Interaction Design in a Mobile Food Recommender System

Mehdi Elahi
Politecnico di Milano, Italy
mehdi.elahi@polimi.it

Mouzhi Ge
Free University of Bozen-Bolzano, Italy
mouzhi.ge@unibz.it

Francesco Ricci
Free University of Bozen-Bolzano, Italy
fricci@unibz.it

Ignacio Fernández-Tobías
Universidad Autónoma de Madrid, Spain
ignacio.fernandezt@uam.es

Shlomo Berkovsky
CSIRO, Australia
shlomo.berkovsky@csiro.au

Massimo David
Free University of Bozen-Bolzano, Italy
david.massimo@stud-inf.unibz.it

ABSTRACT

One of the most important steps in building a recommender system is the interaction design process, which defines how the recommender system interacts with a user. It also shapes the experience the user gets, from the point she registers and provides her preferences to the system, to the point she receives recommendations generated by the system. A proper interaction design may improve user experience and hence may result in higher usability of the system, as well as, in higher satisfaction.

In this paper, we focus on the interaction design of a mobile food recommender system that, through a novel interaction process, elicits users' long-term and short-term preferences for recipes. User's long-term preferences are captured by asking the user to rate and tag familiar recipes, while for collecting the short-term preferences, the user is asked to select the ingredients she would like to include in the recipe to be prepared. Based on the combined exploitation of both types of preferences, a set of personalized recommendations is generated. We conducted a user study measuring the usability of the proposed interaction. The results of the study show that the majority of users rates the quality of the recommendations high and the system achieves usability scores above the standard benchmark.

1. INTRODUCTION

Recommender systems are decision support tools that proactively identify and suggest items, which are expected to be interesting for the users. Recommendations are based on the users' previous interactions with the system and the explicitly provided users' preferences [15]. One important and new application domain for recommender systems is food. This application has recently drawn much attention in the research community due to its potential to improve eating behaviour of users and positively influencing their lives [8, 18, 20, 4, 11]. There is a broad spectrum of available information about food, such as recipe data and cooking instructions. Thus, some applications and websites already provide support functions allowing users to browse recipes and related information. However, most applications only offer generic and non-personalized recipe catalogue browsing support, without tailoring it to the tastes and preferences of individual users.

User preference elicitation is a fundamental and necessary step to go beyond this generic support and generate personalized recipe recommendations. More importantly than in other application domains, such as movies or books, recipe recommendations should not only be based on user's long-term tastes, but also fit their ephemeral preferences, such as the available ingredients or current cooking constraints.

In this paper we address this problem by proposing a preference elicitation approach for food recommender systems that obtains user preferences through a novel and effective interaction design. First, it exploits an integrated *Active Learning* algorithm [5, 6] for selecting the recipes to rate and tag that are estimated to be the most useful for the recommender. The active learning algorithm scores a recipe according to its predicted its rating (using transformed matrix of user-recipe) and then selects the highest scoring recipes. This reveals the users' long-term preferences, i.e., what they usually like to eat or cook. Second, when requested to generate recommendations, the system acquires short-term preferences referring to ingredients the user wants to cook or to include in the meal. The acquired preferences are used by a Matrix Factorization (MF) rating prediction model designed to take into account both tags and ratings [11, 13, 7].

In a real user study, we evaluated the proposed preference elicitation interaction and observed that the users have scored the usability of the system between "good" and "excellent" and assessed the presented recommendations, which are generated on the basis of the elicited preferences, to be of high quality.

Thus, the main contributions of our paper are: (a) a novel *interaction design* that is used to elicit long-term (general) and short-term (session-based) user preferences; and (b) an effective *preference elicitation* method that exploits active learning in the food recommendation domain.

.

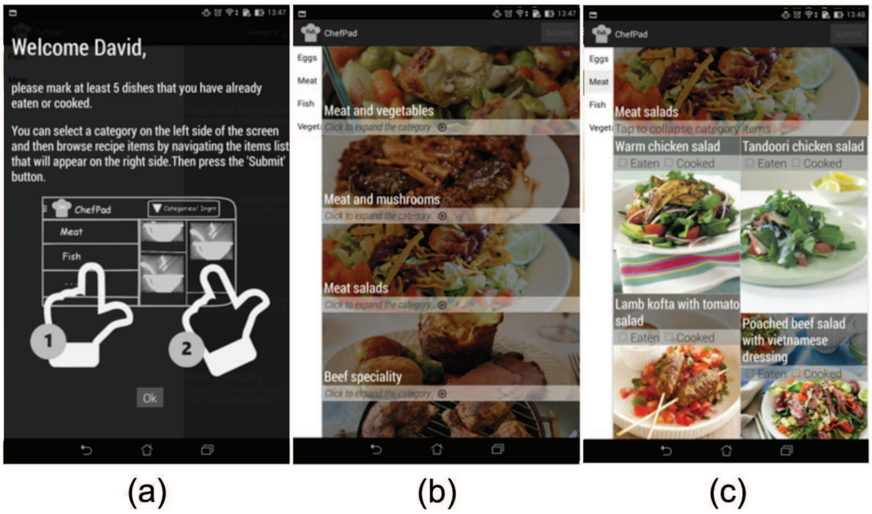

Figure 1: (a) user instructions, (b) browsing food categories, and (c) selecting eaten or cooked recipes.

