Paper:

An Analysis of Group Recommendation Strategies

Shlomo Berkovsky and Jill Freyne

Tasmanian ICT Center, CSIRO
GPO Box 1538, Hobart, Tasmania 7001, Australia
E-mail: firstname.lastname@csiro.au

[Received February 10, 2010; accepted April 10, 2010]

Collaborative filtering recommender systems often suffer from a data sparsity problem, where systems have insufficient data to generate accurate recommendations. To partially resolve this, the use of group aggregated data in the collaborative filtering recommendations process has been suggested. Although group recommendations are typically less accurate than personalized recommendations, they can be more accurate than generic ones, which are the natural fall back when personalized recommendations cannot be generated. This work presents a study that exploits a dataset of recipe ratings from families of users, in order to evaluate the accuracy of several group recommendation strategies and weighting models.

Keywords: recommender systems, collaborative filtering, group recommendations, evaluation

1. Introduction

The sheer volume of online information and the continued growth of online content necessitate Web sites to provide their users with personalized navigation support tools. These tools assist users in identifying items or content most relevant to them, while filtering out irrelevant items, in order to enhance user experience. Collaborative Filtering (CF) [1] is a widely-used statistical recommendation technique that predicts the interest level of a user in previously unseen items. The predictions are based on user models, storing either explicit (e.g., numeric or symbolic ratings) or implicit (e.g., purchasing behaviour or browsing logs) data.

CF is based on the assumption that users, who agreed in the past, are likely to agree in the future [2]. Hence, it analyses the opinions of users, who shared a target user's opinions in the past, in order to predict the target user's future opinions. To do this, CF initially computes the degree of similarity of the target user to all other users, then selects a neighborhood of the most similar users, and finally computes the prediction by aggregating the ratings of the neighbor users for the target item.

One of the main problems experienced by CF recommender systems is data sparsity, i.e., an insufficient amount of user information to generate accurate recommendations [3]. To overcome this, [4] proposed to aggregate sparse individual user data into group data and use the aggregated data for CF recommendations. That is, rather than generating recommendations based on the data available about an individual user, recommendations are based on the data about a group, to which the user belongs. This introduces an interesting trade-off. On one hand, the accuracy of group recommendations may not be as good as of those tailored to an individual user. On the other hand, group recommendations may be more sustainable, as they can be generated when sparsity prevents the generation of individual recommendations. Hence, we consider group recommendations as an alternative to generic recommendations, which are the natural fall back when personalized CF recommendations cannot be generated.

In this work we elaborate on this idea and analyze the performance of CF family recommendations, a particular case of group recommendations. We do this using real-life data logged during a study involving the users of an eHealth portal. The data includes explicit numeric ratings for a set of recipes, provided by families of users that interacted with the portal. We implemented several strategies and weighting models to aggregate individual user data into family data, generated CF recommendations using the aggregated data, and evaluated them against the logged ratings. The evaluation showed that (1) aggregating individual user models is superior to aggregating individual recommendations, (2) role based weighting is superior to a uniform weighting, and (3) the accuracy of recommendations is correlated to the size of the group.

The contributions of this work are four-fold. Firstly, we compare the performance of generic, individual, and group recommendation strategies. Secondly, we compare two strategies for group data aggregation. Thirdly, we compare four models for weighted data aggregation. Finally, we analyze the dependency between the size of groups and the accuracy of the generated recommendations.

The rest of this paper is organized as follows. In section 2 we overview related research on CF and group recommendations. In section 3 we present the group recommendation strategies and weighting models we developed. In section 4 we present and analyze the experimental evaluation. Finally, section 5 concludes this work and outlines future research directions.

2. Collaborative Filtering and Group Recommendations

CF is one of the most popular and widely-used recommendation algorithms. It is based on the notion of *word of mouth* [2], which assumes that users, who agreed in the past, will agree also in the future. In other words, it uses opinions of similar users to generate predictions for a target user. The opinions of users on the items are expressed either as explicit ratings given by users according to a predefined scale or as implicit ratings inferred from the logged user interactions with the system [6].

