Towards a Better Understanding of Negotiation in Group Recommender Systems

Tom Gross
Human-Computer Interaction Group
University of Bamberg
Bamberg, Germany
hci@uni-bamberg.de

Abstract

Group recommender systems identify items that fit all group members' preferences. The final step—the group negotiation for finding consensus on the item to choose—is essential for user satisfaction with the system and its outcome. It typically follows the acquisition of information about users' preferences as well as the generation and presentation of recommendations. This paper contributes to a better understanding of requirements for negotiation support in group recommender systems. In particular, we report on preliminary results of our exploratory study on the effect of three different negotiation conditions on user satisfaction.

CCS CONCEPTS

Human-centered computing~Empirical studies in collaborative and social computing

KEYWORDS

Recommender Systems, Negotiation, Negotiation Timing, Consensus Finding.

ACM Reference format:

Tom Gross. 2019. Towards a Better Understanding of Negotiation in Group Recommender Systems. In *Mensch und Computer 2019 (MuC '19), September 8–11, 2019, Hamburg, Germany.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3340764.3344458

1 Introduction

Group recommender systems (GRS) make suggestions based on all group members' preferences. Traditionally, single-user recommender systems (RS) help solitary users to identify adequate goods or services by offering suitable items from a broad range of alternatives (e.g., suggesting book recommendations to online shoppers). GRS maximise the predicted satisfaction of all group members (e.g., suggesting a ranked list of restaurants for eating out in a group).

RS and GRS generate recommendations by calculating predictions of users' satisfaction with individual items (based on algorithms such as collaborative filtering [3]). MusicFX is a prominent early example of a GRS that automatically adapts the music to the present members of a fitness club while exercising [9]. The recommendations are based on aggregations of individual users' preferred music genres. Yet, most RS and GRS follow the principle of suggestion rather than automation—that is, it is the users who evaluate and compare recommendations and make the final choice. Thereby, the system performs four subtasks: acquiring information about users' preferences; generating recommendations; presenting recommendations; and helping to arrive at a consensus [7].

Previous research emphasised that the latter negotiation for finding consensus is key for GRS; as Jameson and Smyth [7, p. 622] write: "With individual recommenders … the decision process … typically takes place within the mind of a single person. With a group recommender, extensive debate and negotiation may be required". It has been pointed out that the negotiation phase has a strong influence on the overall user satisfaction with the recommendations [5, 10].

The time allocated is the most central aspect of negotiation. Already in early literature on group decision support systems (GDSS) Nunamaker et al. point out: "Efficiency … differs depending upon whether the time available is predefined or open-ended." [12, p. 1328]. Still, in today's GRS negotiation has been under-researched, as Salamo et al. write: "many group recommenders do not explicitly support consensus negotiation" and "consensus remains an open issue for group recommenders." [14, p. 599].

Three alternative approaches for negotiation time in GRS have been suggested [7]: (1) the system takes its top-ranked suggestion and users do not have a choice (i.e., no time); (2) one group member moderates the group during the selection process and also manages the timing (i.e., short time); and (3) group members have an unmoderated open discussion until they eventually come to a concluding selection (i.e., long time).

MuC '19, September 8–11, 2019, Hamburg, Germany

ACM ISBN 978-1-4503-7198-8/19/09…$15.00
https://doi.org/10.1145/3340764.3344458

In this paper we report on preliminary results of a study on the negotiation for consensus finding in GRS. The study explores how short time for negotiation compares to no time and long time for negotiation.

2 Related Work

Over the last two decades numerous recommender approaches (i.e., algorithms and systems) have been developed and evaluated.

Jameson [6] suggests a generic four-phased group process: preference specification, recommendation generation, recommendation presentation, and recommendation decision. For the relevant fourth phase, three principal strategies for finding consensus are named. According to the first strategy, the system automatically takes the top-most recommendation and acts accordingly. According to the second strategy, an agent is in the role to make the final decision. According to the third strategy, the group members have a face-to-face discussion until they find consensus. While the authors explain the strategies and prominent system examples in detail, they do not provide information on how to support groups during negotiation.

Herlocker et al. [5] provide concrete metrics for evaluation of single-user recommender systems. As described before, the satisfaction users might experience with a recommender system is not solely based on the accuracy of the recommendation, but also on the usefulness in the respective domain. Examples for usefulness are novelty and serendipity. However, although novelty and serendipity are important values and contribute to the satisfaction, the authors do not to discuss the influence of supporting the decision-making process actively.

