Social Interaction in the CATS Group Recommender*

Kevin McCarthy, Lorraine McGinty, Barry Smyth & Maria Salamó

Adaptive Information Cluster, School of Computer Science & Informatics,
University College Dublin, Belfield, Dublin 4, Ireland.
{kevin.mccarthy, lorraine.mcginty, barry.smyth, maria}@ucd.ie

Abstract. Co-operative group recommenders aim to help a group of users arrive at a consensus when they need to make a decision in relation to a common goal. The success of any such recommender critically relies on its ability to; (1) gather and accurately model group preferences, (2) enhance group awareness of member preferences, and (3) focus user attention on areas of the recommendation space which are likely to contain recommendation options that are highly relevant to the group as a whole. In this paper we describe how we manage this kind of social interaction within a group travel recommender system through effective use of visual cues and an accurate group preference modelling methodology.

1 Introduction

Our research in the area of social recommender systems concentrates on group recommendation in particular. We are especially interested in co-operative group recommendation contexts whereby the objective is to help a group of users arrive at a consensus when making a decision in relation to a common goal. A good example here is a group of users planning a holiday together; each having their own (often diverse!) *preferences* with respect to what constitutes as an *ideal* holiday for them. The obvious challenge here is how best to aggregate individual preference models, and this is an area of research we have investigated extensively recently [9, 10]. Of course, research by Jameson *et al* [5, 11] has highlighted that supporting users in group recommendation tasks requires a great deal more than identifying an appropriate aggregation function. Having an accurate way of modelling group member preferences only brings us part of the way to supporting users in this kind of a recommendation setting. Other social interaction challenges include: how to support groups of users in the preference elicitation task, how best to educate the group about other member preferences, and how can the recommender draw attention to recommendation candidates that seem to satisfy group preferences.

Our work in this area has greatly been influenced by recent work in the area of social interaction [1, 8], user modelling [6], and intelligent user interfaces [4, 7,

* This material is based on works supported by Science Foundation Ireland under Grant No. 03/IN.3/I361.

13]. For instance, in this paper, we describe how we make extensive use of visual cues to create emphasis and help users locate relevant information, as well as enhance group member awareness of each other's preferences and motivational orientations. We describe our ideas in the context of a group recommender system that we have designed and implemented; the Collaborative Advisory Travel system (CATS). First of all we discuss the social interaction modalities and interface components supported by CATS. Next, in Sections 3 and 4 we summarise our approach to group preference modelling and recommendation retrieval. Section 5 discusses some of the feedback we have collected from user trails, to date, in relation to the effectiveness of the interaction measures we have put in place.

2 Interaction Modalities within our Group Recommender

The interface to the CATS system is shown in Figure 1, implemented as a Web-based client-server application whereby each user interacts using a standard PC. It draws on case base of 5700 European skiing holidays described by 43 features related to the *resort* (e.g., *country, transfer time, lift system, etc.*) and the *accommodation* (e.g., *rating, price, ski room & restaurant facilities, etc.*). A key objective in CATS is to help users understand which ski-packages best suit their needs *and* the needs of all group members. In this section we discuss how individual user and group interaction is supported within CATS; paying particular attention to how we effectively communicate information contained by the group model (through visual means), and how we support group interaction in terms of preference elicitation, annotation and consensus calibration.

2.1 Individual Interaction

In Figure 1 the interface screen, shown to each user, is a map of countries in Europe, with ski resorts marked by mountain range icons. To view a particular resort a user simply clicks on the resort icon. The user is subsequently presented with a *case window* describing a particular skiing package option, see Figure 2. By critiquing a recommendation, a user can express a preference over a specific feature in line with their own personal requirements (e.g., *more red runs*, *cheaper*, *higher star rating for hotel* etc), which affords the user an opportunity to provide informative feedback; see [2–4, 13]. The critiques made by each user are added to their individual preference model and the next case recommended to them is determined by the selection mechanism described in Section 4. Furthermore, when individual users are satisfied with a particular holiday recommendation and wish to draw it to the attention of the other group members, they can do this by adding it to a *stack area*. Further details on the stack aspect of the interface are discussed in the next section.

