Using Group Recommendation Heuristics for the Prioritization of Requirements

Gerald Ninaus
Institute for Software Technology
Graz University of Technology
Inffeldgasse 16b
8010 Graz, Austria
gerald.ninaus@ist.tugraz.at

ABSTRACT

Group recommendation heuristics have been successfully applied in different domains such as Interactive Television and e-Tourism. Our work focuses on the improvement of requirements prioritization techniques on the basis of group recommendation technologies. First, we analyse the impact of preference visibility on the outcome of the requirements prioritization process. In this context we evaluate the resulting software quality and the satisfaction of stakeholders with the requirements prioritization. Second, we analyse standard group recommendation heuristics on a dataset originated from a study conducted at our university. Finally, we propose new heuristics to improve the prediction quality and conclude with an evaluation of the different recommendation heuristics.

Categories and Subject Descriptors

D.2 [**Software Engineering**]: Requirements Engineering; D.2.1 [**Requirements/Specifications**]: Requirements Negotiation; H.5 [**Information Interfaces and Presentation**]: Modelling Environments

General Terms

Requirements Engineering, Human Factors, Experimentation

Keywords

Requirements Prioritization, Recommender Systems, Group Recommendation, Group Decision Making

1. INTRODUCTION

Limited resources in software projects make it necessary to prioritize requirements [7] [13], which eases the planning of subsequent software releases. Furthermore, requirements prioritization supports the conflict resolution between stakeholder preferences. Because of the complexity of the requirements prioritization task, this process requires collaborative work contributed by multiple stakeholders. Systematic requirements prioritization is crucial for the successful completion of a software project since then the available resources are used only for the most relevant requirements. However, establishing consensus about the prioritization of requirements is a challenging task. There are many effects which must be considered, such as cognitive dissonance reduction [5], group polarization [14], and the primacy effect [3]. To simplify this task, tools can recommend prioritizations of requirements as a function of the team preferences. However, if such a prediction of a group decision is only based on the initial preferences of the stakeholders, the prediction quality will be low [11]. A challenge in this context is changing group member preferences during the decision-making process [11]. Another challenge is the primacy effect in preference elicitation [3]: the outcome of this phase will depend on the sequence in which preferences have been inserted. The psychological literature shows that consensus about topics formed early in discussions is cognitive resistant to changes. Additional information added later will be assimilated to already chosen consensus and it is very unlikely that another option is chosen [8]. This phenomenon can be explained by the *assimilating effect* based on the dissonance theory [5] which states *that individuals are motivated to reduce psychological incongruity or discrepancy* that may arise by adding new information to a present perception [2]. The result is that stakeholders will perceive already selected options more attractive than new options [2] and this leads to a bias of group preferences depending on the order of the incoming preferences. Unfortunately, this effect is increased if there is a high group identity, because the fear of exclusion from the group is higher [8]. To reduce this effect, a brainstorming phase in which stakeholders become aware of their own preferences should be established. To raise the willingness of stakeholders to report their honest concerns, this brainstorming phase should be implemented in an anonymous fashion [6]. In a software environment, for example, this can be done if requirements can be added without authentication. First related study results will be presented in the following Sections.

An important issue in the group prioritization process is the factor of *fairness*. The degree of perceived fairness influences the willingness of group members to accept compromises in the resolution of disputes and their trust in other stakeholders [8]. Especially in environments with a high amount of requirements it is necessary to provide tool support in the prioritization process to achieve a maximum degree of perceived fairness. One way to provide such a tool support is to present recommendations of reasonable requirements prioritizations to decision makers.

