

An approach for explaining group recommendations based on negotiation information

Christian Villavicencio, Silvia Schiaffino, Jorge Andrés Díaz-Pace, Ariel Monteserin

ISISTAN (CONICET-UNCPBA), Campus Universitario, Tandil, Argentina

Article Info

Article history:

Received Mar 4, 2022

Revised Jul 5, 2022

Accepted Jul 8, 2022

Keywords:

Explanations

Group recommender systems

Negotiation-based recommendations

Negotiation Information

Transparency

ABSTRACT

Explaining group recommendations has gained importance over the last years. Although the topic of recommendation explanation has received attention in the context of single-user recommendations, only a few group recommender systems (GRS) currently provide explanations for their group recommendations. However, those GRS that support explanations, provide either explanations being highly reliant on the aggregation technique used for generating the recommendation (most of them trying to tackle shortcomings of the underlying technique), or explanations with a rich content but requiring users to provide considerable additional data. In this article, we present a novel approach for providing explanations of group recommendations, which are generated by a GRS based on multi-agent negotiation techniques. An evaluation of our approach with a user study in the movies domain has shown promising results. Explanations provided by our GRS system helped users during the decision-making process, since they modified the feedback given to recommended items. This is an improvement with respect to systems that do not provide explanations for their recommendations.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Ariel Monteserin

ISISTAN (CONICET-UNCPBA), Campus Universitario

Paraje Arroyo Seco, Tandil, Buenos Aires, Argentina

Email: ariel.monteserin@isistan.unicen.edu.ar

1. INTRODUCTION

Nowadays many items tend to be consumed by groups of users rather than by single users (i.e. movies, restaurants and touristic places, among others). As a result, the generation of group recommendations has become a practical need and also a promising research area. Although various techniques have been developed for making recommendations to a group as a whole, only a few of them [1], [2] have targeted the problem of explaining why certain items are being recommended. By explaining the decisions made by a recommender system, it is expected that users will better understand and trust the recommendations.

From the perspective of group recommendation generation, most works have used aggregation techniques to combine preferences, recommendations or user profiles [3], [4]. Thus, the explanations provided by these works are normally tied to the workings of such techniques and often fail at satisfying all group members evenly with their recommendations. Some of the problems of aggregation techniques include: i) they can produce values that might not represent correctly the data being aggregated, especially when the data are small and have a high variance and ii) the decision-making process of the group and the group dynamics [5] are not reflected by the aggregation techniques [6], [7]. To deal with these shortcomings, the multi-agent group recommender system (MAGReS) approach has been recently proposed [8]. MAGReS relies on a multi-agent system

in which the agents use negotiation techniques for producing group recommendations. This approach basically consists of a set of personal agents, each one representing a user in a group, which are involved in a negotiation process in order to determine which item should be recommended to satisfy users evenly. The items (proposals) under negotiation come from items being recommended by a single user recommender system (SUR), which is allocated at each agent and addresses the needs of the corresponding group member.

From the perspective of recommendation explanation, in turn, only a handful amount of group recommender systems (GRS) provide explanations for their recommendations [9]–[10]. Those GRS providing explanations often require either none [2], [11], [12] or considerable additional data [1], [13]. In the first case, the explanations are centered on how satisfied the GRS believes the group will be, and the explanations are generally a result of the aggregation technique used to generate the recommendations or a justification of its shortcomings. For example, an explanation might state that “the rating for movie X is 2.33, because the ratings provided by group members are 1, 1, and 5”. In the second case, the explanations are centered on analyzing additional information, which often must be provided by users, such as the personality of the group members and friendship relationships among them. For example, “Although we have detected that your preference for this item is not very high, your close friend X (who you highly trust) thinks it is a very good choice”. To address these issues and provide explanations that are not tied to aggregation techniques, we propose an approach that is able to generate explanations based on the recommendations negotiated by MAGReS. A key aspect of these explanations is that they are centered not only in data provided by the SUR but also in data captured from negotiation process carried out by MAGReS. The aim of these explanations is to provide transparency and increase users’ trust in the GRS. To do so, we define an approach that can be integrated to the MAGReS architecture, and extend it with explanation skills that rely on information from both the user profiles and the negotiation process.

