a$ and $POIS_b$ are the set of POIs visited by user $a$ and user $b$, respectively. $POIS_{a,b}$ is the set of POIs that are visited both by user $a$ and user $b$. $|POIS_a|$ is the size (the number of elements) of $POIS_a$.

#### *3.2.2 A user similarity measure considering visit frequencies (popularity) of POIs (freq_USim)*

It is obvious that two users accessed a POI visited by a few people might be more correlated than others who share a POI history accessed by many people (Zheng et al., 2009b). For instance, many people have visited the Great Wall and the Forbidden City, two well-known landmarks in Beijing. It might not be the case that all these people are similar to each other. However, if two users visited a POI, which is not very popular, they might indeed share some similar preferences (Zheng et al., 2009b).

As a result, visit frequency of POI is considered when measuring preference similarity between two users. Following is the proposed *freq_USim*:

$$freq\_USim(a,b) = \frac{\sum_{p \in POIS_{a,b}} \frac{1}{F_p}}{\sqrt{(\sum_{p \in POIS_a} \frac{1}{F_p}) * (\sum_{p \in POIS_b} \frac{1}{F_p})}}$$

$POIS_a$ and $POIS_b$ are the set of POIs visited by user $a$ and user $b$ respectively. $POIS_{a,b}$ is the set of POIs which are visited both by user $a$ and user $b$. $F_p$ is the visit frequency of POI $p$ considering all the trajectories, and is measured as the ratio of the number of users visiting $p$ and the number of all users.

In a more general sense, considering visit frequencies of POIs into similarity measurement can be regarded as a specific technique of variance weighting described by Herlocker et al. (1999), which give different weights to different items to reflect their importance in measuring user similarity.

#### *3.2.3 A user similarity measure considering preferences and spatio-temporal behaviour (freq_seq_USim)*

Research in behaviour modelling has shown that users' motion behaviour also affects their activities, such as choosing which place to visit and which route to follow (Millonig and Gartner, 2008; Holden, 2008). In this article, we mainly focus on the sequence relationship between POIs visited (i.e., the ways in which they are visited). Suppose users *a, b* and *c* have the same preference ratings for $POI_A$, $POI_B$ and $POI_C$, users *a* and *b* visit them in the order of $POI_A$ -> $POI_B$ -> $POI_C$, and user c visits them in the order of $POI_B$ -> $POI_A$ -> $POI_C$. It is obvious that user *a* is more similar to *b* than to *c*. Therefore, user similarity based on sequence relationship is explored and considered during selecting similar users for making recommendations.

In literature, different methods have been proposed for measuring trajectory similarity in terms of sequence relationship, such as the Longest Common Subsequence (LCS) approach in Yan and Zeng (2009), Edit Distance on Real Sequence approach in Chen, Özsu, and Oria (2005). In this article, the LCS approach will be used. It finds the longest subsequence (not necessary consecutive) common to all sequences in a set of sequences (Wikipedia, 2011). The longer of the LCS, the more similar between two users when only considering sequence relationship.

Therefore, the LCS-based similarity between two users is measured as:

$$lcs\_USim(a,b) = \frac{len}{\sqrt{N_a * N_b}} * \frac{len}{\sqrt{Gap_a * Gap_b}}$$

*len* is the number of POIs in the LCS, $N_a$ and $N_b$ are the number of POIs visited by user *a* and user *b*, and $Gap_*$ is the index difference between the LCS's first and last POI in each trajectory. The second part of the above measure gives a higher value when the LCS is consecutive in the trajectories.

The overall user similarity based on preferences and spatio-temporal motion behaviour (mainly the ways in which POIs are visited) is defined as:

$$freq\_seq\_USim = \lambda * freq\_USim + (1-\lambda) * lcs\_USim$$

The important weight $\lambda \in (0,1)$ can be estimated and learned from the collected data. For example, we can evaluate several thousand parameterizations (e.g., varying the important weight), and use the best-performing one as the optimised weight.

### *3.3 Making recommendation*

In the following, we integrate the above three similarity measures into the CF process, and design three CF methods for making POI recommendations. We assume that the current user *u* has visited a set of POIs, and currently he/she is at the POI *p*, and asking "which POI to visit next". The steps of each method are as follows.

#### *3.3.1 simple_CF: using simple_USim*

1) Identifying users whose next POI after visiting *p* (the current POI) has not been visited by the current user *u*.

2) For the results of step 1), identify the N most similar users. The *simple_USim* measure is employed.

3) For the N most similar users, aggregating every similar user's next POI after visiting *p* (considering the user similarity value).

4) Selecting the POI with the highest predicted value, and recommending it to the current user *u*.