In our work we investigate different *recommendation heuristics* with respect to the achieved *prediction quality*. We also show how anonymity (preferences are not connected to stakeholder names)

RecSys'12, September 9–13, 2012, Dublin, Ireland

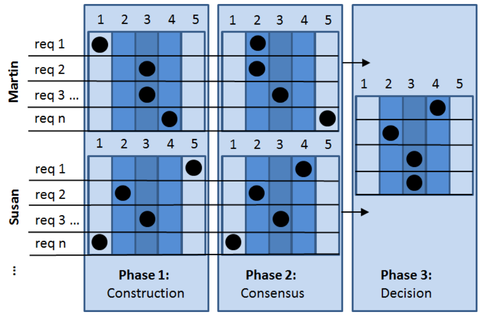

Figure 1: The three decision making phases in IntelliReq.

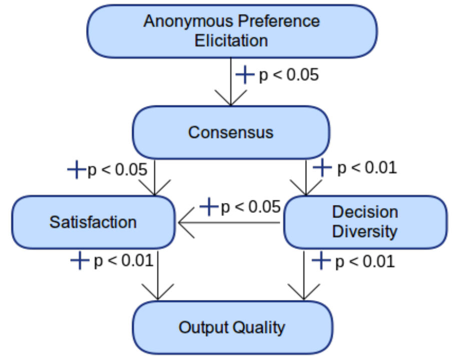

Figure 2: The result of the IntelliReq study (N=39 software teams of size 5-6) as an impact graph. Starting with Anonymous Preference Elicitation each following phase is influenced according to the arrows.

in group decision processes can help to improve the quality of requirements prioritizations. The major contributions reported in this paper are to show that anonymity in decision making influences the quality of the prioritization and to present new heuristics for group recommendations with better prediction quality.

The remainder of the paper is organized as follows. In Section 2 we provide an overview of the IntelliReq decision support environment developed at the Graz University of Technology. In Section 3 we show the impact of the factors *anonymity*, *consensus*, and *decision diversity* on the *quality of the prioritization*, the *stakeholder satisfaction* with the prioritization, and the *quality* of software artefacts. Section 4 covers standard recommendation heuristics including our new proposed heuristics. Section 5 gives an overview of the empirical study conducted at our university. In Section 6 we conclude the paper with an outlook on future work.

2. INTELLIREQ DECISION SUPPORT

IntelliReq is a group decision support tool developed at our university. One feature is the group based requirements negotiation support which will be discussed in detail in the remainder of this paper. The requirements prioritization process consists of three different phases (see Figure 1). The first phase is called *construction* in which the individual stakeholder preferences are collected. The second phase is denoted as *consensus* and is used to discuss and adapt stakeholder preferences. In the third phase (*decision*), the project manager finalizes the group prioritization and can attach explanations for decisions. This explanation can be seen by all group members. To evaluate the impact of anonymized preference elicitation (it is hidden which stakeholder defined which preference), IntelliReq supports two different modes of interaction. One mode uses anonymised preference elicitation, in the other mode user preferences are visible.

3. ANONYMITY, DIVERSITY, AND OUTPUT QUALITY

Within the scope of our study we evaluated the impact of the factors *anonymity*, *consensus*, and *decision diversity* (the standard deviation between the final prioritization results of different requirements) on the *output (software) quality* and the *satisfaction* of the stakeholders with the prioritization. The empirical study has been conducted within the course *Object-oriented Analysis and Design* (N=39 software teams of size 5-6; 15.45% female, 84.55% male) and showed that there are impacts as can be seen in Figure 2. These impacts can be summarized as follows:

- *Anonymous preference elicitation increases consensus of the group:* The hypothesis is that anonymous preference elicitation helps to decrease the commitment [1] which is the degree of a person's unwillingness to change an already articulated opinion. Our study showed that the anonymous preference elicitation leads to a significantly higher degree of consensus between group members.
- *Decision diversity influences the satisfaction of stakeholders:* Our study revealed a significant correlation between a high standard deviation in the prioritization of different requirements and the satisfaction of the stakeholders with the prioritization of the requirements.
- *Consensus increases decision diversity:* Our study results show that a high consensus leads to a high decision diversity. We hypothesise that a dissent makes it necessary to find a compromise in the group. This compromise will often be oriented towards the average of the initial group member preferences. If there are many compromises necessary, the final group prioritization will be located in the middle of the scale. This leads to a small standard deviation in the prioritization of the requirements.
- *Consensus increases satisfaction of stakeholders:* Our study showed that groups with a high consensus regarding the requirements are more often satisfied with the prioritization.
- *Satisfaction and decision diversity increase the output quality:* When evaluating the quality of the software developed by the different groups (teams), we could detect a significant increase of the output quality of software artefacts in situations where there was a high decision diversity and satisfaction with the requirements prioritization.

