Group Decision Support for Requirements Negotiation

Alexander Felfernig, Christoph Zehentner, and Harald Grabner

Institute for Software Technology, Graz University of Technology,
Inffeldgasse 16b, A-8010 Graz, Austria
{alexander.felfernig,christoph.zehentner,harald.grabner}@ist.tugraz.at
http://www.ist.tugraz.at

Abstract. Requirements engineering is one of the most critical phases in software development processes. Requirements are verbalizing decision alternatives which are negotiated by stakeholders. In this paper we present the results of an empirical analysis of the effects of applying group recommendation technologies to requirements negotiation. This analysis has been conducted within the scope of software development projects at our university where development teams were supported with group recommendation technologies when deciding which requirements should be implemented. We summarize the results of this analysis and show how group recommendation can be applied to requirements negotiation.

1 Introduction

Requirements engineering (RE) is considered as one of the most critical phases in software projects and poorly implemented RE is a major risk for the failure of a project [8]. Requirements themselves are a verbalization of decision alternatives regarding the functionality and quality of the software [2]. Related individual as well as group decisions are extremely difficult due to the increasing size of requirement models as well as contradicting preferences of stakeholders [1]. In this paper we analyze the impact of applying group recommendation technologies [9] to improve the quality of decision processes in the context of *requirements negotiation* which is the process of resolving existing conflicts between requirements and deciding which requirements should be implemented. Typical functionalities of group recommender systems are the visualization of the preferences of other group members, recommendations for individual and group decisions, and recommendations for conflict resolutions in the case of inconsistent stakeholder preferences [9]. Our major motivation for applying group recommendation technologies is to improve the *usability* and the *quality of decision support* in requirements engineering environments (especially in the context of requirements negotiation – both are used as subjective measures in our evaluation).

Note that decision models based on rational thinking [11] are not applicable in most requirements negotiation scenarios since stakeholders do not exactly know their preferences beforehand [1]. Furthermore, preferences are not stable but rather change over time which is an important aspect to be taken into account by requirements negotiation environments [1].

For the purpose of supporting preference construction in requirements negotiation we have developed IntelliReq. Teams are allowed to configure the set of requirements that should be implemented. Note that our goal was to develop recommendation technologies which can be flexibly exploited in requirements negotiation; it is not our intention to replace existing requirements negotiation approaches (see, e.g., [3]) but to provide useful extensions.

This paper is organized as follows. In Section 2 we sketch the IntelliReq environment which supports group decision processes in requirements negotiation – for reasons of space limitations we omit screenshots. In Section 3 we present our hypotheses defined for the empirical evaluation of IntelliReq and discuss the corresponding study results. The paper is concluded with Section 4.

2 IntelliReq Environment

2.1 Application Scenario

IntelliReq is a group decision environment that supports computer science students at the Graz University of Technology in deciding on which requirements should be implemented within the scope of their software projects. Typically, a project team consists of 6–8 students who implement a software system with an average effort of about 8 man months. At the beginning of a project, students have to evaluate a set of requirements which have been defined by the course instructors and to figure out which requirements they will implement within the scope of their project (requirements negotiation phase). For example, the task could be the implementation of a tourist recommender application – the corresponding decision alternatives are depicted in Table 1. We will use this simple set of decision alternatives as a working example throughout the paper.

2.2 IntelliReq User Interface & Functionalities

With the goal of supporting the achievement of a common group decision, the IntelliReq user interface supports the following functionalities:

- Each stakeholder is enabled to define, adapt, and store his/her preferences (choices) regarding a given set of decision alternatives (see, e.g., Table 1).
- Each stakeholder can comment on, argue for, and discuss defined preferences.
- Each group can view and discuss recommendations for group decisions determined on the basis of already defined user preferences.
- Define and store a group decision; this is allowed for project managers.
- Each IntelliReq user can evaluate the application; this user feedback has been analyzed within the scope of an empirical study.