In this article, we focus on the different types of explanations that the GRS can generate and how to obtain the information required by the explanations. To evaluate our approach for group explanations, we carried out a series of experiments with real users in the movies domain. First, we performed an experiment with an initial set of explanations that showed promising results but also highlighted aspects of the approach that should be improved. The lessons learned from this experiment helped us enhance the explanation capabilities. Then, we conducted a second experiment with an improved set of explanations, in which users had to evaluate different styles and types of explanations. Some findings of this experiment indicate that users preferred hybrid explanations (combining graphs and text), they considered the different types of explanations useful since they help them to have a better understanding about why items were recommended, and they could learn about the affinity with other users. In addition, half of the users would change their feedback after seeing the explanations, which can be seen as an improvement with respect to systems that do not provide explanations for their recommendations.

The rest of the article is organized as follows. In section 2 we discuss background concepts about explanations and group recommendations, and then we analyze related works. In section 3 we describe our proposed approach for explaining group recommendations, based on the MAGReS architecture. In section 4 we describe the experiments carried out with subjects being exposed to different types of recommendations, and then discuss the main results and lessons learned. Finally, in section 5 we give the conclusions, outline current limitations of the approach, and discuss future work.

2. BACKGROUND AND RELATED WORK

To explain the decisions that an intelligent system take is key to improve the adoption of such decisions by users [14]. Particularly, explanations of recommendations are important because recommender systems are often affected by two problems. The first one is that many recommenders work as “black boxes”, and given a recommendation, the user can trust or doubt it since she/he does not know in what way the system obtained that recommendation [15]. The second one is that recommenders are stochastic processes (since they vary over time) and, therefore, can produce errors no matter how well they have been implemented. In this context, explanations can provide some transparency, giving details of the reasoning and data behind the recommendation [16], [17].

The information provided by an explanation can be of many types, depending on its purpose, which is not always to convince the group members about accepting the recommendation given. That is, the role of explanations is necessary to justify why an item was recommended. According to [15], [18], the purpose

of explaining the recommendations is to help increasing confidence, improving the performance of decision-making and, above all things, granting transparency to users. The explanation can help users to understand the system recommendation process as well as to know its virtues and shortcomings. According to [19] the explanations can serve multiple purposes. For example, if the goal is to expose the reasoning and the data behind the recommendation, then the purpose of the explanations is to increase transparency. If the goal is to explain the user why she/he would want a certain item or not, the explanation contributes to improving the effectiveness of the recommendations. In addition to these two motivations, as explained in [19], [20], other possible objectives that an explanation can pursue include: trust (increase the user's trust in the system), efficiency (help the user to make decisions more quickly), persuasiveness (convince the user to accept the suggestion: to try something and buy something), satisfaction (increase the user's level of enjoyment when using the system and ease of use), and scrutability (if the user is allowed to inform the system about an error in the prediction made).

In addition to the motivations that make the explanation of recommendations necessary, the explanation styles to be used must be taken into account. The explanation style refers to the way in which the explanation is presented to the user and the information it contains. In the literature it is possible to find taxonomies that classify explanations both with respect to their content and the way they present the information, as well as regarding the technique being used to generate the recommendations. For example, in [21] a taxonomy is proposed that classifies the recommendation systems according to the given explanation style: human style, item style, feature style, and hybrid styles, i.e. a combination of some of the previous styles. If we consider the underlying algorithm, recommendation systems are classified in [19], [22], according to the style of explanation they provide, which according to the authors, it is somewhat dependent on the type of algorithm used by the recommendation engine. These styles can be based on: cases, collaborative filtering, based on content, conversational recommendation systems (ACORN [23]), demographic information (INTRIGUE [1]), or knowledge/utility (Qwikshop [24]).