MusicFX [9] is a group recommender system and provides present members of a fitness centre with a shared music recommendation. Users specify their individual music preferences on a local workstation computer—that is, they rate genres on a scale from -2 to 2. When users arrive at the fitness centre they authenticate themselves using an electronic member id card. The system calculates recommendations accordingly. In MusicFX, the users do not negotiate on recommendations because the system automatically plays the top-most tracks to the audience. It is an early system and it increased the variety of music played. Overall MusicFX was perceived as positive, however users manipulated their preferences in order to directly influence the music played (i.e., users changed their preferences for polka for playing tracks of that genre). Even though MusicFX does not provide negotiation support, users' behaviour, which aimed to influence the recommendations, can be interpreted as users' need to participate in the decision-making process.

AGReMo [1] is a mobile group recommender system in the domain of movies. Users rate watched movies on a scale from 0 to 10 in order to specify their preferences. The system also includes contextual information (i.e., location and time). It generates recommendations for a group of users who want to watch a movie together. The recommendations are based on the users' preferences and their current location. An agent is responsible for doing the inputs on the mobile device for the group. While AGReMo is a sophisticated mobile system and it leaves the group time for discussion, it is limited in active participation of all group members (i.e., only the agent can control the application).

So, in general, several GRS are based on the idea that the system generates and presents recommendations to the group, yet that it is the group who takes the final decision. Still, to the best of our knowledge the decision time has not been systematically analysed.

2 Method

We conducted a laboratory study to assess how the amount of time allocated to the negotiation process influences the satisfaction of group members with the finally chosen recommendation in the domain of movies.

2.1 Participants

Thirty (23 male, 7 female) students with an average age of 23.63 years (SD = 2.37) were recruited from our university, and were compensated with bonus course credit. Since the task was to choose a movie from a list of recommended movies in a group, the study required that all participants had already been to movie theatres. Each group comprised three participants, a typical size in the movie domain [13]. Participants signed up as groups with classmates so they had some acquaintance with each other. Informed consent was a prerequisite to participate.

2.1 Material and Preparation

Recommendation database: The goal of preparing real-life data (i.e., movies actually shown in the movie theatres in the city and at the time of the study) that fit to the actual groups of participants required to: (1) generate a personal profile as well as recommendations for the individual participants and (2) aggregate individual to group recommendations.

For step (1) each participant was assigned an individual account at the movie recommendation Web site of our project partner moviepilot[1]. Then they were required to rate a minimum of 100 movies they had actually seen (M = 129.17; SD = 16.43) on a scale from 10 (maximal) to 0 (minimum) at least 24 hours before the study, so we could process the profiles that the system generated based on these ratings applying a collaborative filtering algorithm (we used the classical social filtering algorithm with a nearest neighbour approach [13]).

[1] moviepilot GmbH. http://moviepilot.com

In step (2) we aggregated recommendations for each group to a ranked list of movie recommendations and associated prediction scores (i.e., average satisfaction of the respective group on a scale from 10 to 0) with an average length of 47.60 movies (*SD* = 6.38). For the aggregation we used the least misery approach which maximises the minimus [4].

User questionnaires: We used two types of questionnaires. The post-trial questionnaire concerning the decision outcome had a single item 10-point scale (comparable to the scale of moviepilot the participants were already familiar with) ranging from 10 (very satisfied) to 1 (very unsatisfied) asking how satisfied participants were with the outcome of the group decision (i.e., the chosen movie). The post-test questionnaire concerning the decision process used three 10-point scales of system feedback on all participants choices, reversibility of own choice, and possibility to discuss with other group members from 10 (very important) to 1 (very unimportant) and open questions to assess positive and negative impressions on each condition.

Technical support system for group decision-making: We used the audience response system from Turning Technologies[2] that allowed participants to vote for one of three simultaneously presented movies by pressing a button (A, B, or C) on a remote control.

2.1 Procedure during Study

The groups were sitting around a table mounting a large screen that showed the group recommendations. Each movie recommendation presented the title, poster, and short description of the plot. At the beginning participants provided demographic data (gender and age). Then each group did three repetitions for each of the following three conditions:

- No negotiation: participants were presented with a single recommendation, got time to read, and then rated their satisfaction with the recommendation—all without any time for discussion.
- Short negotiation: participants were presented with three recommendations. They were given time to read them without talking to each other. Then they had 30 seconds to discuss the three movies in their group and individually vote for their favourite one using the remote controls. Any type of verbal and non-verbal communication was allowed. We informed them that they could revise their initial choice within the given time and that the group choice did not have to be unanimous. In case of equally distributed votes across the three recommendations, the system selected the one with the highest prediction quality as decision outcome. Participants saw the decision outcome on the screen and rated their satisfaction with it.
- Long negotiation: the maximum time was increased from 30 seconds to 5 minutes with all other settings unchanged: recommendations, discussion and decisions, rating of satisfaction.

At the end of the study, participants completed a post-test questionnaire giving their positive and negative overall impressions.