2.2 Social Interaction

In addition to being able to make *reactive* recommendations to individual users (i.e., on the basis of their individual critiques) CATS also has the capablilty

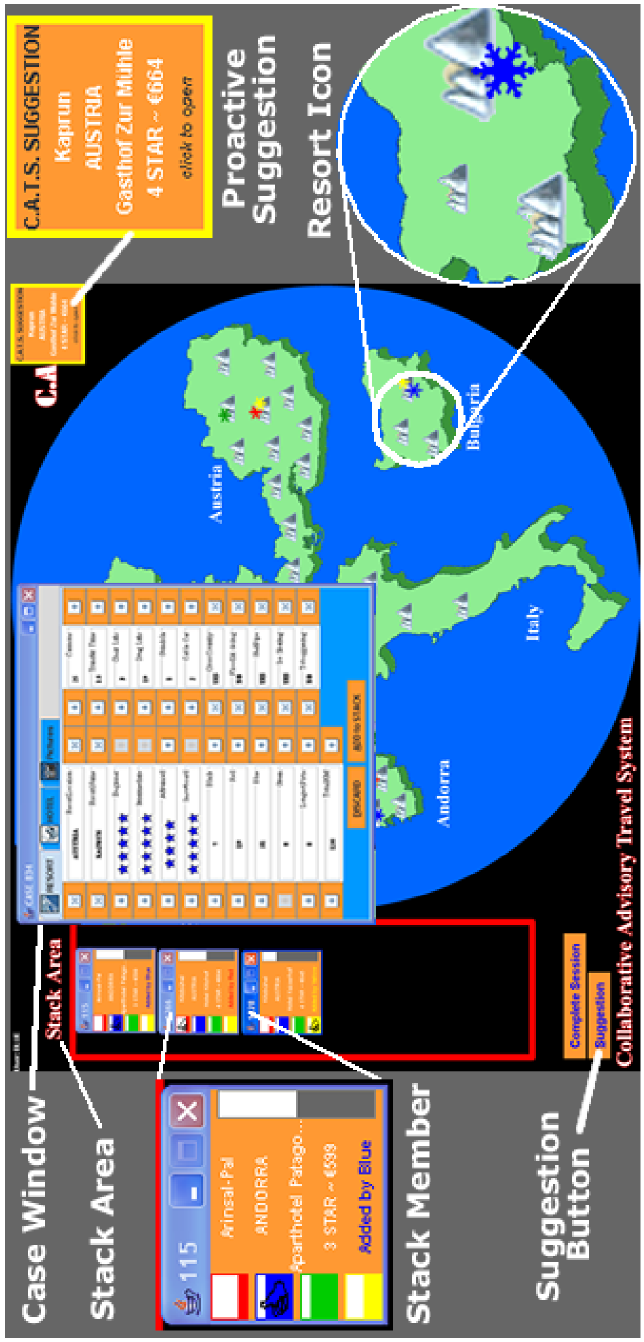

Fig. 1. The main CATS interface.

Fig. 2. The *case window* presents the user with a complete description of a case and is used as the starting point for collecting critiquing-based feedback from each user.

of making *proactive* recommendations to all users (i.e., on the current group preferences model; see Section 3). Briefly, the CATS system constantly compares the preferences of the group with the remaining cases available; that is, cases that have not been previously viewed or discarded by any of the group members. Occasionally, one or more of these cases exceeds a certain critical *compatibility threshold* with respect to the group preference model and when this happens the most compatible case is pro-actively recommended to all users. For example, one such case (for a 4-star hotel in Austria) has been proactively recommended in Fig. 1 and will appear on the map window for all users where they can interact with it in the usual way. The motivation here is to draw their collective attention towards cases that appear to maximally satisfy their preferences.

We briefly mentioned above how each group member can *promote* recommendations to the attention of all group members through the *stack* area of the interface. The stack houses summaries of these case recommendations, as well as displaying compatibility information relating to group compatibility. Essentially, compatibility of each promoted case with the current users individual preference model, IM^U, is shown along the left hand side of each stack member, and a consensus barometer (i.e., wrt the group preference model, $GM^{U_1,\ldots,U_k}$), is presented along the right hand side. This informs all the users in the group about how satisfied the group is with the cases being considered, and helps users to better understand what factors are important to other group members.

The group preference model is also used to dynamically update resort icons on the map interface. Each resort icon is annotated with a colour-coded snowflake to show where each user is concentrated in the product space at the current moment in time, but the resort icon sizes change dynamically also. The size of the resort icon reflects the compatibility rating of its most group-compatible case; that is, the resort case that satisfies the most critiques contained in the group preference model. Thus, the map also communicates the focus of group activity and preferences to all users.