2. RELATED WORK

Several recommender systems for the food domain have recently been developed [9, 18, 19, 20]. For example, Freyne and Berkovsky [9] proposed a food recommender that, through an easy-to-use interface, elicits user preferences and provides personalized recommendations Their system transferred the recipe ratings collected by the system to ingredient ratings and then aggregated the ratings of the ingredients used in a recipe to generate rating predictions.

Elahi et al. [4] proposed a food recommendation model that combines the predicted value of a recipe along different dimensions (user food preferences, nutritional indicators, and ingredients costs) to compute a single utility measure of a recipe. The goal is to consider factors influencing the user's food decisions in order to produce more useful and valuable recommendations. In a follow-up work [11], the authors conducted an offline evaluation of the rating prediction algorithm, which extends MF by using, in addition to ratings, the users' tags assigned to recipes. It was shown that this additional source of information about the user preferences allowed the proposed method to outperform other state-of-the-art algorithms, e.g., those proposed in [10].

In general, the user's preferences that are collected and used by a recommender can be either long-term (general preferences) or short-term (session-based and ephemeral). While obtaining both preference types is crucial, many recommender systems do not distinguish between the two. In fact, there are few studies that taken this consideration into account. Ricci and Nguyen proposed in [14] a mobile recommender system in travel domain, which elicits both general long-term preferences (e.g., explicitly defined by users) and short-term preferences in the form of critiques expressing more detailed session-based preferences. More recently, short term preferences were found to depend on the recommendation context and many context-aware approaches have been proposed to better suit the needs of the users [1].

It is worth noting that RSs research often focused on the improvement of the prediction model, by assuming that the preference elicitation process is completed. Hence, they ignore the complete user-system interaction, required for building a real-world recommender system. To address this limitation, this paper focuses on the interaction design, mainly for the preference elicitation: long-term and session-based.

3. USER-RECOMMENDER INTERACTION

We designed a complete human-computer interaction for collecting user preferences, in the form of recipe ratings and tags [4]. An Android-based prototype was developed, in order to implement this interaction. The first step is a general preference elicitation, aimed at collecting the long-term (stable) user preferences, i.e., what she generally likes to cook (or eat). This step includes two stages: (1) the system asks the user to specify the recipes she cooks at home and, (2) the user assigns ratings and tags to the recipes she experienced.

Upon logging in the system, the user can browse the full catalogue of recipes and mark those that she has eaten before (see Figure 1). Users can navigate through the recipe categories and sub-categories in order to find the desired recipe, e.g., 'Beef' → 'Roasted Beef' → 'Roasted Beef with Salad'. Inside each category there is a list of recipes mapped to this category. When the user finds one of them she can mark it as 'Eaten'or 'Cooked' by clicking the check box.

After that, a selection of the recipes that the user marked as eaten or cooked, is presented to the user for rating and tagging. This allows the system to acquire knowledge about the general user preferences. However, the system also needs to deeper explore the user's preferences and it presents additional recipes for the user to rate and tag. These are found by predicting what the user might have eaten, but did not mark in the first step. In order to find such recipes, we use active learning. For this, the rating dataset is transformed into a binary format indicating only whether the user rated an item: null entries are mapped to 0, and not null entries to 1. Then, using a factor model, predictions are computed for all the values mapped to 0, and for each user the items with the highest prediction are shown to the user [5, 6].

Figure 2-a shows the rating and tagging interface. This interface uses the classical 5-star Likert scale. The users are also requested to "explain" the core motivations for their ratings by assigning tags to recipes. Users can either tag a

