CO-075

INCORPORATING CONTEXT-AWARE COLLABORATIVE FILTERING INTO LOCATION BASED SERVICES

HUANG H., GARTNER G.
Vienna University of Technology, VIENNA, AUSTRIA

1. INTRODUCTION

Providing context-aware services in Location Based Services is still very challenging. Currently, context-awareness in mobile guides mostly relies on an adaptation engine to determine the appropriateness of POIs for satisfying user's needs and context. However, building the adaptation engine has to undergo a long process of knowledge acquisition, which is very time-consuming and impractical for lots of LBS.

Additionally, the increasing ubiquity of GPS-enabled devices has led to the collection of large spatio-temporal datasets, such as trajectories. These trajectories may reflect the perspective and experiences of other people who solve their spatial tasks (e.g., choosing which POI to visit next) in this situation. It is obvious in our daily life that experiences from past users (especially similar users) in similar context can help current users to efficiently solve their problems (Wexelblat 1999). Therefore, by aggregating the trajectories, LBS can provide users with smart services, such as providing social affordance for making decisions. However, little research has addressed these considerations.

Collaborative filtering (CF, "Amazon-like recommendation") is a promising solution for the above problems. The goal of this paper is to investigate methods of introducing context-aware CF (CaCF) into LBS to provide context-aware recommendations. Specifically, we aim at applying CaCF methods on the highly available spatio-temporal trajectories to enhance visitors with context-aware POI recommendations. With CaCF, smart services like "in similar context, other people similar to you often visited POI A" can be provided in LBS applications.

2. METHODOLOGY

CaCF aggregates what similar users chose in similar context in the past for recommendation. Several key issues have to be considered when providing CaCF in LBS: annotating user profiles with context, measuring similarities between contexts, and incorporating context information into the CF process. In addition, methods of measuring user similarity based on trajectories are proposed.

2.1 Identifying Relevant Context Parameters

Context-dependent user profiles are important for context-aware recommendation. For annotating user profiles with context, a main question has to be answered: which context parameters are relevant and thus needed to be modeled.

We adopt the interactional perspective on context (Dourish 2004). Something is context (parameter) only if user's decision-making (e.g., choosing which POIs to visit), interaction with the system, or the behavior of the system depends on it, otherwise it is just a feature of the world (Winograd 2001). Based on this understanding, a two-stage method to identify relevant context parameters is designed: a preliminary set of context parameters can be identified from literature or brainstorming. Then, data about visitors (i.e., trajectories) are collected and annotated with the preliminary set of context parameters. The final set of context parameters can be created by refining the preliminary set according to the collected data. The basic strategy of refining is to analyze how some key aspects (e.g., the number of visited POIs, the length of the visit, and the duration of visit) of users' trajectories differ with different values of each context parameter in the preliminary set.

2.2 Measuring User Similarity

For each user, a serial of POIs visited by him/her can be identified from his/her trajectory. Therefore, a simple user similarity measure is adopted. We measure similarity between two users by comparing POIs they visited. 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. 2009). As a result, the visited popularity of a POI is considered when measuring similarity between users.

2.3 Measuring Context Similarity

The similarity of the context (situation) in which the trajectory is made with the current context of the active user (who asks for recommendations) determines the usefulness of this trajectory in recommending POIs for him/her. Two methods are explored for measuring context similarity: Local-Global Approach (LGA) and Statistic-Based Approach (SBA).

2.3.1 Local_Global Approach (LGA)

The local-global principle is often employed in case-based reasoning to measure the similarity between complex case representations (i.e., context or situation in our case) consisting of attributes (i.e., context parameters in our case) with various different value types. According to this principle it is possible to decompose the entire similarity computation in a local part only considering local similarities between single attribute values, and a global part computing the global similarity for whole cases based on the local similarity assessments (Stahl 2003). As a result, for measuring similarity between two contexts (situations), the following steps are applied: 1) for each relevant context parameter, calculate its local similarity. Context hierarchies are employed for measuring local similarity. 2) The global similarity between the two contexts is calculated as the sum of the product of every local similarity and its importance weight. The important weights are learned from the collected data.

2.3.2 Statistic-Based Approach (SBA)

The second approach adopts a machine-learning technique. With the method proposed in section 2.1, relevant context parameters can be identified. By varying values for each parameter, all different kinds of situations can be identified. In the following, we propose an approach to measure the similarity between any two situations.