#### *3.3.2 freq_CF: using freq_USim*

The steps are the same as steps in simple_CF, except that the user similarity in step 2) is measured with the *freq_USim*.

#### *3.3.3 freq_seq_CF: using freq_seq_USim*

The steps are the same as steps in simple_CF, except that the user similarity in step 2) is measured with the *freq_seq_USim*.

With the above CF methods, Amazon-like POI recommendations can be provided in mobile guides.

## 4. Evaluation and discussion

Offline experiments, which evaluate algorithms on historical data, are often employed for assessing recommendation systems in literature (Jannach et al., 2011). In this section, the proposed methods are evaluated in an offline experiment using two historical GPS datasets. Section 4.1 describes these two datasets. In section 4.2, we discuss the experiment setting. The results are presented in section 4.3. We discuss and summarise the results in section 4.4.

### *4.1 Datasets*

Two datasets are used for the experimental evaluation: zoo dataset, and urban dataset. Both of them are GPS trajectory dataset. The zoo dataset was collected in Vienna Zoo (Schönbrunner Tiergarten, Austria) in the first half of 2010. The urban dataset was shared by prof. dr. ir. S.C. van der Spek from Delft University of Technology, who tracked visitors in the city centre of Delft (Netherlands) in the second half of 2009. These two datasets reflect different scenarios and scales, which may help us evaluate the proposed methods in a more comprehensive way. For example, the urban scenario is typical for outdoor mobile recommendation applications, while the zoo scenario is similar to some indoor scenarios, such as museums.

For every trajectory, we extract the following information: POIs visited and their orders, duration of the trip, and length of the trip. To simplify the process of identifying the visited POIs from every trajectory, 36 POIs (candidate stops) are defined in the zoo by considering the layout of the zoo and GPS accuracies. The SMoT method developed by Alvares et al. (2007) is employed to extract the list of POIs visited by each user from his/her trajectories. Similarly, we extract a series of POIs from every trajectory in the urban dataset.

We only consider trajectories with at least 8 POIs for our experimental evaluation. In total, for the zoo dataset, we have 39 trajectories. The average number of POIs visited is 13.9 (ranging from 8 to 23), with a standard deviation of 4. For the urban dataset, we obtain 91 trajectories. The average number of POIs visited is 10.8 (ranging from 8 to 18), with a standard deviation of 2.2.

### *4.2 Experiment setting*

We use the two datasets to evaluate the predictive performance of the proposed CF methods: simple_CF, freq_CF, and freq_seq_CF. A location-based method is implemented as a benchmark. The location-based method identifies POIs that are close to the current location, and have not been visited by the current user. It randomly recommends one of these POIs to the current user. This location-based method is designed based on the observation that when no/little previous knowledge about a place is available, mobile guides often recommend users with POIs that are near the current location.

Due to the small size of our datasets, we use a leave-one-out validation: For example, we train the four prediction models on 38 of the 39 visitors for the zoo dataset, and test them on the remaining visitor (i.e., the active user).

Precision and recall are the most popular metrics for evaluating information retrieval systems. Herlocker et al. (2004) point out that recall is impractical to measure in a recommendation system. Therefore, in this article, we employ precision to evaluate the proposed methods. In the proposed methods, we only recommend the top one POI to the active user. Therefore, precision is either 1 or 0, depending on whether the recommended POI is actually viewed immediately by the active user or not. We average the precision values for each method to represent its predictive performance. In other words, the predictive performance of each method is measured as the ratio of the number of correct recommendations (i.e., the recommended POI is actually viewed immediately by the active user) and the number of recommendation processes (i.e., 39 for the zoo dataset, and 91 for the urban dataset).

In order to identify the optimised value for the important weight $\lambda$ in the freq_seq_CF method, we evaluate several thousand parameterisations (e.g., varying the important weight), and use the best-performing one for our final experiment.

In the experiment, we evaluate how the predictive performance of the proposed CF methods differs when predicting POIs at different places of a trip (i.e., the 1$^{st}$ last, the 2$^{nd}$ last, the 3$^{rd}$ last, the 4$^{th}$ last, and the 5$^{th}$ last). For making recommendation for an active user, we run the benchmarking method (location-based method, LBM) 400 times, and use the average as the predictive performance for the current prediction to the current active user. The evaluation can help us answer the following questions:

1) Does the proposed CF methods perform better in recommending POIs than the LBM (LBM versus simple_CF, freq_CF and freq_seq_CF)?