The main stages of the CF recommendation process are [1]: (1) recognizing commonalities between users by computing inter-user similarity; (2) selecting most similar users; (3) generating predictions for unseen items by aggregating opinions of the most similar users; and (4) recommending the items with the highest prediction scores. The main advantage of CF is that it is domain agnostic and independent of the representation of users and items, i.e., CF systems can generate recommendations for any items, e.g., movies, images, or text. Hence, CF is appropriate to a variety of domains and applications [7].

CF recommender systems often suffer from a data *sparsity* problem, which prevents the system from generating accurate predictions due to insufficient user data [3]. Two cases of the sparsity problem should be differentiated: the new user problem, when the number of ratings provided by a new user is insufficient for the identification of similar users and reliable recommendations generation to that user, and the new item problem, when the number of ratings from users for a new item is insufficient for the generation of reliable recommendations for that item to other users. In either case, the data sparsity problem results in low-quality recommendations, inevitably having a negative impact on user experience and system trust.

In contrast to *hybrid* recommendation approaches [8] aimed at overcoming the sparsity problem of CF, [4] proposed to use group based data as a means to enrich sparse individual user data. Group recommendations are mostly generated using two strategies: *aggregated models* – aggregate individual user data into group data and generate predictions based on the group data, or *aggregated predictions* – aggregate the predictions for individual users into group predictions. These strategies differ in the timing of the aggregation, as shown in **Fig. 1**. Specifically, the former [9, 10] aggregates the data of the group members *before* the CF prediction and then generates recommendations basing on the aggregated data. Alternatively, the latter [11, 12] treats group members as individuals for the prediction generation and *afterwards* aggregates the predictions to generate group recommendations. In both cases, various ways to aggregate the individual data into group data can be considered [13].

As discussed in [4], the selection between the two aggregation strategies depends on external factors, such as the ability to examine or negotiate group preferences, coverage of the recommendations, privacy considerations, and ability to explain the recommendations. However, in many circumstances several strategies could be applied. Hence, the primary aim of this work is to determine experimentally which data aggregation and group recommendation strategy is appropriate when dealing with groups that are made up of individuals within a nuclear family structure – a particular case of a strongly connected group.

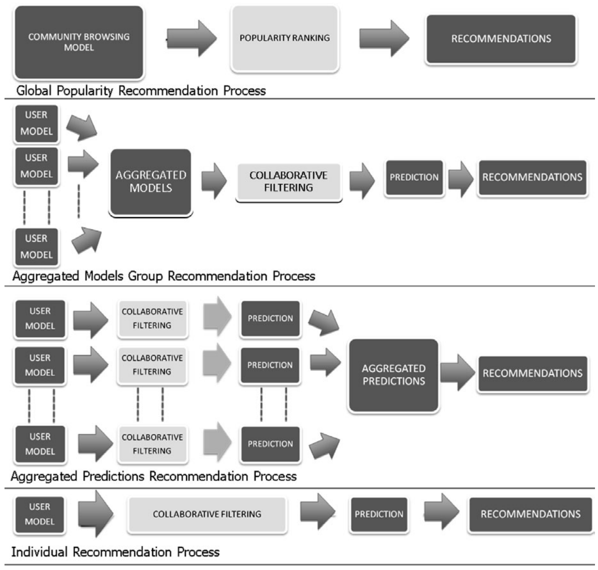

Fig. 1. Recommendation generation process.

3. Recommendation Strategies

We investigate the four recommendation strategies shown in **Fig. 1**. The *generic* strategy exploits the wisdom of the crowd and recommends the most popular items to all users. The *aggregated models* and *aggregated predictions* group strategies exploit the above two group recommendation algorithms. Finally, the *personalized* strategy exploits the standard CF algorithm.