We assume that if visits in a situation (e.g., A) are similar to visits in another situation (e.g., B), then these two situations can be considered as similar. As a result, similarity between different contexts (situations) can be measured as some statistical metrics, e.g., the distance of visits in situation A and visits in situation B.

2.4 Making recommendations

Context information can be incorporated into CF by contextual pre-filtering, contextual post-filtering, and contextual modeling (Adomavicius and Tuzhilin 2010). In this paper, we mainly focus on contextual pre-filtering and contextual modelling. As a result, four kinds of CaCF methods are designed (the current user is finishing the current POI *p*, and asking “which POI to visit next”):

Method2_1: Using Local-Global Approach and contextual pre-filtering (LGA_CP_CaCF)

1) Identifying users whose next POI after visiting p (the current POI) hasn’t been visited by the current user.

2) Filtering users whose context similarities with the current user do not exceed a threshold. Context similarity is measured by the LGA method.

3) For the results of step 2, identify the N most similar users. The user similarity measure proposed in section 2.2 is employed.

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

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

Method2_2: Using Local-Global Approach and contextual modeling (LGA_CM_CaCF)

1) The same as step 1 in Method2_1 (LGA_CP_CaCF).

2) For the results of step 1, identify the N most useful users. The usefulness is measured by considering both context similarity and user similarity.

$\text{Utility}(a,b) = \lambda * SIM_{user}(a,b) + (1-\lambda) * SIM_{conx}(C_a, C_b)$

Where C_a and C_b are the contexts of user a and b. SIM_{user} (a,b) is calculated using the LGA method in section 2.3.1. SIM_{conx} is measured by the LGA method.

3) For the N most useful users, aggregating every useful user’s next POI after visiting p (considering usefulness value).

4) The same as step 5 in Method2_1 (LGA_CP_CaCF).

Method3_1: Using Statistic-Based Approach and contextual pre-filtering (SBA_CP_CaCF)

The steps are the same as steps in LGA_CP_CaCF, except that the context similarity in step 2 is measured by the SBA method.

Method3_2: Using Statistic-Based Approach and contextual modeling (SBA_CM_CaCF)

The steps are the same as steps in LGA_CM_CaCF, except that the context similarity in step 2 is measured by the SBA method.

With the above CaCF methods, context-aware recommendations can be provided in LBS applications.

3. EVALUATION AND RESULTS

In this section, we discuss our experimental evaluation. The data collection and processing are discussed in section 3.1. Section 3.2 employs the proposed method in section 2.1 to identify relevant context parameters for our CaCF methods. We describe the experiment setting in section 3.3. The evaluation and results are presented in section 3.4.

3.1 Data Collection and Analysis

Thanks to the cooperation with Vienna Zoo, we collected trajectories in the zoo. We encouraged visitors to carry GPS loggers with them while walking through the zoo. Before they start, we recorded some additional information (e.g., context information) about them, i.e., weather condition (sunny or rainy), age (>=45 or <45), time pressure (has two hours to stay in the zoo? Yes or No), annual ticket (Yes or No), first time in the Zoo (Yes or No), and companion with small children (Yes or No). In total, 41 valid trajectories of different kinds of visitors were collected.

For each trajectory, the following information is modeled and extracted:

<ID, visited POIs and their orders, number of visited POIs, length of visit, duration of visit, whether the user visited the mountain or not, age, first time in the Zoo, companion with small children, time pressure, annual ticket, weather >

3.2 Identifying Relevant Context Parameters

The recorded context information (i.e., < “age”, “first time in the Zoo”, “companion with small children”, “time pressure”, “annual ticket”, “weather”>) can be viewed as the preliminary set of context parameters. In the following, we apply the proposed method in section 2.1 to identify relevant context parameters from this preliminary set.

We mainly compare the follow key aspects of visits among different situations: number of visited POIs, length of visit, duration of visit, and frequency of visiting the mountain. In order to test whether the differences among different conditions for each context parameter are significant, we employ the F-test. The F-Test indicates that only the variances (e.g., number of visited POIs, length of visit, and duration of visit) of the different groups of “age” and “weather” are significantly different. As a result, only “age” and “weather” are considered as relevant context parameters, and taken as the final set of context parameters.