ID	Question	Decision Alternatives
1	which application domain?	20 destinations in Austria; 20 world-wide
2	which type of persistence management?	relational databases; XML; Java objects
3	which type of user interface?	text-based; Java Swing; Web application
4	which recommendation algorithms?	knowledge; collaborative & content-based
5	evaluation by whom?	by: students; other universities; instructors
6	type of user manual?	HTML-based; .pdf based
7	type of acceptance procedure?	live-demo; presentation with screenshots

Table 1. Example decisions to be taken by the project teams.

	with recommendation	without recommendation
preference view	version 1	version 3
no preference view	version 2	version 4

Table 2. The 4 *IntelliReq* versions. The variation points are: *group recommendation supported* (*yes/no*) and *preferences of other team members are visible* (*yes/no*).

3 Empirical Study

In order to evaluate the provided IntelliReq functionalities, we conducted an empirical study within the scope of the course *Object-oriented Analysis & Design* organized at the Graz University of Technology. The major focus of this study was to analyze the impact of group decision technologies on the dimensions *usability* of the system and *quality of decision support*.

3.1 Study Design

For the purpose of the empirical study we provided the IntelliReq environment in *four* versions. In order to analyze our hypotheses, we decided to implement a 2x2 study with the variation points *group recommendations available – recommendations are determined by majority voting* (*yes/no*) and *preferences of other users visible* (*yes/no*) – these versions are shown in Table 2. Both, group recommendations and preference visibility, are key functionalities provided by state of the art group recommendation environments [9, 13]. On the basis of this empirical study we wanted to investigate to which extent these functionalities are applicable within the scope of requirements negotiation.

N=293 participants (computer science students at the Graz University of Technology, 23.1% female and 76.9% male) selected their preferred requirements using the IntelliReq environment. The participants of the study were assigned to one of 56 different groups (the development teams) and defined (stored) 3733 individual preferences and 101 group decisions. For each development team the last stored group decision was interpreted as the final decision; after the published deadline no further adaptations of the taken decisions were possible. After a user had successfully articulated his/her requirements, he/she had the possibility to give feedback on the *usability* and the *decision support quality* of IntelliReq on a 10-point Likert scale.

3.2 Study Hypotheses

The empirical study is based on hypotheses derived from existing research in requirements engineering [1, 3], group recommender systems [9, 5], and decision & social psychology [4, 6, 12]. The list of hypotheses is shown in Table 3.

hypothesis	description
H1	*group recommendations* improve system usability
H2	*group recommendations* improve quality of decision support
H3	*group recommendations* trigger more discussions
H4	*preference visibility* deteriorates perceived usability
H5	*preference visibility* deteriorates perceived quality of decision support
H6	*preference visibility* triggers less preference adaptations
H7	*preference visibility* triggers a decision bias
H8	winning strategy: use *group recommendation* but no *preference visibility*
H9	unconsidered preferences: –usability and –quality of decision support

Table 3. Hypotheses used for evaluating the IntelliReq environment.

Group Recommendation (Hypotheses 1–3) Existing research in the field of recommender systems [9, 5] points out the potential of group recommendation technologies to significantly improve the quality of group decision processes. *First* we wanted to investigate the potential of group recommendation technologies to improve the quality of the dimensions *usability* and *decision support* in a requirements negotiation scenario. With *Hypothesis 1* we express the assumption that recommendation technologies can improve the overall system quality in terms of *usability*. *Hypothesis 2* expresses the assumption that recommendation technologies can help to improve the perceived *quality of decision support*. *Second* we wanted to know whether the availability of group recommendations has an influence on the frequency of applying discussion functionalities (*Hypothesis 3*) – the underlying assumption is that the availability of group recommendations intensifies discussions between group members. This phenomenon is well known and exploited by critiquing-based recommenders where the system proposes recommendations and the user can give feedback in terms of critiques [14]. Studies in social psychology show that frequent information interchange can improve the quality of group decisions [6, 12].