3 Modelling Group Preferences

As mentioned earlier, having an accurate way of aggregating the individual preference models of all users is an essential aspect of any co-operative group recommender. In our work each user U is associated with an *individual preference model*, IM^U, that is made up of the critiques that they apply throughout the course of the recommendation session (see Equation 1).

$$IM^U = \{I_1, ..., I_n\} \tag{1}$$

As new critiques are made by the user, their preference model is updated. This involves the addition of new critiques but may also involve the removal of past critiques if they conflict with, or are subsumed by the most recent critique. For example, if a user had previously indicated a *Price < \$600* critique and a new *Price < \$500* critique is later applied then the earlier critique will be removed to reflect the users refined *Price* preference. Similarly, if a user had previously indicated a *Price < \$600* critique but the new critique is for *Price > \$650*, then the earlier conflicting critique is deleted. In this way the user's preference model reflects their most recent preferences. In addition, a group preference model, $GM(U_1, ..., U_k)$, is also maintained by combining the individual user models and associating critiques with the users who contributed them as shown in Equation 2 such that I_j^U refers to the j^{th} critique in the preference model for user U_i.

$$GM^{U_1,...,U_k} = \{I_1^{U_1}, ..., I_n^{U_1}, ..., I_1^{U_k}, ..., I_m^{U_k}\} \tag{2}$$

During recommendation it will sometimes be necessary (as we will see in the next section) to leverage part of the group preference model, usually the model less some individual user's critiques. Thus we will often refer to the *partial group model* or the *members model*, MM^U, to be the group model without the critiques of a particular user U as shown in Equation 3. This means that the group preference model is based on the preference models for individual users at a given point in time and after they have been processed to remove inconsistent or *redundant* critiques. We have chosen not to repeat this processing over the group preference model and therefore it is possible, indeed likely, that the group preference model will contain conflicting preferences, for example. Of course, during recommendation these inconsistencies will have to be managed by preferring cases that are maximally compatible with the overall group model.

$$MM^U = GM^{U_1,\ldots,U_k} - IM^U \tag{3}$$

4 Generating Group Recommendations

Our approach to recommendation selects items/products for an individual U_i in the context of some group of users G based on U_i's individual preference model and on the group preference model. In this section we will consider recommendations from the point at which U_i critiques a recently recommended case. The critiqued case (c_p) is often referred to as the current preference case. The job of the recommender system is to pick a new case that is compatible with the latest critique while similar to c_p.

The first step is to temporarily filter-out cases that are not compatible with the current critique (applied by U_i) [3] . This leads to a set of *recommendation candidates*. The standard approach to critiquing only ranks these candidates according to their similarity to the critiqued case (c_p), irrespective their compatibility with prior critiques. We, instead, use the incremental critiquing method [12], which uses the user's preference model to influence future recommendations.

The result is that recommendation candidates are ranked such that they are similar to c_p, compatible with the current critique, and also such that they are compatible with past critiques in so far as is possible. To do this, each candidate recommendation, c_r, is scored according to its compatibility to the user's current preference model, as shown in Equation 4. Essentially, this compatibility score is equal to the percentage of critiques in the user's model that are satisfied by c_r; for example, if c_r is a \$1000 vacation case then it will satisfy a *price* critique for less than \$1200 (I_i) and so $satisfies(I_i, c_r)$ will return 1.

$$compatibility(c_r, U_i) = \frac{\sum_{\forall I \in U} satisfies(I, c_r)}{|U|} \tag{4}$$

The *quality* of a case c_r with respect to a preference case c_p, is a weighted sum of preference similarity and critique compatibility. When a user U critiques c_p the next case recommended will be the one with the highest quality score; see Equation 5. By default, for incremental critiquing $\alpha = 0.5$ to give equal weight to preference similarity and critique compatibility.

$$quality(c_p, c_r, U) = \alpha * compatibility(c_r, U) + (1 - \alpha) * similarity(c_p, c_r) \tag{5}$$

$$c_{rec} = argmax_{c_r}(quality(c_p, c_r, IM^U, MM^U)) \tag{6}$$

Our group recommender is based on incremental critiquing but adapted to include the preferences of the other group members ($MM^{U_i} = G - [U_i]$) in the quality metric, as well as the preferences of the user applying the critique, to select a recommendation (c_{rec}) according to Equation 6. To do this, we compute a new compatibility score, for a recommendation candidate c_r, as shown in Equation 7 and combine this with similarity to the preference case (c_p) as in Equation 8. The β parameter controls how much emphasis is placed on individual versus group compatibility while α controls the emphasis that is placed on compatibility versus preference similarity; by default we set both parameters to 0.5. In this way, the case that is recommended after critiquing c_p will be chosen because it is compatible with the critique, similar to c_p, and compatible with both the user's own past critiques and the critiques of other users. Thus we are implicitly treating past critiques as *soft constraints* for future recommendation cycles [14]. It is not essential for recommendation candidates to satisfy all of the previous critiques (individual or group), but the more they satisfy, the better they are regarded as recommendation candidates.