When working with a GRS, explanations become even more important. Given the many ways in which a recommendation for a group can be obtained, and the conflicts of interest that may exist within a group, it is natural to think that group members would like to understand how they came to the recommendation they obtained and, in particular, how attractive the recommended items are for the other members of the group [25]. Currently there are only a few GRSs that provide explanations for their recommendations. While some works [2] focus on explaining the process for generating the group recommendations, others focus on explaining the recommendations generated [1]–[13]. In the rest of this section, we consider the second category of approaches, since from the perspective of group members this kind of explanations is easy to understand, because they do not need to know about the techniques used to generate the group recommendation.

In [1], the INTRIGUE GRS is proposed, which is able to recommend travel destinations to individuals and groups. INTRIGUE also provides explanations for its recommendations as a way to increase the users' trust in the recommendations. Particularly, the explanations pursue two objectives: i) to justify why an item is being recommended to the group and ii) to inform the group members about the existence of possible conflicts of interest. In [11], the authors propose a GRS that provides explanations for its recommendation. The explanations are built to inform the user about: i) the aggregation technique used when generating the group recommendation; ii) the estimated group rating for the recommended item; and iii) the preferences of the group members that were prioritized when computing the group ratings for the recommended items. POLYLENS [12] is an GRS that generates group recommendations in the movies domain by using a recommendation aggregation approach). To increase the transparency regarding the decision made, POLYLENS provides simple explanations that show the predicted individual (group members) and group ratings for the movies recommended. At last, in [13] the authors proposed an approach called Make-It -Personal. Differently from other expert systems, this approach not only generates explanations about the group recommendation but also about the social reality of the group for whom the recommendation was generated. Make-It-personal is able to produce two types of explanations: textual social explanations (TSE) and graphical social explanations (GSE).

In general, we can see that the works above only focus on informing how satisfied the GRS considers that the group will be with its recommendation. This is due to the lack of data that could be used to create better explanations. Existing approaches use aggregation techniques and thus they cannot explain much. Other approaches, like [13], provide richer explanations but require the users to provide additional (personal) data (e.g. with regard to their social relationships). We argue that a common problem shared by all the approaches is that, due to their reliance on aggregation techniques when generating the group recommendation, they tend

to use their explanations as a way of convincing unsatisfied group members to accept the recommendations so that other group member(s) will then be happy. In this context, the main difference between our proposal and these approaches is that our explanations are entirely based on: i) information related to the rating predictions made by a SUR and ii) in data extracted from a negotiation process. Additionally, the explanations provided do not only aim to inform the users about how satisfied the system believes they will be with the recommendation (like most of the analyzed approaches do), but also aim to identify and inform possible conflicts of interest among the group members.

3. PROPOSED APPROACH

We depart from the assumption of a GRS that works as a multi-agent system (MAS). In this MAS, each agent acts on behalf of a group member and maintains a profile with the user's preferences. An agent is capable of: i) predicting the rating the user would assign to an item not yet rated and ii) generating a ranking of "interesting items" for the user (items the user would like). To do so, each agent internally delegates some recommendations functions on a SUR. Initially, the user's preferences are the ratings assigned by the user to the items she rated in the past. Furthermore, these agents can engage in a negotiation process to try to reach a consensus on the most satisfying items for the group of users they represent. This negotiation process is multi-lateral and single-issue [26]. An instance of a MAS-based GRS using negotiation is the MAGReS approach [8], although other implementations are also possible. The MAGReS architecture is schematized in Figure 1.