The *generic* strategy implements simple mechanisms, which guide users to the most popular, i.e., highly rated items [14]. Each item $item_i$ is assigned a prediction score $pred(item_i)$ based on ratings $rat(u_x, item_i)$ of n_i users in $u_x \in U$, who rated $item_i$, as shown in Eq. (1).

$$pred(item_i) = \frac{\sum_{x \in U} rat(u_x, item_i)}{n_i} \quad . \; . \; . \; . \; . \quad (1)$$

The group based *aggregated models* strategy [15] initially constructs a family rating $rat(f_a, item_i)$ for family f_a and item $item_i$ by aggregating the individual ratings $rat(u_x, item_i)$ of family members $u_x \in f_a$, who rated $item_i$, according to their relative weight $\omega(u_x, f_a)$, as shown in Eq. (2).

$$rat(f_a, item_i) = \frac{\sum_{x \in f_a} \omega(u_x, f_a)\, rat(u_x, item_i)}{\sum_{x \in f_a} \omega(u_x, f_a)} \quad . \quad (2)$$

Then, CF is applied to the family models, as shown in Eq. (3). A prediction $pred(f_a, item_i)$ for family f_a

and $item_i$ is generated by computing similarity degree $sim(f_a, f_b)$ between the target family f_a and other families $f_b \in F$ and aggregating family ratings $rat(f_b, item_i)$ according to the similarity degree $sim(f_a, f_b)$.

$$pred(f_a, item_i) = \frac{\sum_{f_b \in F} sim(f_a, f_b)\, rat(f_b, item_i)}{\sum_{f_b \in F} sim(f_a, f_b)} \quad (3)$$

Finally, $pred(f_a, item_i)$ is assigned to the family members, i.e., $pred(u_x, item_i \mid u_x \in f_a) = pred(f_a, item_i)$.

The *aggregated predictions* strategy [15] initially generates individual prediction $pred(u_x, item_i)$ for user u_x and item $item_i$ using the standard CF algorithm, as shown in Eq. (4). The prediction is generated by computing the degree of similarity $sim(u_x, u_y)$ between the target user u_x and other users $u_y \in U$ and aggregating individual ratings $rat(u_y, item_i)$ of users, who rated $item_i$, according to the similarity degree $sim(u_x, u_y)$.

$$pred(u_x, item_i) = \frac{\sum_{y \in U} sim(u_x, u_y)\, rat(u_y, item_i)}{\sum_{y \in U} sim(u_x, u_y)} \quad . \quad (4)$$

Then, the process becomes group focused. To generate family prediction $pred(f_a, item_i)$ for family f_a and item $item_i$, individual predictions $pred(u_x, item_i)$ of family members $u_x \in f_a$ are aggregated according to their relative weight $\omega(u_x, f_a)$, as shown in Eq. (5).

$$pred(f_a, item_i) = \frac{\sum_{x \in f_a} \omega(u_x, f_a)\, pred(u_x, item_i)}{\sum_{x \in f_a} \omega(u_x, f_a)} \quad . \quad (5)$$

Finally, $pred(f_a, item_i)$ is assigned to the family members, i.e., $pred(u_x, item_i \mid u_x \in f_a) = pred(f_a, item_i)$.

The *personalized* strategy examines individual users regardless of their families. For each user u_x, each item $item_i$ is assigned a prediction score $pred(u_x, item_i)$ using the standard CF algorithm [1], as shown in Eq. (4).

Here we consider a personalization task of recommending the top k items, i.e., k items having the highest prediction scores that maximize $\prod_{i=1}^{k} pred(u_x, item_i)$. Note that the *generic* strategy generates one list of recommendations for all users, the group based *aggregated models* and *aggregated predictions* strategies – one list for each family, and the *personalized* strategy – one list for each user.

When the ratings and predictions of individual users are aggregated in Eqs. (2) and (5), each user is assigned a relative weight $\omega(u_x, f_a)$. We developed four models for weighting the contribution of individual users. The *uniform* model weights users uniformly, i.e., $\omega(u_x, f_a) = 0.25$ if the family contains 4 members. Three other models are role based. Here, a role refers to a user's function within a family: *applicant* – the adult, who initiated the family participation in the study, *partner* – the other adult in the family, or *child*. The *heuristic* model presumes that $\omega(u_x, f_a)$ is defined solely by the user's role. Hence, an applicant's weight is $\omega(u_x, f_a) = 0.5$, as they are likely to be highly interested in the portal, a partner's weight is $\omega(u_x, f_a) = 0.3$, as they are likely to be reasonably interested in the portal, and a child's weight is $\omega(u_x, f_a) = 0.1$, as they are not as likely to be interested in the portal.

Fig. 2. Recipe rating interface.