Visible User Preferences (Hypotheses 4–7) Existing research in the field of group-based recommendation points out the advantages of preference transparency in group decision making [9]. In contrast, literature in social psychology points out the fact that suboptimal outcomes of group decision processes are correlated with the visibility of individual preferences of other group members [12, 6]. The reason for groups not being able to take optimal decisions (hidden-profile identification problem) is explained by an insufficient discussion of unshared information which is triggered by the initial disclosure of individual preferences (focus shift from information interchange to preference comparison). *First* we wanted to investigate whether the group-wide visibility of individual preferences has an influence on the perceived usability and decision support quality (*Hypotheses 4* and *5*). *Second* we wanted to figure out whether the group-wide visibility of individual preferences has an influence on the frequency of preference adaptation (*Hypothesis 6*). One underlying assumption here is that persons follow the phenomenon of *social proof* [4], i.e., are doing or accepting things that others already did (accepted). The other underlying assumption is that persons tend to stick with their current decision due to the phenomenon of *consistency* [4], i.e., the effect that published personal opinions are changed less often. *Third*, a lower frequency of information exchange can lead to a different decision outcome [6]. With *Hypothesis 7* we wanted to investigate whether the group-wide visibility of preferences can lead to a decision bias (due to *social proof* [4]), i.e., whether preference visibility has an influence on the decision outcome.

Winning Strategy (Hypothesis 8) We wanted to provide an answer to the question which of the four different IntelliReq versions will be evaluated best regarding usability and quality of decision support. With *Hypothesis 8* we want to express the assumption that group recommendations improve system usability and decision support quality. In contrast, making preferences of other group members visible in the group decision process deteriorates the system evaluation. Consequently, *version* 2 (see Table 2) should be evaluated best.

Distance Matters (Hypothesis 9) Finally, we wanted to provide an answer to the question whether the distance of a users's preference to the final group decision has an impact on the overall system evaluation. With *Hypothesis 9* we express the assumption that users with a low number of considered requirements will not be satisfied with the system usability and the decision support quality.

Group recommendation heuristics The *majority* rule is a simple but very effective heuristic in group decision making [7]: each decision is taken conform to the majority of the votes of the team members. In addition to the majority rule, there exist a couple of heuristics which can be applied when generating recommendations for groups, for example, the *fairness* heuristic which guarantees that none of the group members will be disadvantaged. In the final part of our empirical study we will compare the *prediction quality* of different group recommendation heuristics in the context of our requirements negotiation scenario.[1]

3.3 Study Results

In order to identify statistically significant differences in the user quality feedback depending on the used IntelliReq version we conducted a series of two-sample t-tests. We will now discuss the results of our analysis.

Hypothesis H1 has to be rejected since the *usability* of IntelliReq versions with recommendation support (version 1 and version 2 in Table 2) is only better on the descriptive level (p=0.17, avg. 7.0, std.dev. 1.17) compared to versions without a recommendation support (avg. 6.42, std.dev. 2.47).

[1] Note that due to limited number of subjects (N=293) we were not able to compare the different recommendation heuristics w.r.t. the dimensions usability and quality of decision support. Such comparisons will be in the focus of future work.

Hypothesis H2 can be confirmed since we could detect a significant better evaluation of the IntelliReq *decision support* for recommendation-enhanced versions ($p<0.001$, avg. 7.07, std.dev. 2.03) compared to versions without a recommendation support (avg. 5.21, std.dev. 2.96).

Hypothesis H3 can be confirmed as well since the number of comments on individual preferences is significantly higher in versions which provided group recommendations ($p<0.0015$, avg. 7.96, std.dev. 5.90 vs. avg. 3.53, std.dev. 2.71). Thus we can interpret group recommendations as a stimulating elements for information interchange among group members which is a key factor for high-quality group decisions [12, 6].