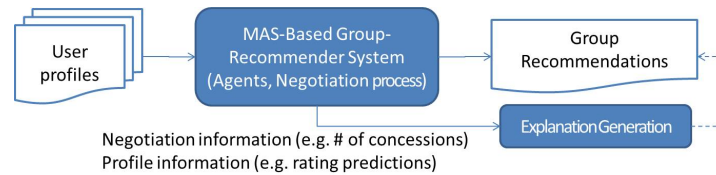


Figure 1. Group recommendation approach plus explanations

More formally, let $A = \{ag_1, ag_2, \dots, ag_n\}$ be a finite set of N cooperative agents, and let $X = \{x_1, x_2, \dots, x_m\}$ be a finite set of potential agreements or proposals, each one of them containing an item that can be recommended to one of the agents. Each agent $ag_i \in A$ has a *utility function* $U_i : X \rightarrow [0, 1]$ that maps proposals to its satisfaction value. This utility function captures the preferences of the user represented by the agent for different items. Each agent internally relies on a SUR to generate a ranking containing the items (candidate proposals) that the agent can propose to the group. The ranking is sorted in descendant order according to the utility value of the item. This way, the set X can be seen as the union of the rankings produced for all the agents, plus a special agreement called conflict deal, which yields utility 0 for all the agents and will be chosen as the worst possible outcome (no agreement is possible). The negotiation proceeds in rounds, in which each agent makes a proposal (i.e., a given item) to the other agents. In case all the agents agree on the proposal, in terms of their respective utilities, the negotiation ends with an agreement, which is interpreted as a group recommendation. There is an agreement when one agent makes a proposal that is at least as good (regarding utility) for any other agent as their own current proposals. In such a case, the other agents will accept the proposal. In case no agreement exists, one (or more) agents should make a concession. A concession means that an agent seeks an inferior proposal (in terms of utility), with the hope of reaching an agreement. When none of the agents can make concessions, the negotiation ends with conflict.

Conceptually, the MAS is seen as a GRS that takes a set of user profiles as inputs and generates a list of recommended items as output. Each item in this list is generated by a negotiation process, and it is intended to satisfy the group of users as a whole. However, these items carry no explanations for the users of the group. The explanation generation (EG) approach proposed in this work complements the MAS by allowing users to know how the recommendations were decided by the GRS. It should be noticed that MAGReS is a particular instance of a GRS. To provide explanations, EG processes information collected by the GRS with regards to: i) the negotiation process and ii) the rating predictions made by the SUR systems. We refer to all this information as source data. From this information, the EG approach is capable of generating two types of explanations, which we believe contribute to the understanding and transparency of the GRS. The first explanation type is related to the information extracted from the negotiation process. Within this type, we can distinguish two

sub-types, namely: affinity explanation and explanations with respect to the amount of items recommended. The second type is based on the information extracted from the SUR. Within the second type, we can also distinguish two sub-types, namely: satisfaction explanations and explanations with respect to the group members' profiles, which are also related to the negotiation process.

The mechanism for generating the explanations is based on predefined templates (one per explanation type) that might include textual and graphical parts. The textual part contains placeholders, which can be substituted for actual values (e.g., average user satisfaction) or user information (e.g., user name). There are different metrics to evaluate text generation. Particularly, the quality of the free-text explanations can be evaluated in terms of readability, based on frequently used readability measures, such as: Gunning Fog Index, Flesch Reading Ease, Flesch Kincaid Grade Level, Automated Readability Index, and Smog Index [27]. We ran different tests on our templates and obtained satisfactory values of the measures. For example, the average value for the Flesch Kincaid [28] test was 80 (being 100 the maximum). Each type of explanation is detailed in the following subsections.

3.1. Affinity explanations

This type of explanations seeks to inform the group members about how compatible their interests are, so as to help exposing conflicts of interest. To generate these recommendations, we take information from the negotiation process carried out by the agents in the MAS. Particularly, we build the explanation with the following information: the amount of concessions made by each agent, the amount of proposals made by each agent, and the amount of proposals that each agent accepted to another agent. Notice that we do not need to access to private information of users to build this type of explanation. The affinity explanations are generated with two levels of granularity: user-user and group. At user-user granularity level, the explanations aim to inform what the affinity levels between each pair of user are. This kind of explanation is built considering how the agent that represents each group member interacts during the negotiation process. Particularly, given two users u_1 and u_2 , and two agents ag_1 and ag_2 (where ag_i represents u_i), we analyze the proportion of the proposals uttered by ag_1 that were accepted or rejected by ag_2 , and vice versa. The affinity between two users is defined by $Aff_{user-user}(u_i, u_j) = |p|p \in P_{u_j} \wedge u_i \text{ accepted } p| / |P_{u_j}|$, where P_{u_j} is the set of proposals uttered by the agent ag_j during the negotiation.