2.1 Design

In summary, the experiment was a 3 conditions x 3 repetitions within-group design leading to 9 trials per group. The groups preceded the conditions in a fixed order to avoid unwanted irreversible effects of randomisation. It is well known in the literature on experimental design—both within HCI and beyond—that in repeated measure designs randomisation of conditions is vital. This is true in most cases, yet there are exceptions. For instance, when comparing prolonged and increasing use of drugs [2]. Martin writes: 'If you have chosen to make a circumstance into a random variable, you must be sure that it varies in a truly random way, because not all events that appear random are really so.' [8, p. 28]. Our design is based on those reasons—and particularly the fact that loss aversion would lead to unbalanced reactions in randomised conditions [15].

Targeting at high external validity, we used real-life data. Each movie recommendation needed to be unique per trial (i.e., presenting a group with the same movie in more than one trial would add to the overall time to think and negotiate it). Recommendations were distributed systematically across trials.

3 Results

A one-way repeated-measures analysis of variance (ANOVA) was performed to compare the users' satisfaction with the chosen movie in the 3 conditions. It shows that user satisfaction differs across conditions ($F(2, 58) = 26.05$, $p < .01$, $\eta_p^2 = 0.47$). Paired samples t-tests were conducted and reveal significant differences in users' satisfaction between the no negotiation condition and the two others (*t's(29) < -4.65, p's < .01, d's > 0.85*; effect sizes were calculated using Morris and DeShon equation 8 [11]). No significant differences between the two conditions with negotiation allowed could be found. A Bonferroni adjustment was applied to avoid Type I errors. Table 1 depicts the average users' satisfaction ratings and negotiation times per condition.

Variable	Conditions		
	No Negotiation	30 sec Negotiation	5 min Negotiation
Satisfaction	4.72 (1.22)	6.89 (1.54)	7.89 (1.12)
Actual Duration	-	25.08 (4.46)	80.27 (33.63)

Table 1. Satisfaction ratings and negotiation times by condition.

[2] Turning Technologies LLC. http:// turningtechnologies.com

The analysis of the actually used time in conditions with negotiation shows that in (2) a considerable share of the available 30 seconds was used (M = 25.08 s; SD = 4.46). In (3) the used time was below the allowed 300 seconds ranging from 39 s to 160 s (M = 80.27 s; SD = 33.63). Unanimous choice in (1) was 70%, and in (2) was 77%.

The analysis of the post-test questionnaire data reveals the following for the three 10-scale questions: How important was the option to see your team members' choice: M = 5.37, SD = 2.71; How important was the option to change your choice during the discussion process: M = 5.57, SD = 2.49; and How important was the option to communicate with your team members: M = 9.07, SD = 1.41.

The open questions for pros and cons of the three conditions brought answers with respect to recommendations and negotiation. The 3 conditions got the following comments:

No negotiation: 18 (8+, 10-) comments on recommendations and 12 (5+, 7-) on negotiation process. Comments on the positive side were typically were quite short such as 'simple and fast', 'quick results', 'new [previously unknown] movies', but also 'no group dynamics, no group pressure'. Some more critical comments were: 'no option to contradict', 'no compromise possible', and 'problematic if none of the recommendations fit'.

Short negotiation: 0 on recommendations and 18 (10+, 8-) on negotiation process. Here some positive comments were 'fast decisions', 'short timeframe avoids lengthy discussions', and 'possible to find a compromise'. However, some comments were rather critical: 'short duration is difficult for shy people who need to warm up', and 'not enough time for compromise', and 'not suitable for groups of more than three people since some people might talk too much'.

Long negotiation: 3 (0+, 3-) on recommendations and 16 (13+, 3-) on negotiation process. On the positive side participants told us: 'can have discussions', 'can talk about contents as well as reviews', 'no time pressure', and even 'the more time, the better'. Yet, some concerns were: 'five minutes are too long for most decisions', 'too much time', and 'lengthy discussion can still lead to negative outcomes'.

Other comments were diverse and sometimes contradictory—for instance, one participants said: "If you are undecided, it is very helpful to hear the others' preferences.", whereas another participants commented: "In situations with contradictory opinions social mechanisms get activated that move you to give in, which would not happen without a negotiation.". And, finally, another participant pointed out: "The context is important, such as cinema versus DVD versus television, since it is not only the time but also the money you invest.".

4 Discussion and Conclusions

Our study has identified a strong need for communication and a sweet spot of negotiation time of 60-120 seconds in negotiating movies suggested by a GRS. This seems plausible since in this setting we deal with what Nunamaker et al. call "conflicts of viewpoint between essentially friendly parties" as opposed to "conflicts of interest between conflicting parties" [12, p. 1325].

This communication is complementary to other approaches such as the "consensus negotiation strategies" of Salamo et al. [14] who suggest to mainly have the system suggest consensus, based on statistics and content analysis. From the participant's open feedback it became clear that in the domain of movies there is a users' need for understanding the soft facts in the group such as the current mood and motivation to go to the cinema; aspects which cannot be completely covered by the system.