4. GROUP DECISION HEURISTICS

We hypothesise that the effects described in the previous section could be enhanced by the use of group recommendation heuristics. Consequently, we want to evaluate the best heuristics to predict stakeholder voting behaviour in the requirements prioritization process. In the afore mentioned study conducted at our university we collected a dataset of approximately 10.000 decisions including the preferences of stakeholder and the resulting prioritizations of the software teams. With this dataset we evaluated different group recommendation heuristics [9] and compared predicted with real decisions taken by the participants of our study. For the evaluation we used the precision metric according to Formula 1. The following subsections describe the different applied heuristics that are based on following parameters: **median** m**, standard deviation** sd**, heuristic** h and **requirement** r.

$$precision(h) = \frac{\#\ correctly\ predicted\ group\ preferences(h)}{\#\ predicted\ group\ preferences(h)} \quad (1)$$

Least Distance

This heuristic determines (recommends) for each requirement r_i the priority value p with the lowest distance to the other elements in the set of distinct user preferences (PREF) where $p \in PREF$. This criteria is formalized in a corresponding evaluation function (see Formula 2 and Table 1).

$$Priority(r) = selectmin(p, \sum_{pref \in PREF} |pref - p|) \quad (2)$$

	Martin	Susan	Peter	Pauline	recom. priority
r_1	1	3	1	**2**	**2**
r_2	5	**4**	**4**	3	**4**
r_3	**2**	1	3	1	**2**

Table 1: Application of *Least Distance* heuristic.

Majority Voting

This heuristic recommends a priority value which represents the majority of stakeholder votes related to a specific requirement. An example for the recommendation result of the majority heuristic is given in Table 2.

	Martin	Susan	Peter	Pauline	recom. priority
r_1	**1**	3	**1**	2	**1**
r_2	5	**4**	**4**	3	**4**
r_3	2	**1**	3	**1**	**1**

Table 2: Application of *Majority Voting* heuristic.

Average Value

This heuristic determines the average value (see Formula 3) and round the result (see Formula 4) of the declared stakeholder preferences for each requirement. This value is taken as a recommendation of the priority value for the corresponding requirement.

$$AVG(r) = \frac{\sum_{i=1}^{\#user} pref(i,r)}{\#user} \quad (3)$$

$$Priority(r) = \begin{cases} rounddown(AVG(r)) & AVG(r) < 0.5 \\ roundup(AVG(r)) & AVG(r) \geq 0.5 \end{cases} \quad (4)$$

Random Priority Selection

The random heuristic has been integrated only for evaluation purposes. This heuristic has - as expected - the weakest performance, which can be seen in the *Study Result* Section.

Median Based

Inspired by a survey of the voting behaviour of six-person juries conducted by the University of Chicago [12], we implemented a new heuristic called *Median Based Heuristic* (see Formula 5). In the original study [12], jury members were asked about their initial preferences for the penalty. These initial punishment preferences were compared with the final decision after group deliberating. The finding was that *there was a severity shift for the high-punishment cases, and a leniency shift for the low-punishment cases* [12]. Such choice shifts are explained by the *Group Polarization Theory* [14].