For example, if agent ag_1 accepts 9 of 10 proposals uttered by ag_2 , $Aff_{user-user}(u_1, u_2)$ will be 0.9, but if agent ag_2 accepts 2 of 6 proposals uttered by ag_1 , $Aff_{user-user}(u_2, u_1)$ will be 0.33. These values of affinity indicate that the preferences of u_1 are closely related to the preferences of u_2 , but not vice-versa. To materialize this information in an explanation, we use text templates according to the affinity level. We empirically define three levels of affinity: low, medium and high based on two thresholds $th_{affLow} = 0.5$ and $th_{affHigh} = 0.8$. Then, the *affinity* level of u_1 with respect to u_2 is *high*, but the *affinity* level of u_2 with respect to u_1 is *low* ($0.33 < th_{affLow}$). Table 1 shows the templates to generate affinity explanations according to the affinity level.

Table 1. Templates for affinity explanations

Type	Template	Level
user-user	"From what we saw in {userName1} preferences profile, it seems he/she does not like the same kind of movies {userName2} likes."	low
	"According to what we saw in {userName1} preferences profile, he/she may like some of the movies {userName2} likes."	medium
	"It seems that the users {userName1} and {userName2} may have some interests in common."	medium
	"It seems that {userName1} likes the same kind of movies {userName2} likes."	high
	"It seems that the users {userName1} and {userName2} have many interests in common."	high
group	"According to what we saw when generating the recommendation, it seems that the user {userName} does not like many of the movies the rest of the group members like."	conflicting
	"It seems that {userName} would like almost all the movies liked by the other members of the group."	emphatic

At group granularity level, we aim to inform that there are conflicting or emphatic members in the group. A conflicting member is a user who has rejected most of the proposals received during the negotiation. In contrast, an emphatic member is a user who has accepted most of the proposals received during the negotiation. To determine this, we compute the proportion between the proposals accepted by u_i and the total number of proposals uttered during a negotiation: $Aff_{group}(u_i) = |p|p \in P \wedge u_i \text{ accepted } p| / |P|$, where P is the set of proposals. Moreover, we define empirically two thresholds $th_{conflicting} = 0.2$ and $th_{emphatic} = 0.8$.

Thus, if $Aff_{group}(u_i) < th_{conflicting}$ the user is considered as a *conflicting* member, and if $Aff_{group}(u_i) > th_{emphatic}$ the user is considered as a *emphatic* member. In Table 1, we can also see the templates to generate affinity explanations at group granularity level.

Moreover, we define an additional explanation within the scope of a single negotiated item. Figure 2(a) shows an example of this explanation. Here, the users can observe information about the number of times that each user concede and information about how many proposals are accepted or rejected by each group member.

3.2. Explanations with respect to the amount of items recommended

This type of explanation is generated when the GRS is not able to generate a recommendation containing the requested amount of items. This might occur if the interests of the group members are so conflicting that the agents are not able to reach an agreement. In consequence, the negotiation process ends with a conflict. The template used to communicate this explanation is "It seems that the interest of the group members $\{severityLevel\}$. Because of this, it was $\{difficultyLevel\}$ to find recommendations suitable for the group and so we could only produce $\{recsGenCounts\}$ recommendation (requested $\{recsExpectedCount\}$)". In this template, the variable *severityLevel* indicates the severity of the conflicts occurred during the negotiation. The severity level is computed taking into account the number of recommendation generated by the negotiation process (*recsGenCounts*) and the number of recommendations requested by the group (*recsExpectedCount*): $severity = recsGenCounts / recsExpectedCount$. We define three levels of severity: *high* if $0 \leq severity < 0.33$; *medium* if $0.33 \leq severity < 0.66$; and *low* if $0.66 \leq severity < 1$. Notice that if *recsGenCounts* is equals to *recsExpectedCount* this kind of explanation is not generated.

3.3. Satisfaction explanations

To complement the explanations generated from the negotiation information, we define well-known explanations inspired by the approaches analyzed in section 2. They aim to increase the transparency of the recommendation process, by clarifying the reason why the items were recommended to the group. They also aim to persuade the group members to consider items that they might not know but might like. The satisfaction explanations are also generated with two different levels of scope (item or recommendation) and granularity (per-user or group). Table 2 shows the templates used to generate the explanations. Particularly, when the scope is item, the satisfaction explanations can be presented using three different styles, namely: textual, graphical and hybrid (it combines text and graphics, as depicted in Figure 2(b).