Two other weighting models are derived from the logged user interactions with the portal. The weights assigned to users refer to their activity $act(u_x)$, i.e., number of ratings $rat(u_x, item_i)$, as an indicator of their interest in the portal. The *role based* model weights users according to the relative activity of users having the same role across the entire community, as shown in Eq. (6).

$$\omega(u_x, f_a) = \frac{\sum_{y \in U} act(u_y) \mid role(u_y) = role(u_x)}{\sum_{y \in U} act(u_y)} \quad (6)$$

The *family log* model weights users according to their relative activity of users in their family only, as shown in Eq. (7).

$$\omega(u_x, f_a) = \frac{act(u_x)}{\sum_{y \in f_a} act(u_y)} \quad . \; . \; . \; . \; . \; . \; . \; . \quad (7)$$

4. Evaluation

The evaluation was carried out using a dataset of explicit ratings for recipes, provided by families of users that interacted with an experimental eHealth portal. The aim of the analysis was to determine which recommendation strategy is appropriate for a group recommender. Specifically, we aimed to compare the accuracy of two group recommendation strategies, four weighting models, and assess the implications of a group size on the accuracy of recommendations. Results obtained for a considerably smaller dataset were presented in [15].

4.1. Experimental Setting

The data was gathered during a three week study of family usage of the portal. A byproduct of the study was the capture of a dataset of explicit ratings for a set of recipes sourced from the CSIRO Total Wellbeing Diet book [5]. Users were asked to provide their opinions on how much recipes appealed to them (**Fig. 2** shows the rating interface). Their explicit symbolic ratings provided on a *hate* to *love* scale were converted into numeric ratings on a 1 to 5 scale.

Table 1 summarizes the dataset. The columns denote, respectively, overall number of users, overall number of families, number of families in which $n = 1, 2, 3$, or 4 users interacted with the portal,[1] number of recipes in the dataset, number of ratings captured, and data sparsity (ratio between the number of ratings captured and overall

1. Families having only 1 active user were excluded from the testing set and used only in the training set.

Table 1. Experimental dataset.

N_{users}	N_{fam}	$N_{fam,n=1}$	$N_{fam,n=2}$	$N_{fam,n=3}$	$N_{fam,n=4}$	N_{items}	$N_{rat(u,i)}$	$sparsity$
169	108	70	19	12	7	136	3305	14.38%

Table 2. Comparison of recommendation strategies.

metric	*weighting*	*generic*	*aggregated models*	*aggregated predictions*	*personalized CF*
F1	uniform	0.191	0.289	0.213	0.376
	heuristic	0.191	0.300	0.228	0.376
	role based	0.191	0.331	0.231	0.376
	family log	0.191	0.343	0.238	0.376
MAE	uniform	0.216	0.186	0.210	0.175
	heuristic	0.216	0.185	0.209	0.175
	role based	0.216	0.184	0.208	0.175
	family log	0.216	0.183	0.208	0.175
coverage	uniform	100%	97.63%	93.55%	85.41%
	heuristic	100%	97.65%	93.56%	85.41%
	role based	100%	97.63%	93.54%	85.41%
	family log	100%	97.62%	93.57%	85.41%

number of possible ratings). Notably, the distribution of recipe ratings was not uniform: 883 were rated *hate*, 1352 – *don't like*, 741 – *neutral*, 254 – *like*, and 75 – *love*.

For each user/family, a one-off similarity matrix with other users/families was computed using Cosine Similarity [1]. Using these matrices, 5 most similar users/families were selected, leave-one-out recipe rating predictions were computed, and 16 recommendation lists were produced: for the four recommendation strategies (*generic*, *aggregated models*, *aggregated predictions*, *personalized*) and four weighting models (*uniform*, *heuristic*, *role based*, *family log*). The recommendations were evaluated using *F1*, *precision@k*, Mean Absolute Error (*MAE*), and *coverage* metrics [16].