In a similar fashion, our algorithm calculates the recommendations depending on the median of the initial preferences. In our study, the preferences are distributed on a six point scale. In this context, the algorithm has three possible states (see also Formula 5):

- The median is one or two: Calculate the average (see Formula 3) of the preferences and round **down** the result
- The median is three or four: Use the *Least Distance* heuristic (see Formula 2).
- The median is five or six: Calculate the average (see Formula 3) of the preferences and round **up** the result

$$Priority(r,m) = \begin{cases} rounddown(AVG(r)) & m = 1,2 \\ LDIS(r) & m = 3,4 \\ roundup(AVG(r)) & m = 5,6 \end{cases} \quad (5)$$

Ensemble Based

For a further improvement of the prediction quality we introduced a combination of different heuristics. For each recommendation task the algorithm calculates the *Majority* (MAJ), the *Least Distance* (LDIS), and the *Median Based* (MDB) heuristic. If two heuristics recommend the same result, this result is the final recommendation. If there are three different recommendations, the *Median Based* heuristic is the final recommendation. The heuristic is shown in Formula 6.

$$Priority(r,m) = \begin{cases} MDB(r,m) & MDB(r,m) = MAJ(r) \\ MDB(r,m) & MDB(m,r) = LDIS(r) \\ LDIS(r) & LDIS(r) = MAJ(r) \\ MDB(r,m) & else \end{cases} \quad (6)$$

Standard Deviation Based

The hypothesis for this heuristic is that the best heuristic depends on the degree of conformity of the initial preferences. For example, if the group has a high degree of consensus, the *Least Distance* heuristic will perform best. If there is a high dissent, the *Average* heuristic will perform better. The highest standard deviation in our dataset is 2.50. We divided the dataset into three subsets. Each subset has the same standard deviation range ($\frac{2.50}{3} \approx 0.84$). Next we tested the heuristics mentioned in this paper on the dataset to find the heuristic which performs best on each subset (see Table 3). For this evaluation we used the complete dataset of the *Consensus* and *Decision* phases (see Figure 1) which included the anonymous and the non-anonymous setting. The combination of the different heuristics can be found in Formula 7.

$$Priority(sd,r,m) = \begin{cases} LDIS(r) & sd < 0.84 \\ MDB(r,m) & 0.84 \geq sd < 1.67 \\ AVG(r) & 1.67 \geq sd \leq 2.50 \end{cases} \quad (7)$$

Thereafter we generated a function which uses the heuristic depending on the standard deviation and the results of Table 3.

Group	SD From	SD To	Best Heuristic
1	0	0.84	LDIS
2	0.85	1.67	MDB
3	1.68	2.50	AVG

Table 3: Groups based on *Standard Deviation*.

5. EMPIRICAL STUDY

An overview of the study results can be found in Table 4.[1] Although the *Median Based* heuristic is outperformed by the *Least Distance* (see Table 4), the *Median Based* performs better in an environment with visible preferences. In the study conducted at the University of Chicago [12], on which this heuristic is based on, the same observation was made in decision making processes where group member preferences are visible. The *Standard Deviation* heuristic is out of competition as this heuristic is defined for this specific dataset. Future studies with new datasets will show whether this combination of different heuristics will have an improved prediction quality. When comparing the prediction quality of the remaining heuristics, the *Ensemble* heuristic performs best.

Heuristic	Consensus	Decision
LDIS (least distance)	0.619	0.733
MAJ (majority voting)	0.576	0.719
AVG (average value)	0.617	0.702
RAN (random selection)	0.167	0.188
MDB (median based)	0.618	0.732
ESB (ensemble)	0.629	**0.739**
SDB (sd based)	**0.636**	0.722

Table 4: Comparison of heuristics of the *Consensus* and the *Decision* phase (best results are marked bold).

6. CONCLUSION AND FUTURE WORK

The heuristics discussed in this paper are basic group recommendation heuristics [9]. Our initial analysis shows the applicability of these heuristics in terms of prediction quality. As part of our future work we want to investigate the positive effect of recommendations on the dimensions *decision diversity*, *satisfaction with requirements prioritization*, and the *software quality* resulting from the requirements prioritization process. In addition, we want to conduct an in-depth evaluation of the user acceptance of the determined group recommendations - up to now only the majority voting based recommendations has been analysed [4].