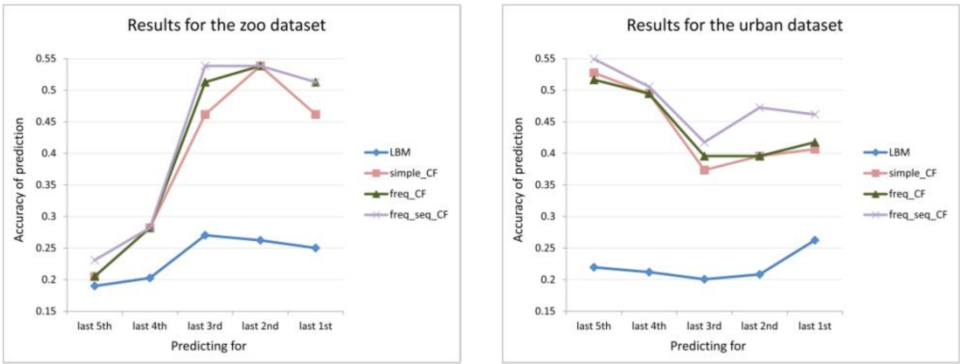
2) Does considering visit frequencies (popularity) of POIs and spatio-temporal behaviour in the CF process improve the predictive performance (simple_CF versus freq_CF and freq_seq_CF)?

3) How does the predictive performance of the proposed methods change when predicting POIs at different places of a trip?

### *4.3 Results*

The results of the experiment are summarised in Figure 1. In the following, we mainly focus on analysing the results with respect to the questions in section 4.2.

**Figure 1** The predictive performance of the proposed CF methods changes when predicting POIs at different places of a trip: zoo scenario (left), and urban scenario (right). A location-based method is implemented as a benchmark.




***CF methods versus LBM:*** For both datasets, when predicting POIs at different places of a trip, the proposed CF methods (simple_CF, freq_CF, and freq_seq_CF) always perform considerably better than the LBM. Specifically, for both datasets, the differences between the proposed CF methods and LBM are statistically significant[1] (*zoo dataset*: p=0.06 for LBM versus simple_CF, p=0.069 for LBM versus freq_CF, p=0.056 for LBM versus freq_seq_CF; *urban dataset*: p=0.001 for LBM versus simple_CF, p=0.0005 for LBM versus freq_CF, p=0.00004 for LBM versus freq_seq_CF). In short, the proposed CF methods can provide more appropriate recommendations (i.e., with more accurate POI predictions) than the LBM method.

***Considering visit frequencies of POIs and spatio-temporal motion behaviour:*** Figure 1 also shows that, among different CF methods, for both datasets, freq_seq_CF always performs the best, followed by freq_CF and simple_CF. Specifically, when considering the predictive performance, freq_seq_CF is 7.89% better than simple_CF for the zoo dataset, and 9.5% for the urban dataset. For both datasets, the overall performance of freq_CF is at least as good as the performance of simple_CF. Therefore, considering visit frequencies of POIs and spatio-temporal motion behaviour in the CF process can improve the predictive performance.

***Predicting POIs at different places of a trip:*** It is also important to note, for both datasets, the predictive performance of different CF methods at different places of a trip is correlated with that of the LBM. The performance of LBM reflects the POI layout (i.e., the connectedness of POIs), and how constrain the space is (the number of neighbours). In other words, the performance of LBM improves when the space is highly constrained, as there are few POIs available for recommendation. As a result, the divergent predictive performance of different CF methods at different places of a trip might be explained by the influence of POI layout. This confirms Bohnert et al. (2008)'s findings for exhibit recommendation in museums: the predictive performance of recommendation methods is influenced by the structure of a place.

## *4.4 Discussion and summary*

Simple location-based methods are often employed when no/little previous knowledge about a place is available. Compared to location-based methods that only consider neighbouring relationships between POIs, the proposed CF methods utilise other people's experiences/opinions for assisting current user's decision-making (i.e., choosing which POI to visit next). Opinions from other users may reflect their perceptions on the fitness/appropriateness of a particular item for the context. If most of the similar users like/visit that particular item, it can be considered as a matching item for the current user. Moreover, CF solely relies on user feedback and requires no previous domain knowledge. Therefore, the CF methods can provide more accurate POI recommendations than simple location-based methods, especially in dynamic environments and situations without previous domain knowledge, which is very common in real world applications. The results in Figure 1 confirm the above expectation.

We also expected that visit frequencies of POIs and spatio-temporal motion behaviour of the user (i.e., the way in which POIs are visited) have some contributions in defining the similarity between users. Considering visit frequencies of POIs and spatio-temporal motion behaviour in a CF process can help to locate similar users more accurately, and therefore improve the accuracy of prediction. This is also confirmed by the results in Figure 1.