We denote using $\mathbb{V}$ the set of positively rated recipes, i.e., rated *neutral*, *like*, or *love*, and using $\mathbb{R}$ the set of recipes with positive prediction scores (3 or higher). Hence, *precision* of the recommendations is computed by $\frac{|\mathbb{V}\cap\mathbb{R}|}{|\mathbb{R}|}$ and *recall* by $\frac{|\mathbb{V}\cap\mathbb{R}|}{|\mathbb{V}|}$. When the size of $\mathbb{R}$ is k, the precision metric is referred to as *precision@k*. Combining the precision and recall metrics yields the *F1* metric, which represents their harmonic mean, as shown in Eq. (8).

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad . \ . \ . \ . \ . \ . \ . \quad (8)$$

We compute *MAE* as the average difference between the predicted and logged score for user u_x and item $item_i$, normalized by the size of the range of scores $R_d \in [1 \ldots 5]$ in the dataset, as shown in Eq. (9).

$$MAE = \sum_{x \in U} \sum_{i \in I} \frac{|\ pred(u_x, item_i) - rat(u_x, item_i)\ |}{|\ R_d\ |} \quad (9)$$

The *coverage* of the recommendations reflects relative portion of items, for which a prediction was generated. It is computed by dividing the number of successfully predicted items by the overall number of items in the dataset.

4.2. Recommendation Strategies and Weighting Models

The first question relates to the accuracy of the recommendations generated using the generic versus the group based versus the personalized strategies. **Table 2** shows the average classification accuracy *F1*, predictive accuracy *MAE*, and coverage scores obtained for each recommendation strategy and weighting model.

As expected, the results show that the *personalized* strategy performed best, outperforming both group based strategies, which in turn outperformed the *generic* strategy across all weighting models and both the accuracy metrics (higher *F1* and lower *MAE*). Analysing the coverage, the *generic* strategy obtained 100%, the group based strategies obtained 97% and 93%, and the *personalized* strategy obtained 85%. **Fig. 3** shows *precision@k* obtained for k ranging from 1 to 9 for the *uniform* weighting model.[2] Similarly, group based strategies are inferior to the *personalized* and superior to the *generic* strategy.

In summary, the accuracy of personalized CF recommendations is highest, but the coverage is lowest. Hence, it is the most accurate strategy once it can be applied. Group strategies have a reasonably good accuracy and coverage; they are applicable once dense group data is available. Generic recommendations can always be generated, although their accuracy is lowest.

Next, we compare the performance of the group based *aggregated models* and *aggregated predictions* strategies. **Table 2** shows that the *F1* score of the *aggregated models*

2. Similar results obtained for other weighting models.

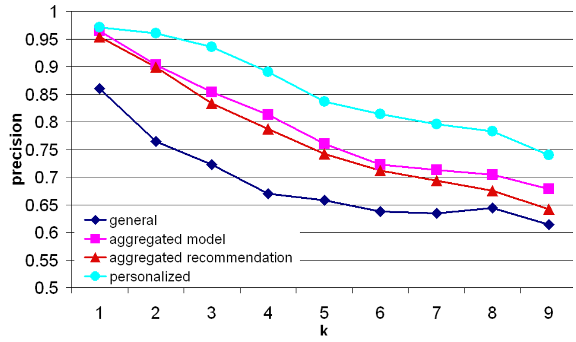

Fig. 3. Recommendation strategies – *Precision@k* vs. *k*.

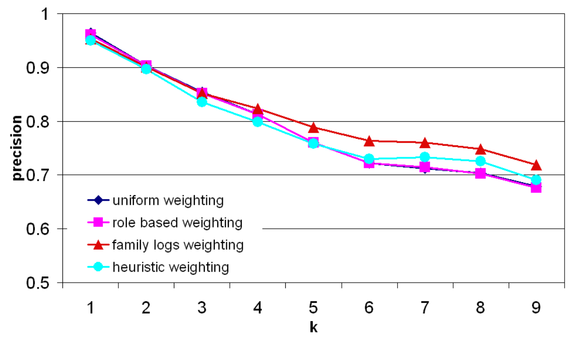

Fig. 4. Weighting models – *Precision@k* vs. *k*.