Hypotheses H4 and *H5* can not be confirmed since users with no access to the preferences of other group members did not provide a significantly better rating for usability and quality of decision support. However, on the descriptive level the evaluation of versions *without* preference visibility for all group members is better (e.g., usability, avg. 7.0, std.dev. 2.08) compared to versions that make preferences visible (e.g., usability, avg. 6.46, std.dev. 2.09).

Hypothesis H6 can be confirmed since the number of adapted individual preferences is significantly *lower* in versions with access to the personal preferences of other group members ($p<0.001$). This can be explained by the fact that – due to preferences visible for other users – the current user inclines to be *consistent* [4] with his/her original requirements, i.e., the willingness to change articulated preferences decreases if preferences are accessible for other users [4].

Hypothesis H7 can be confirmed since users having access to the preferences of other group members articulate preferences which are more similar to the final group decision (avg. 0.28, std.dev. 0.09 vs. avg. 0.43, std.dev. 0.13). Being confronted with the preferences of other group members, persons base their decisions on the already known preferences and do not focus on a discussion of unshared information which is extremely important for finding optimal decisions [6]. There is a significant biasing effect due to the visibility of preferences ($p<0.001$). This effect can be explained by the phenomenon of social proof [4] which triggers group members to do things or accept things that other group members are doing (accepting).

Hypothesis H8 can not be confirmed. However, users with recommendation support and without insight into the preferences of other users provided the highest ranking for both, *usability* (avg. 7.62, std.dev. 1.84) and *quality of decision support* (avg. 7.11, std.dev. 2.06). Versions with recommendation support outperform versions without recommendation support in terms of *decision support quality* ($p<0.001$) and versions with recommendation support and without a view on the preferences of other users clearly outperform all other versions in terms of *usability* ($p<0.001$).

Hypothesis H9 can be confirmed since users with preferences having a higher distance from the final group decision rated the IntelliReq environment significantly worse in terms of *usability* ($p<0.05$). This result conforms to the *win-lose* situations discussed in [3] which typically turn into *lose-lose* situations. We could not detect a difference in the evaluation of the *quality of decision support.*

3.4 Comparison of Group Recommendation Heuristics

In our empirical study we applied the *majority voting* heuristic [7] for determining group recommendations. In addition to the majority heuristic we wanted to evaluate and compare different other group recommendation heuristics (see, e.g., [10]) w.r.t. to their applicability for our requirements negotiation scenario – for our comparison we used the following ones:

- RAND (randomized recommendation): a recommendation where each individual prediction has been generated randomly.
- LDM (least distance member): the preferences (selections) of the group member with the lowest distance to the preferences of all other group members is used as the group recommendation.
- FAIR (fairness): at least one preference of each group member is taken into account when generating the group recommendation.
- MP (most pleasure): for each question (see, e.g., Table 1) each possible answer is rated regarding its difficulty (in our case in terms of effort in man-months estimated by instructors). The alternative with the lowest overall difficulty is used as group recommendation.
- GBCF (group-based collaborative filtering): group decisions (of other groups) which are similar to the personal preferences of the members of the current group are used as group recommendation.
- MAJ (majority voting): decisions (preferences) supported by a majority of group members are integrated in the final group recommendation.
- MIN (minority voting): decisions (preferences) supported by a minority of group members are integrated in the final group recommendation.

On the basis of the data (individual preferences and taken group decisions) we compared these seven decision heuristics w.r.t. their *prediction quality* (see Table 4). This evaluation shows that (as expected) RAND and MIN should not be taken into account as serious heuristics for predicting user preferences. For our dataset, the MAJ heuristic outperforms all other decision heuristics in terms of the *average distance between predicted and actual group decision*.[2]

4 Conclusions

In this paper we have presented the results of an empirical study which investigated the impact of group recommendation technologies applied in the context of requirements negotiation. We introduced the IntelliReq decision support environment which is used at the Graz University of Technology for supporting group decision processes in small-sized software projects (6–8 team members). The major results of this experiment where that group recommendation technologies can improve the perceived usability and quality of decision support. It is not recommended to disclose the preferences of individual group members at the beginning of a decision process since the knowledge of the preferences of other group members can result in an insufficient discussion of unshared information.