[1] In this article, the statistical tests performed are independent group two-tailed t-tests. Due to the small size of the zoo dataset, $p < 0.1$ is used to denote statistical significance for the zoo dataset. For the urban dataset, we use $p < 0.05$ to indicate statistical significance.

In summary, the proposed CF methods (simple_CF, freq_CF and freq_seq_CF) can provide more suitable POI recommendations than simple location-based method in mobile guides. Also considering visit frequencies of POIs and spatio-temporal motion behaviour into the CF process can improve the predictive performance.

## 5. Conclusions and future work

In this article, methods of introducing collaborative filtering (CF) into LBS are proposed. To be more specific, CF methods are applied on the highly available GPS trajectories to enhance visitors with Amazon-like POI recommendations in mobile guides. Three different CF methods are proposed: simple_CF, freq_CF and freq_seq_CF. These methods are also implemented and evaluated with two real-world GPS datasets: a zoo dataset (Austria) and a urban dataset (Netherlands).

The experiment shows that compared to the simple location-based method, the proposed CF methods, which use other people's experiences/opinions (e.g., trajectories), can provide more appropriate POI recommendations in mobile guides. We also show that considering visit frequencies of POIs and spatio-temporal motion behaviour (mainly the ways in which POIs are visited) in the CF process can improve the predictive performance.

Our next step is to collect more trajectory data in both outdoor and indoor to evaluate the proposed methods. We propose that with more trajectories available, the predictive performance of the CF methods will be significantly improved. We are also interested in identifying more important features reflecting motion behaviour from other research fields, and considering them in the CF process. As context-awareness plays a key role in LBS, context-aware CF methods will be also explored to provide more appropriate POI recommendations.

## Acknowledgement

Haosheng Huang is a recipient of the Marshall Plan Scholarship of the Austrian Marshall Plan Foundation. He thanks Prof. Martin Raubal from ETH Zürich for his supervision and discussion, and prof. dr. ir. S.C. van der Spek from Department of Urbanism in Delft University of Technology for sharing the urban dataset ("tracking Delft I: walking patterns in the city centre", http://bk.tudelft.nl/uot/TD09). We also gratefully acknowledge the contributions of students from the LBS class (VU 126.099, 2010SS) in collecting the data in Vienna Zoo. We also thank the editors and two anonymous reviewers for their constructive comments.

## Reference

Adomavicius, G., and Tuzhilin, A. (2005) 'Towards the next generation of recommender systems: A survey of the state-of-the-art and possible extensions', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 17 No. 6, pp.734-749.

Alvares, L., Bogorny, V., Kuijpers, B., de Macedo, J., Moelans, B. and Vaisman, A. (2007) 'A Model for Enriching Trajectories with Semantic Geographical Information' in *Proceedings of the 15th annual ACM international symposium on Advances in geographic information systems*, ACM Press.

Bohnert, F., Zukerman, I., Berkovsky, S., Baldwin, T., and Sonenberg, L. (2008) 'Using Interest and Transition Models to Predict Visitor Locations in Museums', *AI Communications*, Vol. 21 No. 2-3, Special Issue on Recommender Systems, pp. 195-202.

Chen, L., Özsu, M., and Oria, V. (2005) 'Robust and Fast Similarity Search for Moving Object Trajectories' in *Proceedings of the 2005 ACM SIGMOD International conference on Management of data*, ACM Press, pp. 491-502.

Choudhury, M., Feldman, M., Amer-Yahia, S., Golbandi, N., Lempel, R. and Yu, C. (2010) 'Automatic construction of travel itineraries using social breadcrumbs' in *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, ACM Press.

Desrosiers, C., and Karypis, G. (2010) 'A comprehensive survey of neighborhood-based recommendation methods', in Ricci, F. et al (Eds.), *Recommender Systems Handbook*, Springer, pp. 107-144.

de Spindler, A., Norrie, M., Grossniklaus, M. and Signer, B. (2006) 'Spatio-temporal proximity as a basis for collaborative filtering in mobile environments' in *Proceedings of UMICS'06*, Luxemburg, June 5-9, 2006.

Gonotti, F., Nanni, M., Pedreschi, D. and Pinelli, F. (2007) 'Trajectory pattern mining' in *Proceedings of KDD'07*, ACM Press, pp. 330-339.

Herlocker, J. L., Konstan, J. A., Borchers, A. and Riedl, J. T. (1999) 'An Algorithmic Framework for Performing Collaborative Filtering' in *Proceedings of the 22nd Annual International ACM SIGIR Conference*, ACM Press, pp. 230-237.