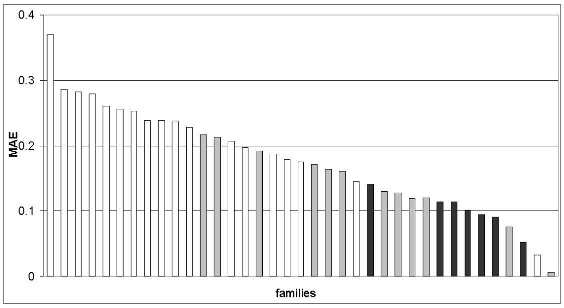

Fig. 5. *MAE* vs. family size.

is higher than that of the *aggregated predictions* strategy across all weighting models. Also the *MAE* score of the *aggregated models* is lower than that of the *aggregated predictions*. This is supported by **Fig. 3**, which shows that the *aggregated models* consistently obtains higher *precision@k* than the *aggregated predictions*. The coverage of the *aggregated models* is also higher than that of the *aggregated predictions* strategy. Hence, aggregating individual user models into group models and generating group recommendations using the group data is preferable to generating individual recommendations using the sparse models and aggregating them into group recommendations.

Next, we examine the effect of the weighting models. The weighting models can be partitioned into two groups: the static *uniform* and *heuristic* models, and the adaptive *role-based* and *family log* models. **Table 2** shows that the adaptive models outperform the static models across both the accuracy metrics (higher *F1* and lower *MAE*). Hence, adaptive weighting models, which assign weights based on the logged user interactions, lead to more accurate recommendations than static models, which assign predefined weights. As expected, the weighting models affect only the accuracy of the recommendations and their impact on the *coverage* is negligible.

Figure 4 shows the *precision@k* obtained for various values of k for the *aggregated models* strategy.[3] For low k, the results of all the models are very high and comparable. The models separate at $k = 3$, with the static models becoming less accurate. Eventually, the adaptive models outperform the static models, with the *family log* model demonstrating the highest *precision@k*. Hence, the *family log* model is the most appropriate weighting model (note its higher *F1* and lower *MAE* scores in **Table 2**) due to its *localized* nature: the weights are assigned according to interactions logged across the family, rather than across the entire community of users.

4.3. Group Size Dependency

In addition to comparing the performance of group recommendation strategies and weighting models, we investigated performance fluctuations across families. We analyze the dependency between the size of a family and accuracy of recommendations.[4]

Figure 5 shows the *MAE* scores obtained for the recommendations generated for various families sorted in decreasing order of *MAE*, where each family is colour coded according to its number of members: white bars represent 2 user families, grey – 3 user families, and black – 4 user families. The accuracy of recommendations generally increases with family size. Most families with high *MAE* are 2 user families, while 3 and 4 user families, i.e., grey and black bars, tend to have lower *MAE*, i.e., more accurate recommendations. The correlation between the *MAE* of recommendations and the number of family members is -0.656. Hence, the accuracy of recommendations improves with the number of family members and amount of data available, since the data of large families are denser than of small families.

5. Conclusions

Data sparsity is a well known problem for CF recommender systems. To resolve it, individual user data can be aggregated into group data, and these can be used in the CF recommendation process. In this work we analyzed CF group recommendations using a dataset of recipe ratings of families of users gathered by an eHealth portal.

3. Similar results obtained for the *aggregated predictions* strategy.

4. These results refer to the *family model* strategy and *family log* weighting model. Similar results obtained for the *aggregated predictions* strategy and other weighting models.

The contributions of the work are four-fold: (1) evaluation of generic, group based, and personalized recommendation strategies, (2) in-depth comparison of two group based strategies, (3) comparison of four weighing models, and (4) analysis of a family size dependency.

The results showed that the group based strategies were superior to the generic, but inferior to the personalized CF strategy. These outcomes were consistent across all the metrics. Comparison of the group strategies showed that aggregating individual models into group models was preferable to aggregating individual recommendations into group recommendations. These outcomes were also consistent across all the metrics. Comparison of the four weighting models showed that weighting users according to the logged interaction was preferable to a predefined weighting, while focusing on group interaction was more accurate than on community interaction. The accuracy of group recommendations was discovered to depend on the size of groups, such that it increased with the number of users in a group.

In the future, we plan to investigate the feasibility of sequential group recommendations. Often, recommendations are not provided on an ad-hoc basis, but users have prolonged interactions with the system. Hence, it is important to the prolonged interactions differently, e.g., to compensate users, whose satisfaction levels were lowest in previous interactions. Also, we plan to investigate group based social dynamics. Different groups may have complex social intra-group relationship at play. We will investigate how these relationships, e.g., family roles, ages, and compromises, affect group recommendations.