3.3 Experiment Setting

We use the dataset in section 3.1 to evaluate the predictive performance of the proposed CaCF methods: LGA_CP_CaCF, LGA_CM_CaCF, SBA_CP_CaCF, and SBA_CM_CaCF. In order to experimentally study whether including context information in a CF can improve the predictive performance, we also implement a non-contextual CF method (non_CaCF, i.e., LGA_CP_CaCF ignoring step 2).

Due to the small size of our dataset, we use the leave-one-out validation. Accuracy is used to evaluate the performance of the methods, and is defined as the ratio of the number of corrected recommedations (i.e., the predicited POI is actually viewed immediately by the active user) and the number of recommendation processes.

Two kinds of evaluations are performed. The first evaluation studies how the predictive performances of the proposed CaCF differ among different sets of context parameters. This evaluation is very useful for testing the effectiveness of the method proposed in section 2.1 (i.e., identifying relevant context parameters). The second evaluation focuses on how the predictive performances of the proposed CaCF methods differ when prediciting POIs at different places of a visit (i.e., the 1st last, the 2nd last, the 3rd last, the 4th last, and the 5th last). It can help us to answer the following questions: 1) Does including context information in a CF for LBS improve the CF predictive performance (context-aware CF vs. non-contextual CF)? 2) How do the predictive performances of the proposed methods change when predicting POIs at different places of a visit?

3.4 Results

Figure 1 shows the results of how the predictive performances of the proposed CaCF methods change among different sets of context parameters.

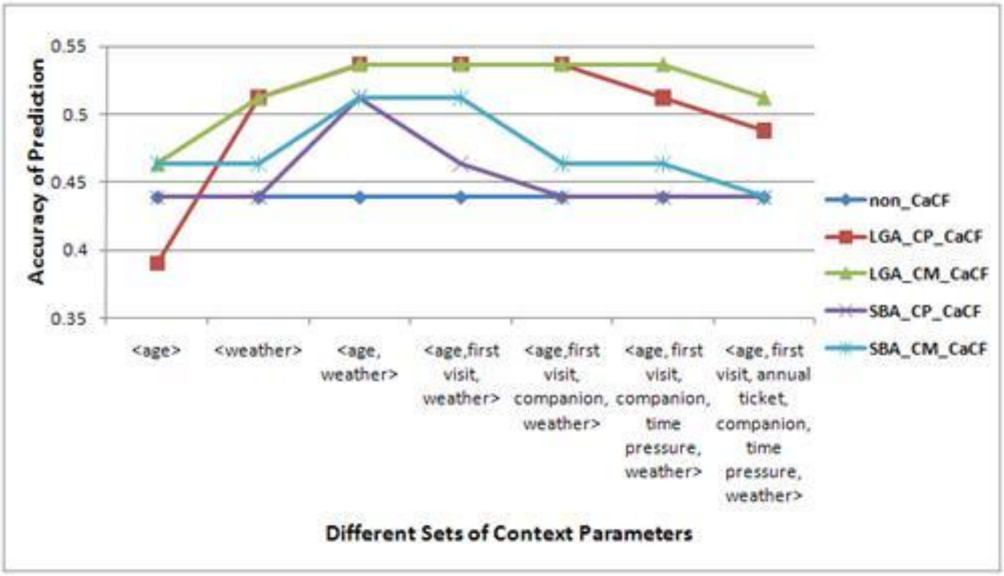

Figure 1. *the predictive performances of the proposed CaCF methods change among different sets of context parameters (when predicting for the last POI of every trajectory)*

Figure 1 shows that the accuracy of CaCF methods (using "<age, weather>" (i.e., the proposed set of context parameters in section 3.2)) is at least as high as the accuracy of CaCF methods using other sets of context parameters. Figure 1 also shows that for all different sets of context parameters(except LGA_CP_CaCF using"<age>"), the performances of CaCF methods are considerably better than the performances of non-contextual CF method (i.e., non_CaCF).

Figure 2 shows the results of how the predictive performances of the proposed CaCF methods change when predicting POIs at different places of a visit.