(a)



(b)

Figure 2. Examples of explanations with scope single item: (a) affinity explanation and (b) satisfaction explanation (Granularity: per-user and group. Style: hybrid)

Table 2. Templates for satisfaction explanations

Scope	Granularity	Template
Item	User	<p>"We believe that {userName} will be {satisfactionLevel} with this movie. In fact, we think that {userName} would rate this movie {starsQuantifier} {starsCount} stars {starsQuantifierAproxAfter}."</p> <p>"In our opinion, {userName} will be {satisfactionLevel} with this movie, as according to our estimations {userName} would rate this movie {starsQuantifier} {starsCount} stars {starsQuantifierAproxAfter}."</p> <p>"Given that we estimate that {userName} would rate this movie {starsQuantifier} {starsCount} stars {starsQuantifierAproxAfter}, we believe that {userName} will be {satisfactionLevel} with this movie."</p> <p>"It was impossible to determine whether {userName} will be satisfied with this item or not. We have assumed that he/she will not. He/She may not have rated enough items."</p>
	Group	<p>"We believe that the group will be {satisfactionLevel} with this movie. In our opinion, the group would rate this movie {starsQuantifier} {starsCount} stars {starsQuantifierAproxAfter}."</p> <p>"In our opinion, the group will be (satisfactionLevel) with this movie. In fact, we think that the group would rate this movie {starsQuantifier} (starsCount) stars (starsQuantifierAproxAfter)."</p> <p>"With regard to this item, we think that the group would rate it {starsQuantifier} {starsCount} stars {starsQuantifierAproxAfter}. Thus, in our opinion the group will be (satisfactionLevel) with this movie."</p>
Recommendation	User	<p>"We believe that {userName} will be {satisfactionLevel} with this recommendation."</p> <p>"In our opinion, {userName} will be {satisfactionLevel} with this recommendation."</p> <p>"We think that {userName} will be {satisfactionLevel} with the movies recommended."</p> <p>"It was impossible to determine whether {userName} will be satisfied with this recommendation or not. We have assumed that he/she will not. For more information, head to the "Preferences Profiles" Section."</p>
	Group	<p>"We believe that the group will be (satisfactionLevel) with this recommendation."</p> <p>"In our opinion, the group will be {satisfactionLevel} with the movies recommended."</p> <p>"Overall, we believe that the group will be {satisfactionLevel} with this recommendation."</p> <p>"According to our rating estimations, we think that the group will be {satisfactionLevel} with the movies we recommended."</p>

3.4. Explanations with respect to the group members' profiles

Finally, our approach generates explanations to inform the group about those group members for whom their preference profiles do not have enough information for the recommender to make predictions. This is important because the recommender cannot predict the preferences of the users, and this aspect has a strong impact on the negotiation process. In these situations, the agents do not have proposals to utter and, consequently, they have to accept any proposal received from the other group members.

4. EVALUATION

We carried out a series of experiments with subjects to evaluate the different types of explanations generated by the EG approach in the movies domain. The experiments had the following goals: i) determine whether the explanations provided are useful for users and ii) considering the information provided in the explanations, assess whether users would consider changing the feedback given to the recommended items. Note that the second goal is more ambitious than the first one, as it would indicate that the GRS can persuade the user about item decisions. We used a reduced version of the MovieLens latest dataset, provided by GroupLens [29], which contains: 15.816 users, 3.257 movies and 780.327 ratings. The experiments were performed using a web application in order to provide explanations for the recommendations.