Furthermore, we want to analyse whether if the *Standard Deviation* heuristic (as described in Formula 7) has the same prediction quality on other data sets . We also want to use this heuristic for group recommendation in our upcoming IntelliReq user studies. An important task for future work is the analysis of preference reversal [2] in requirements engineering and the impact on the prioritization quality and the stakeholders' satisfaction with the prioritization. In this context we want to develop new methods to improve stakeholder satisfaction.

Another topic of interest is the different interpretation of preferences. Are the group members arguing of the same topic? Does everybody has the same understanding of a given term? It has been shown that team members often enter a decision process from different viewpoints. Therefore it is necessary to find a consensus on the interpretation of shared information. This is especially important as the interpretation of issues has an massive impact on the decision making and is therefore considered as crucial [10].

[1]Due to the limited space only the results without differentiation between visible and non-visible preferences are shown.

7. REFERENCES

[1] R. B. Cialdini. The science of persuasion. *Scientific American*, pages 76–81, Feb. 2001.

[2] J. R. Curhan, M. A. Neale, and L. Ross. Dynamic valuation: Preference changes in the context of face-to-face negotiation. *Journal of Experimental Social Psychology*, 40(2):142 – 151, 2004.

[3] A. Felfernig, G. Friedrich, B. Gula, M. Hitz, T. Kruggel, G. Leitner, R. Melcher, D. Riepan, S. Strauss, E. Teppan, and O. Vitouch. Persuasive recommendation: serial position effects in knowledge-based recommender systems. In *Proceedings of the 2nd international conference on Persuasive technology*, PERSUASIVE'07, pages 283–294, Berlin, Heidelberg, 2007. Springer-Verlag.

[4] A. Felfernig, C. Zehentner, G. Ninaus, H. Grabner, W. Maalej, D. Pagano, L. Weninger, and F. Reinfrank. Group decision support for requirements negotiation. In L. Ardissono and T. Kuflik, editors, *Advances in User Modeling*, volume 7138 of *Lecture Notes in Computer Science*, pages 105–116. Springer Berlin / Heidelberg, 2012.

[5] L. Festinger. *A Theory Of Cognitive Dissonance*. Stanford University Press, 1957.

[6] P. Grünbacher. Collaborative requirements negotiation with easywinwin. In *Proceedings of the 11th International Workshop on Database and Expert Systems Applications*, DEXA '00, pages 954–960, Washington, DC, USA, 2000. IEEE Computer Society.

[7] H. Hofmann and F. Lehner. Requirements engineering as a success factor in software projects. *IEEE Software*, 18(4):58–66, Jul/Aug 2001.

[8] E. A. Lind, L. Kray, and L. Thompson. Primacy effects in justice judgments: Testing predictions from fairness heuristic theory. *Organizational Behavior and Human Decision Processes*, 85(2):189–210, July 2001.

[9] J. Masthoff. Group recommender systems: Combining individual models. In *Recommender Systems Handbook*, pages 677–702. 2011.

[10] S. Mohammed and E. Ringseis. Cognitive diversity and consensus in group decision making: The role of inputs, processes, and outcomes. *Organizational Behavior and Human Decision Processes*, 85(2):310 – 335, 2001.

[11] E. Molin, H. Oppewal, and H. Timmermans. Modeling group preferences using a decompositional preference approach. *Group Decision and Negotiation*, 6:339–350, 1997. 10.1023/A:1008612828114.

[12] D. Schkade, C. R. Sunstein, and D. Kahneman. Deliberating about Dollars: The Severity Shift. *Ssrn Electronic Journal*, 2000.

[13] K. Wiegers. First Things First: Prioritizing Requirements. *Software Development Online*, 7:48–53, Sept. 1999.

[14] J. A. Zuber, H. W. Crott, and J. Werner. Choice shift and group polarization: An analysis of the status of arguments and social decision schemes. *Journal of Personality and Social Psychology*, 62(1):50 – 61, 1992.