Herlocker, J. L., Konstan, J. A., Terveen, L. G. and Riedl, J. T. (2004) 'Evaluating collaborative filtering recommender systems', *ACM Transactions on Information Systems*, Vol. 22 No.1, pp. 5-53.

Holden, A. (2008) *Environment and tourism*, 2nd ed., Routledge, London.

Horozov, T., Narasimhan, N., and Vasudevan, V. (2006) 'Using location for personalized POI recommendations in mobile environments' in *Proceedings of the International Symposium on Applications on Internet 2006*, IEEE Computer Society, pp. 124-129.

Höök, K. (2003) 'Social navigation: from the web to the mobile', in Szwillus, G. and Ziegler, J. (Eds.), *Mensch & Computer 2003: Interaktion in Bewegung*, Vieweg +Teubner, pp. 17-20.

Jannach, D., Zanker, M., Felfernig, A. and Friedrich G. (2011). *Recommender systems – an introduction*, Cambridge University Press, New York.

Letchner, J., Krumm, J. and Horvitz, E. (2006) 'Trip Router with Individualized Preferences (TRIP): Incorporating personalization into route planning' in *Proceedings of the IAAI'06*, AAAI Press, pp. 1795–1800.

Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W, and Ma, W. (2008) 'Mining user similarity based on location history' in *Proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, ACM Press, pp. 298-307.

Li, L., Lee, F., Chen, Y., and Cheng, C. (2009) 'A multi-stage collaborative filtering approach for mobile recommendation' in *Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication*, ACM Press, pp. 88-97.

Millonig, A. and Gartner, G. (2008) 'Ways of walking – Developing a pedestrian typology for personalised mobile information systems', in Gartner, G. and Rehrl, K. (Eds.), *Location Based Services and TeleCartography II - from Sensor Fusion to Context Models*, Springer, pp. 79-94.

Nasirifard, P. and Hayes, C. (2011) 'Tadvise: A Twitter Assistant Based on Twitter Lists', in Datta, A. et al. (Eds.), *Social Informatics*, Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 153-160.

Nichols, D. (1997) 'Implicit rating and filtering' in *Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering*, Budapest, Hungary, pp. 31-36, ISBN: 2-912335-04-3.

Phelan, O., McCarthy, K., Bennett, M. and Smyth, B. (2011) 'Terms of a Feather: Content-Based News Recommendation and Discovery Using Twitter', in Clough, P. et al. (Eds.), *Advances in Information Retrieval*, Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 448-459.

Raper, J., Gartner, G., Karimi, H., and Rizos, C. (2007) 'Applications of location-based services: a selected review', *Journal of Location Based Services*, Vol. 1 No. 2, pp. 89-111.

Resnick, P., and Varian, H. (1997) 'Recommender systems', *Communications of the ACM*, Vol. 40 No. 3, pp. 50-58.

van Setten, M., Pokraev, S., and Koolwaaij J. (2004) 'Context-aware recommendations in the mobile visitor application COMPASS' in *Proceedings of AH 2004*, Springer, pp. 235-244.

Wexelblat, A. (1999) *Footprints: Interaction history for digital objects*. PhD thesis, MIT Program in Media Arts & Sciences.

Wikipedia (2011) *Longest common subsequence problem*. Available at: http://en.wikipedia.org/wiki/Longest_common_subsequence_problem (Accessed 10 October 2011).

Yan, P., and Zeng, D. (2009) 'Spatial movement pattern discovery with LCS based path similarity measure' in *Proceedings of the 15th Americas Conference on Information Systems (AMCIS)*, San Francisco, CA, August 6-9.

Yoon, H., Zheng, Y., Xie, X. and Woo, W. (2011) 'Social itinerary recommendation from user-generated digital trails', *Personal and Ubiquitous Computing*. doi:10.1007/s00779-011-0419-8.

Yuan, J., Zheng, Y., Zhang, C., Xie, W., Xie, X. and Huang, Y. (2010) 'T-drive: driving directions based on taxi trajectories' in *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM Press, pp. 99-108.

Zheng, Y., Liu, L., Wang, L., and Xie, X. (2008) 'Learning transportation mode from raw GPS data for geographic applications on the Web' in *Proceedings of WWW 2008*, ACM Press, pp. 247-256.

Zheng, Y., Zhang, L., Xie, X., and Ma. W. (2009a) 'Mining interesting locations and travel sequences from GPS trajectories' in *Proceedings of WWW 2009*, ACM Press, pp. 791-800.

Zheng, Y., Chen, Y., Xie, X., and Ma, W. (2009b) 'GeoLife2.0: A Location-Based Social Networking Service' in *Proceedings of the 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware*, IEEE Computer Society, pp. 357-358.