Acknowledgements

This research is funded by the Australian Government through the Intelligent Island Program and CSIRO Preventative Health Flagship. The Intelligent Island Program is administered by the Tasmanian Department of Economic Development and Tourism. The authors acknowledge the contributions of Mac Coombe, Nilufar Baghaei, Stephen Kimani, Greg Smith and Dipak Bhandari.

References:

[1] J. Konstan, B. Miller, D. Maltz, J. Herlocker, L. Gordon, and J. Riedl, "GroupLens: applying collaborative filtering to Usenet news." Communications of the ACM, Vol.40, No.3, pp. 77-87, 1997.

[2] U. Shardanand and P. Maes, "Social information filtering: algorithms for automating 'word of mouth,'" Int. Conf. on Human Factors in Computing Systems, pp. 210-217, 1995.

[3] Z. Huang, H. Chen, and D. Zeng, "Applying associative retrieval techniques to alleviate the sparsity problem in collaborative filtering," Trans. on Information Systems, Vol.22, No.1, pp. 116-142, 2004.

[4] A. Jameson and B. Smyth, "Recommendation to groups," P. Brusilovsky, A. Kobsa, and W. Nejdl, eds., The Adaptive Web Methods and Strategies of Web Personalization, pp. 596-627, Springer, 2007.

[5] M. Noakes and P. Clifton, "The CSIRO Total Wellbeing Diet Book," Penguin, 2005.

[6] S. Berkovsky, "Decentralized mediation of user models for a better personalization," Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 404-408, 2006.

[7] J. B. Schafer, J. Konstan, and J. Riedl, "Recommender systems in e-commerce," Int. Conf. on Electronic Commerce, pp. 158-166, 1999.

[8] R. D. Burke, "Hybrid recommender systems: Survey and experiments," User Modeling and User-Adapted Interaction, Vol.12, No.4, pp. 331-370, 2002.

[9] J. Freyne and B. Smyth, "Cooperating search communities," Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 101-110, 2006.

[10] Z. Yu, X. Zhou, Y. Hao, and J. Gu, "Tv program recommendation for multiple viewers based on user profile merging," User Modeling and User-Adapted Interaction, Vol.16, No.1, pp. 63-82, 2006.

[11] J. Masthoff, "Group modeling: Selecting a sequence of television items to suit a group of viewers," User Modeling and User-Adapted Interaction, Vol.14, No.1, pp. 37-85, 2004.

[12] M. O'Connor, D. Cosley, J. A. Konstan, and J. Riedl, "Polylens: a recommender system for groups of users," European Conf. on Computer Supported Cooperative Work, pp. 199-218, 2001.

[13] C. Senot, D. Kostadinov, M. Bouzid, J. Picault, A. Aghasaryan, and C. Bernier, "Analysis of strategies for building user group profiles," Int. Conf. on User Modeling, Adaptivity, and Personalization, pp. 40-51, 2010.

[14] P. Brusilovsky, G. Chavan, and R. Farzan, "Social adaptive navigation support for open corpus electronic textbooks," Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 24-33, 2004.

[15] S. Berkovsky, J. Freyne, and M. Coombe, "Aggregation trade offs in family based recommendations," Australasian Conference on Artificial Intelligence, pp. 646-655, 2009.

[16] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," Trans. on Information Systems, Vol.22, No.1, pp. 5-53, 2004.

Name:
Shlomo Berkovsky

Affiliation:
Research Scientist, Tasmanian ICT Centre, CSIRO

Address:
GPO Box 1538, Hobart, Tasmania 7001, Australia

Brief Biographical History:
2007- Research Fellow, Melbourne University
2008- Research Scientist, CSIRO

Name:
Jill Freyne

Affiliation:
Research Scientist, Tasmanian ICT Centre, CSIRO

Address:
GPO Box 1538, Hobart, Tasmania 7001, Australia

Brief Biographical History:
2006- Research Fellow, University College Dublin
2009- Research Scientist, CSIRO