$$\begin{aligned} GCompatibility(c_r, IM^U, MM^U) = \beta * compatibility(c_r, IM^U) + \\ (1-\beta) * compatibility(c_r, MM^U) \end{aligned} \tag{7}$$

$$\begin{aligned} quality(c_p, c_r, IM^U, MM^U) = \alpha * GCompatibility(c_r, IM^U, MM^U) + \\ (1-\alpha) * similarity(c_p, c_r) \end{aligned} \tag{8}$$

5 Evaluating the Social Interaction Modalities in CATS

We carried out a small-scale user study to evaluate our prototype CATS system. Multiple trials, each involving groups of 4 users, were carried out with 12 computer science graduate students, with varying degrees of interest and experience when it came to skiing[1]. Each group of users were instructed to behave as if they were really trying to find a skiing holiday to go on together. For each trial user interactions and recommender activity was recorded, and at the end of each session each group had to complete an extensive questionnaire covering issues such as: their personal level of satisfaction with the final consensus reached and the usefulness of the various social visual cues presented.

In previous work [10, 9] we provide a thorough description of the evaluation setup and methodology. Our previous discussions concentrated on the recommendation accuracy and efficiency results we found; using criteria such as session length and overall preference satisfaction. We found that our approach to group recommendation effectively generates recommendations that satisfy group needs. Here, we would like to focus on our key findings with respect to the *usefulness* of the various social interaction modalities in CATS. It is conceivable that presenting users with dynamically changing interfaces, that incorporate a diverse range of interaction modalities, could have the adverse effect of confusing them rather than supporting them in their co-operative task. We were sensitive to the

[1] All trials were conducted in the computer laboratories at the School of Computer Science & Informatics at UCD Dublin, Ireland

notion that users could well be overwhelmed by the *busy* combination *proactive & reactive recommendations*, *dynamic annotation and resizing of icons*, as well as the various summaries, barometers and indicators in *the ever-changing stack area*. The critical point to make here is that if our user groups do not find these features intuitive and useful then here there is little point in including all of these in the final CATS interface. It is always very difficult to accurately measure subjective functionality of this kind. Like many others, we relied on the feedback we gathered from individual user questionnaires and interaction logs. One of the key things we were interested in was *Do users actually find the stack area useful?* The textual responses gathered in response to this question were extremely positive. Every user (i.e., 100% agreement) found it a very useful resource for drawing the attention of other group members to their own preferred cases, and 75% of users routinely accessed recommendations of other group members from this area. We found that the average user places between 3 and 4 cases on the stack per session, and that the number of case accesses from the stack area varies over the trials (from a min of 15% to a max of 100%).

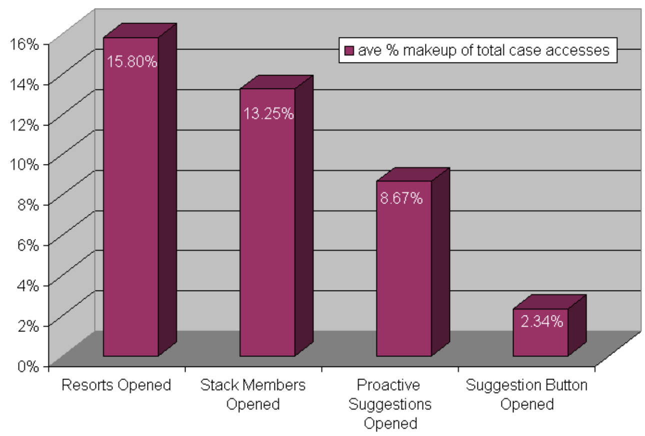

Fig. 3. Comparing the social interaction features of the CATS system.

When asked if they found the proactive recommendations and resort icons annotations and resizing useful users were, once again, largely positive. There was 60% agreement amongst users that the resort icon changes were useful as a means of emphasising to the group the part of the recommendation space that was currently in focus. Incidently, those users who were not in full agreement had neglected to notice the dynamic resizing feature of these resorts. Once a proactive suggestion is made, we found that users respond to these suggestions approximately 26% of the time on average. Only 2 of the 12 users indicated they were not happy with the general quality of these recommendations.

None of our trialists reported feeling confused at any point, although we did find that some users far preferred one interaction modality over another. For instance, those users whose recommendation accesses came primarily from the stack and proactive recommendation component (rather than though direct interaction with resort icons in the map), often did not notice the dynamic changes in resort icons. We do not see this as a negative result. We understood that every user had preferred interaction mode, and that this varied amongst trailists. The fact that some *creatures of habit* concentrated on one or two interaction modes (and was oblivious to others) is okay, so long as they did not do this because they were confused. Overall, our logs revealed that each of the interaction components were utilised considerably more often by users to access cases than the standard 'suggest' button interaction. Figure 3 shows how users are especially attracted to the map and stack areas which play a vital role in communicating strong group preferences.