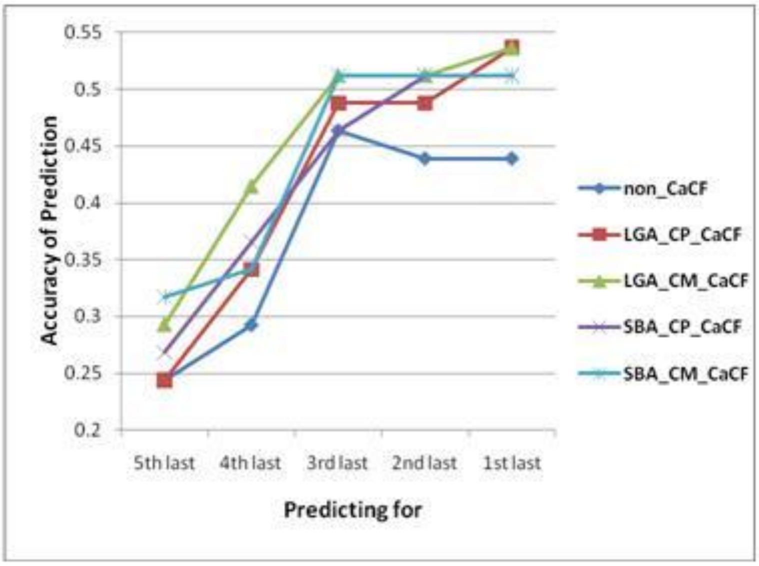

Figure 2. *the predictive performances of the proposed CaCF methods change when prediciting POIs at different places of a visit (using "<age, weather>" as context parameters)*

Figure 2 shows a upwards trend for the accuracy of all CaCF methods and non-contextual methods when the positions of the predicted POI increase. It also shows that when prediciting POIs at different places, the overall performance of contextual modelling approach (i.e., LGA_CM_CaCF, and SBA_CM_CaCF) is at least as good as the performance of contextual pre-filtering approach (i.e., LGA_CP_CaCF, and SBA_CP_CaCF).

In summary, when including context information in the CF process, choosing a suitable set of relevant context parameters is very important and may affect the predictive performance. The proposed method to identify relevant context parameters is feasible and useful, and using the proposed "<age, weather>" can achive a higher accuracy for all the designed CaCF methods. Most importantly, the proposed CaCF methods provide better predictive performance than non-contextual CF, that means including context information in a CF for LBS can improve the predictive performance.

4. CONCLUSIONS AND FUTURE WORK

In this paper, methods of introducing context-aware collaborative filtering (CaCF) into Location Based Services (LBS) are proposed. To be more specific, CaCF methods are applied on the highly available spatio-temporal trajectories to enhance visitors with context-aware POI recommendations in LBS.

Four CaCF methods are designed for LBS applications: LGA_CP_CaCF, LGA_CM_CaCF, SBA_CP_CaCF, and SBA_CM_CaCF. The results of the experiment show that the proposed CaCF methods are feasible and useful for providing context-aware recommendation in LBS applications. From the experiment, following conclusion can be drawn: including context information in the CF process can improve the predictive performances.

Our next step is to evaluate the proposed methods with different kinds of trajectories (from both indoor and outdoor). We are also interested in exploring a more complex user simiarlty measure in considering spatio-temporal behavior to provide more accurate results.

ACKNOWLEDGEMENT

Haosheng Huang is a recipient of the Austrian Marshall Plan Scholarship. Special thanks to Prof. Georg Gartner at the Vienna University of Technology, and Prof. Martin Raubal at the University of California, Santa Barbara for their supervision. Huang also gratefully acknowledges the contributions of students from the LBS class for collecting the data in Vienna Zoo.

REFERENCES

Adomavicius, G. and Tuzhilin, A. (2010): Context-aware recommender systems. In: Ricci, F., Rokach, L., Shapira, B. and Kantor, P. (eds), Recommender Systems Handbook. Springer, 217-253.

Dourish, P. (2004): What we talk about when we talk about context. Personal and Ubiquitous Computing, 8(1), 19–30.

Stahl, A. (2003): Learning of knowledge-intensive similarity measures in case-based reasoning. PhD thesis, University of Kaiserslautern.

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

Winograd, T. (2001): Architectures for context. Human-Computer Interaction, 16(2/4), 401-419.

Zheng, Y., Chen, Y., Xie, X. and Ma, W. (2009): GeoLife2.0: A Location-Based Social Networking Service. In: Proc. of the 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, p.357-358, May 18-20, 2009