It also corroborates that communication for consensus finding can only follow a profound acquisition of information about users' preferences as well as the generation and presentation of recommendations. In classical GDSS these prerequisites of communication had already been identified as "general principles to structure the negotiation process" such as "generate many alternatives before judging" and "use objective data and criteria" [12, p. 1326].

The visual feedback on the others' choice and the reversibility of the own choice were considered less important. This might be due to the fact that users were in the same location at the same time and could be perceived differently in distributed scenarios where verbal and non-verbal communication is computer-mediated.

Finally, the movie domain is characterised by rather easy decisions on single items (i.e., we did not look at sequences of movies to select), where the complexity of the items is low to moderate (i.e., number of parameters that influence the choice), and the investment of users is rather low (i.e., time and money). For the future it would be interesting to explore more complex domains such as GRS in the travel domain with multiple items, high complexity, and high investment.

ACKNOWLEDGMENTS

We thank the members of the Cooperative Media Lab as well as the participants of the study, and the reviewers for their comments.

REFERENCES

[1] Beckmann, C. and Gross, T. AGReMo: Providing Ad-Hoc Groups with On-Demand Recommendations on Mobile Devices. In Proceedings of the European Conference on Cognitive Ergonomics - ECCE 2011 (Aug. 24-26, Rostock, Germany). ACM, N.Y., 2011. pp. 179-183.

[2] Field, A. and Hole, G. How to Design and Report Experiments. Sage Publications Ltd., London, UK, 2003.

[3] Goldberg, D., Oki, B., Nichols, D. and Terry, D.B. Using Collaborative Filtering to Weave an Information Tapestry. Communications of the ACM 35, 12 (Dec. 1992). pp. 61-70.

[4] Gross, T. Supporting Informed Negotiation Processes in Group Recommender Systems. i-com - Journal of Interactive Media 14, 1 (Jan. 2015). pp. 53-61.

[5] Herlocker, J.L., Konstan, J.A., Terveen, L. and Riedl, J. Evaluating Collaborative Filtering Recommender Systems. ACM Transactions on Information Systems 22, 1 (Jan. 2004). pp. 5-53.

[6] Jameson, A. More Than the Sum of Its Members: Challenges for Group Recommender Systems. In Proceedings of the Working Conference on Advanced Visual Interfaces - AVI 2004 (May 25-28, Gallipoli, Italy). ACM, N.Y., 2004. pp. 48-54.

[7] Jameson, A. and Smyth, B. Recommendation To Groups. In Brusilovsky, P., Kobsa, A. and Nejdl, W., eds. The Adaptive Web. Springer-Verlag, Heidelberg, 2007. pp. 596-627.

[8] Martin, D.W. Doing Psychology Experiements. Cengage Learning, Belmot, CA, 2007.

[9] McCarthy, J.F. and Anagnost, T.D. MUSICFX: An Arbiter of Group Preferences for Computer-Supported Collaborative Workouts. In Proceedings of the ACM 1998 Conference on Computer-Supported Cooperative Work - CSCW'98 (Nov. 14-18, Seattle, WA). ACM, N.Y., 1998. pp. 363-372.

[10] McNee, S.M., Riedl, J. and Konstan, J.A. Being Accurate is Not Enough: How Accuracy Metrics Have Hurt Recommender Systems. In Extended Abstracts of the Conference on Human Factors in Computing Systems - CHI 2006 (Apr. 22-27, Montreal, Canada). ACM, N.Y., 2006. pp. 1097-1101.

[11] Morris, S.B. and DeShon, R.P. Combining the Effect Size Estimates in Meta-Analysis with Repeated Measures and Independent-Groups Designs. Psychological Methods 7, 1 (Mar. 2002). pp. 105-125.

[12] Nunamaker, J.F., Dennis, A.R., Valacich, J.S. and Vogel, D.R. Information Technology for Negotiating Groups: Generating Options for Mutual Gain. Management Science 37, 10 (Oct. 1991). pp. 1325-1346.

[13] O'Connor, M., Cosley, D., Konstan, J.A. and Riedl, J. PolyLens: A Recommender System for Groups of Users. In Proceedings of the Seventh European Conference on Computer-Supported Cooperative Work - ECSCW 2001 (Sept. 16-20, Bonn, Germany). Kluwer Academic Publishers, Dordrecht, 2001. pp. 199-218.

[14] Salamo, M., McCarthy, K. and Smyth, B. Generating Recommendations for Consensus Negotiation in Group Personalisation Services. Personal and Ubiquitous Computing 16, 5 (June 2012). pp. 597-610.

[15] Tversky, A. and Kahneman, D. Advances in Prospect Theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty 5, 4 (1992). pp. 297-323.