6 Conclusions

Earlier, in Section 1, we highlighted that while having an accurate way of modelling group preferences and generating relevant recommendations is a critical component for any co-operative group recommender, it is not the *only* challenge. Our Collaborative Advisory Travel system (CATS) takes an approach to cooperative group recommendation that: (1) uses a variety of social interaction features to communicate group, as well as individual, preferences and activity, and (2) constructs a reliable group-preference model by combing critique histories in order to generate recommendations on a proactive and reactive basis. Previous performance evaluations have indicated that CATS effectively translates the often competing preferences of a group of individual users into a recommendation that broadly satisfies the whole group. In this paper we have discussed how we gather preference information, enhance awareness, and focus the attention, of a group of users using visual cues and a sophisticated group preference modelling methodology. In our evaluation trials users responded positively to the various social interfacing elements and recommendation strategies implemented by the CATS group recommender. We have found that user groups make regular use a social interaction features such as the *stack area* and *dynamic resizing of map components* and *resort annotation*, and so seem willing to accommodate emerging group preferences.

References

1. Trent Apted, Judy Kay, and Aaron Quigley. A study of elder users in a face-to-face collaborative multi-touch digital photograph sharing scenario. In *The University of Sydney TR567*, 2005.
2. R. Burke. Interactive Critiquing for Catalog Navigation in E-Commerce. *Artificial Intelligence Review*, 18(3-4):245–267, 2002.
3. R. Burke, K. Hammond, and B.C. Young. The FindMe Approach to Assisted Browsing. *Journal of IEEE Expert*, 12(4):32–40, 1997.
4. B. Faltings, P. Pu, M. Torrens, and P. Viappiani. Design Example-Critiquing Interaction. In *Proceedings of the International Conference on Intelligent User Interface(IUI-2004)*, pages 22–29. ACM Press, 2004. Funchal, Madeira, Portugal.
5. A. Jameson. More than the sum of its members: Challenges for group recommender systems. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, pages 48–54, Gallipoli, Italy, 2004.
6. A. Jameson. User modeling meets usability goals. In Liliana Ardissono, Paul Brna, and Antonija Mitrovic, editors, *UM2005, User Modeling: Proceedings of the Tenth International Conference*, pages 1–3. Springer, Berlin, 2005.
7. A. Jameson, S. Baldes, and T. Kleinbauer. Two methods for enhancing mutual awareness in a group recommender systems. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, Gallipoli, Italy, 2004.
8. Stephen G. Kobourov, Kyriacos Pavlou, Justin Cappos, Michael Stepp, Mark Miles, and Amanda Wixted. Collaboration with diamondtouch. In *10th International Conference on Human-Computer Interaction (INTERACT)*, pages 986–989. Rome, Italy, 2005.
9. K. McCarthy, L. McGinty, B. Smyth, and M. Salamó. The Needs of the Many: A Case-Based Group Recommender System. In M. H. Goker and T. R.Roth-Berghofer, editors, *Proceedings of the European Conference on Case Based Reasoning (ECCBR 2006)*. Springer Verlag, 2006. Turkey.
10. K. McCarthy, M. Salamó, L. Coyle, L. McGinty, and B. Smyth & P. Nixon. CATS: A Synchronous Approach to Collaborative Group Recommendation. In *Proceedings of the FLAIRS 2006 Conference*, pages 86–91. Springer Verlag, 2006. Florida, USA.
11. C. Plua and A. Jameson. Collaborative preference elicitation in a group travel recommender system. In *Proceedings of the AH 2002 Workshop on Recommendation and Personalization in eCommerce*, pages 148–154, Malaga, Spain, 2002.
12. J. Reilly, K. McCarthy, L. McGinty, and B. Smyth. Incremental Critiquing. In M. Bramer, F. Coenen, and T. Allen, editors, *Research and Development in Intelligent Systems XXI. Proceedings of AI-2004*, pages 101–114. Springer, 2004. UK.
13. S. Sherin and H. Lieberman. Intelligent Profiling by Example. In *Proceedings of the International Conference on Intelligent User Interfaces (IUI 2001)*, pages 145–152. ACM Press, 2001. Santa Fe, NM,.
14. M. Stolze. Soft Navigation in Electronic Product Catalogs. *International Journal on Digital Libraries*, 3(1):60–66, 2000.