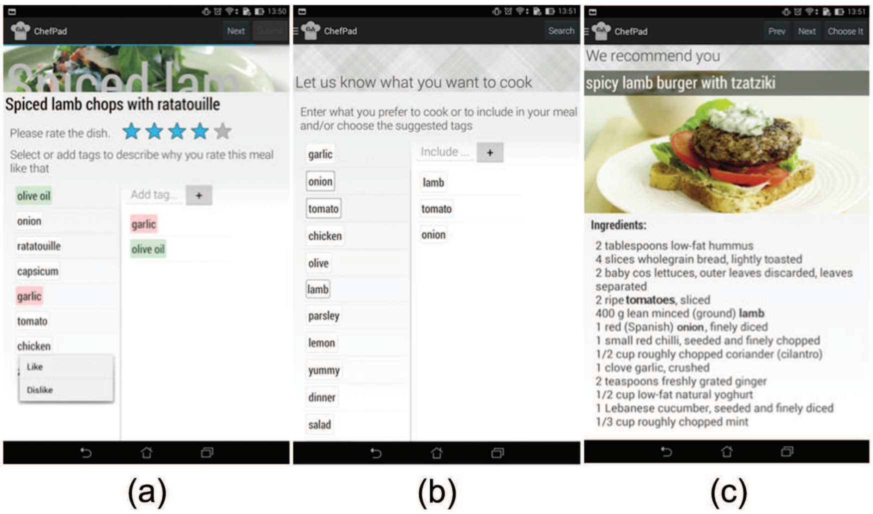

Figure 2: (a) general preference elicitation, (b) session-based preference elicitation, and (c) recommendation.

recipe with the suggested tags or add their own tags. At the recommendation time, session-specific preference are elicited (see Figure 2-b). The user enters the core ingredient she wants to include in the recipe. This is done by selecting a keyword from the list of suggestions derived from food ingredients and popular tags assigned by other users.

Then, the recommendations leverage both types of the collected user preferences, long-term and session-specific. The long term preferences are exploited by a custom MF rating prediction model [13], which uses the tagging information [7]. Each user is associated with a vector that models her latent features and each recipe is modeled by a vector that contains its latent features. Then, the rating of a user for an item is predicted by computing the inner product of the user and item vectors. To exploit the short-term model, the system post-filters the recommendations according to the current user preferences. The recipes with the highest rating are presented to the user one by one. When the user selects a recommended recipe, the system presents the required ingredients and detailed cooking instructions (see Figure 2-c).

4. USER STUDY

The main goal of the evaluation was to assess whether the system can effectively assist users in finding recipes that suit their preferences. For the user study, we designed a usage scenario a task that was formulated as follows: “You want to avoid everyday routine meals. You can use this application to discover new recipes that suit your taste”.

The users were asked to use the mobile application and complete a questionnaire referring to two performance indicators: perceived quality of recommendations quality and system usability. The first part of the questionnaire measured the level of user satisfaction with the recommendations. We used a validated instrument based on a set of questions developed by Knijnenburg et al. [12]. The second part of the questionnaire aimed at collecting the users' impression of the usability of the system. Here, we exploited the System Usability Scale (SUS) questionnaire [17]. The overall usability scores range from 0 to 100 and the bench-

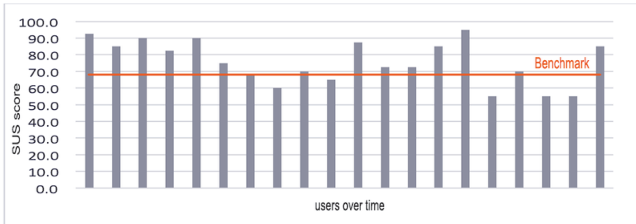

Figure 3: SUS results.

mark value is 68, which is the average SUS score computed over 500 usability studies [16].

In our experiment 20 subjects used the system and completed the questionnaire. They were either computer science researchers or non-academic people. 60% of subjects were male and 40% were female, the age range was 23 to 50, and the ethnical background varied across the subjects (Italy, France, USA, Germany, China, and more).

We first present the perceived recommendation quality results. The survey measures the recommendation quality using 7 questions on a Likert scale from 0 to 4, where 4 is the highest score. Thus, the maximum overall quality score is 28. The average perceived recommendation quality score across the 20 subjects was 19 and the median was 19 (see [3] for more details on the calculation). We observed that the maximal recommendation quality score was 26, and the minimal was 12. Thus, we can conclude that, on average, the users agreed that the recommendations were well-chosen and suited their preferences.

For the SUS usability score, we observed that for 75% of subjects the SUS score was higher than the 68 point benchmark (see Figure 3). The system achieved overall average SUS score of 75.50 and the median was 73.75, which is well above the benchmark. We observed that the minimal usability score of 55, and the maximal was 95. According to these results we can conclude that the system usability was considered between “good” and “excellent” [2].

We have computed the average replies for all the SUS statements and observing the statements with the highest average values, we can report that the users have evaluated the system easy to learn and easy to use. They also believe that various components were well-integrated into the system. On the other hand, by observing the statements with the lowest values we can state that users think that they have to learn a lot before they can use the system properly and they may need technical person for that. Our explanation for this result is that we need to improve further the interface and provide more explanations, so that users can better learn and understand the usage of the components in the system.