[2] The IntelliReq dataset is available (anonymized): www.ist.tugraz.at/ase/intellireq.

heuristic	avg. dist. (all)	avg. dist. (rec.)	avg. dist. (no rec.)
RAND	0.55 (0.04)	0.55 (0.04)	0.55 (0.04)
LDM	0.31 (0.21)	0.26 (0.22)	0.35 (0.19)
FAIR	0.35 (0.23)	0.33 (0.24)	0.38 (0.22)
MP	0.47 (0.20)	0.47 (0.18)	0.46 (0.23)
GBCF	0.31 (0.19)	0.31 (0.21)	0.33 (0.17)
MAJ	0.27 (0.18)	0.22 (0.19)	0.32 (0.16)
MIN	0.80 (0.16)	0.81 (0.17)	0.79 (0.16)

Table 4. Average distances of recommended group decisions to the final group decision. Distances are measured in terms of the share of individual predictions different from the group decision (*rec.* = IntelliReq versions 1 and 2; *no rec.* = IntelliReq versions 3 and 4; *all* = all IntelliReq versions).

References

1. B. Alenljung and A. Persson. Decision-making activities in the requirements engineering decision processes: A case study. In *ISD 2005*, pages 707–718, 2005.
2. A. Aurum and C. Wohlin. The fundamental nature of requirements engineering activities as a decision-making process. *Information and Software Technology*, 45(14):945–954, 2003.
3. B. Boehm, P. Gruenbacher, and R. Briggs. Developing groupware for requirements negotiation: Lessons learned. *IEEE Software*, 18(3):46–55, 2001.
4. R. Cialdini. The science of persuasion. *Scientific American*, (284):76–81, 2001.
5. A. Felfernig, M. Mandl, M. Schubert, W. Maalej, and F. Ricci. Recommendation and decision technologies for requirements engineering. In *2nd Intl. Workshop on Recommendation Systems for Software Eng. (RSSE10)*, pages 11–15, 2010.
6. T. Greitemeyer and S. Schulz-Hardt. Preference-consistent evaluation of information in the hidden profile paradigm: Beyond group-level explanations for the dominance of shared information in group decisions. *Journal of Personality and Social Psychology*, 84(2):332–339, 2003.
7. R. Hastie and R. Kameda. The robust beauty of majority rules in group decisions. *Psychological Review*, 112(2):80–86, 2005.
8. H. Hofmann and F. Lehner. Requirements engineering as a success factor in software projects. *IEEE Software*, 18(4):58–66, 2001.
9. A. Jameson, S. Baldes, and T. Kleinbauer. Two methods for enhancing mutual awareness in a group recommender system. In *ACM Intl. Working Conference on Advanced Visual Interfaces*, pages 48–54, Gallipoli, Italy, 2004.
10. J. Masthoff. Group recommender systems: Combining individual models. In F. Ricci, L. Rokach, B. Shapira, and P. Kantor, editors, *Recommender Systems Handbook*, pages 677–702. Springer, 2011.
11. D. McFadden. Rationality for economists? *Journal of Risk and Uncertainty*, 19(1):73–105, 1999.
12. A. Mojzisch and S. Schulz-Hardt. Knowing other's preferences degrades the quality of group decisions. *Jrnl. of Personality and Social Psych.*, 98(5):794–808, 2010.
13. M. O'Connor, D. Cosley, J. Konstan, and J. Riedl. Polylens: A recommender system for groups of users. In *European Conference on Computer-Supported Cooperative Work*, pages 199–218, 2001.
14. P. Pu and L. Chen. User-involved preference elicitation for product search and recommender systems. *AI Magazine*, 29(4):93–103, 2008.