5. CONCLUSION AND FUTURE WORK

In this paper, we illustrated the preference elicitation process of a novel food recommender system [11]. Our system generates recommendations by exploiting tags and ratings in a MF algorithm. In our study, we collected user evaluations of the recommendation quality and system usability. Both measurements were found to be positive. This means that the proposed preference elicitation process and system interaction are liked by users.

Considering that this is a preliminary study, this paper has several limitations. First, the evaluation is performed on the whole system rather than on preference elicitation. Since the prediction model was already tested in another study [11], this work mostly focuses on preference elicitation as the main component of user interaction. Second, we have not compared our system with alternative preference elicitation processes. Our current result mostly reflects the users' direct perception of their interaction with the system. Third, we admit the limited number of subjects in the user study. In the future, we plan to increase the number of participants in the study. Also, we plan to extend the recommendation model by considering nutritional factors, e.g., the required calories and proteins, in order to build a health-aware recommender system.

6. REFERENCES

[1] G. Adomavicius and A. Tuzhilin. Context-aware recommender systems. In *Recommender systems handbook*, pages 217–253. Springer, 2011.

[2] A. Bangor, P. Kortum, and J. Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, 4(3), 2009.

[3] M. Braunhofer, M. Elahi, F. Ricci, and T. Schievenin. Context-aware points of interest suggestion with dynamic weather data management. In *Information and Communication Technologies in Tourism 2014*, pages 87–100. Springer International Publishing, 2014.

[4] M. Elahi, M. Ge, F. Ricci, D. Massimo, and S. Berkovsky. Interactive food recommendation for groups. In *Poster Proceedings of the 8th ACM Conference on Recommender Systems, RecSys 2014, Foster City, Silicon Valley, CA, USA, October 6-10, 2014*. 2014.

[5] M. Elahi, F. Ricci, and N. Rubens. Active learning strategies for rating elicitation in collaborative filtering: A system-wide perspective. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(1):13, 2013.

[6] M. Elahi, F. Ricci, and N. Rubens. Active learning in collaborative filtering recommender systems. In *E-Commerce and Web Technologies*, pages 113–124. Springer International Publishing, 2014.

[7] I. Fernández-Tobías and I. Cantador. Exploiting social tags in matrix factorization models for cross-domain collaborative filtering. In *Proceedings of the 1st Workshop on New Trends in Content-based Recommender Systems, Foster City, California, USA*, pages 34–41, 2014.

[8] J. Freyne and S. Berkovsky. Intelligent food planning: personalized recipe recommendation. In *IUI*, pages 321–324. ACM, 2010.

[9] J. Freyne and S. Berkovsky. Intelligent food planning: personalized recipe recommendation. In *IUI*, pages 321–324. ACM, 2010.

[10] J. Freyne and S. Berkovsky. Evaluating recommender systems for supportive technologies. In *User Modeling and Adaptation for Daily Routines*, pages 195–217. Springer, 2013.

[11] M. Ge, M. Elahi, I. Fernaández-Tobías, F. Ricci, and D. Massimo. Using tags and latent factors in a food recommender system. In *Proceedings of the 5th International Conference on Digital Health 2015*, pages 105–112. ACM, 2015.

[12] B. P. Knijnenburg, M. C. Willemsen, Z. Gantner, H. Soncu, and C. Newell. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22(4-5):441–504, 2012.

[13] Y. Koren and R. Bell. Advances in collaborative filtering. In F. Ricci, L. Rokach, B. Shapira, and P. Kantor, editors, *Recommender Systems Handbook*, pages 145–186. Springer Verlag, 2011.

[14] F. Ricci and Q. N. Nguyen. Acquiring and revising preferences in a critique-based mobile recommender system. *Intelligent Systems, IEEE*, 22(3):22–29, 2007.

[15] F. Ricci, L. Rokach, and B. Shapira. Introduction to recommender systems handbook. In F. Ricci, L. Rokach, B. Shapira, and P. Kantor, editors, *Recommender Systems Handbook*, pages 1–35. Springer Verlag, 2011.

[16] J. Sauro. Measuring usability with the system usability scale (sus). http://www.measuringusability.com/sus.php. Accessed: 2013-01-15.

[17] J. Swarbrooke and S. Horner. *Consumer behaviour in tourism*. Routledge, 2007.

[18] C.-Y. Teng, Y.-R. Lin, and L. A. Adamic. Recipe recommendation using ingredient networks. In *Proceedings of the 4th Annual ACM Web Science Conference*, pages 298–307. ACM, 2012.

[19] M. Trevisiol, L. Chiarandini, and R. Baeza-Yates. Buon appetito: recommending personalized menus. In *Proceedings of the 25th ACM conference on Hypertext and social media*, pages 327–329. ACM, 2014.

[20] R. West, R. W. White, and E. Horvitz. From cookies to cooks: Insights on dietary patterns via analysis of web usage logs. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1399–1410. International World Wide Web Conferences Steering Committee, 2013.