4.1. Initial evaluation

Initially, we developed a prototype of the approach and conducted a first evaluation with a small set of users, in order to test the different types of explanations and get feedback from the users. The lessons learned during this experiment drove the definition of the explanations described in section 3. During the first semester of 2018, we conducted an experiment with 34 participants using a recommendation application for movies. These participants were students of computer science engineering in their fourth year of the career, Ph.D. students in computer science and researchers, between 20 and 40 years old. As group recommendation systems, we used both a traditional preference aggregation approach and a multi-agent negotiation approach. The experiment comprised six stages, namely:

- Login in to the application and populate the individual profiles by rating movies: the participants were asked to rate at least 15 movies, with 0 to 5 stars. The participants could choose among 300 movies to be rated. These movies were selected considering that: a) the genres of the movies were diverse; and b) they were released after the date of birth of the participants (most of them were born during the 80's) or they were famous movies).
- Group formation: we formed groups of 3 people randomly.
- Group recommendation generation: each group was asked for 10 recommendations.
- Evaluation of recommendations (without looking at explanations): once the recommendation is generated, users can provide feedback for each item recommended. This feedback can be provided individually and group-wise, and it is expressed as a rating (1 to 5 stars).
- Evaluation of explanations: the recommender system provides explanations for the items recommended. Users have to review the explanations generated and determine whether they would change the feedback provided in *stage 5*.
- Users were asked to answer a questionnaire regarding the explanations.

The experiment involved 18 groups, 15 of these group had 3 members and 3 groups had 2 members. The types of explanations that were available in the Web application were regarding satisfaction and flexibility. Regarding satisfaction, for each item recommended to the group, the system showed information about the estimated group satisfaction and the individual satisfaction of each group member. This type of explanation only included the graphical style of the explanation when the scope was an item as shown in section 3.3. Regarding flexibility, the system gives information about the percentage of concessions made by each member of the group, and the amount of proposals accepted and rejected by each group member. This was also a subset of the Affinity explanations defined in section 3.1. Particularly, this type of explanation included the explanation within the scope of a single negotiated item.

The questionnaire that the users had to answer consisted of the following questions:

- Do you consider that the explanations provided are useful? Why?
- Do you consider that the explanations provided enable you to understand how the system works?
- Would you change the feedback provided once you have seen the explanations?
- What is your opinion about the format of the explanations?

After analyzing the answers given by the participants with respect to the usefulness and format of the explanations, we drew a number of observations that were considered to improve the EG approach. Half of the users considered that the explanations provided were useful because, in this way, they could know which users were satisfied and which ones were not. The other half indicated that they were only interested in getting a recommendation, without any need of explanations. All users agreed in that the explanations enabled them to understand better how the GRS works. However, half of them argued that as users, they were not interested in knowing how the system works. Half of the users answered that they would not change the feedback provided after seeing the explanations. The other half indicated that they would change the feedback after seeing the explanations, particularly for unseen movies. Regarding the format of explanations, half of the users agreed with the information provided. Others indicated they would like to have access to the ratings provided by the other group members. Finally, some users would prefer an explanation provided in natural language rather than graphic format. Taking into account these lessons learned, we refined the different types of explanations obtaining those described in section 3.

4.2. Evaluation with improved explanations

The participants of the experiments were 48 students of computer science engineering in their last years of the career and Ph.D. students in computer science. Their ages ranged between 23 to 42 years. The participants finally formed 32 groups of 3 members since each participant was part of at least two groups. The participants were required to perform a 7-stage process, as described in section 4.1.

The evaluation was carried out in two parts, each one covering a subset of the questions [30]. Figure 3 shows the results of the most relevant questions. The first part covered questions *Q1* to *Q7* and was centered on the item-level explanations, which are exclusively satisfaction explanations (with *scope* set to item). For the experiment, we allowed the participants to see all three styles (textual, graphical and hybrid) of recommendations, and asked them which one of the styles they preferred. The favourite style was the *hybrid* one (76%), since it allowed to see at a first glance how good the GRS thought the recommendation will be for them, but it also provided a few additional details in form of text. The participants also stated that the *textual* style was a bit repetitive, which was expected due to the low number of template options available at the time of the experiment. In consequence, only 2% of the participants preferred textual explanations, and 22% of the participants preferred only graphic explanations. In response to *Q3*, 53% of the users said that the explanations were accurate. We later analyzed why the remaining 47% considered the explanations inaccurate. We think that one reason could be the low quality of the predictions made by the individual recommender system (SUR) used in the experiments.

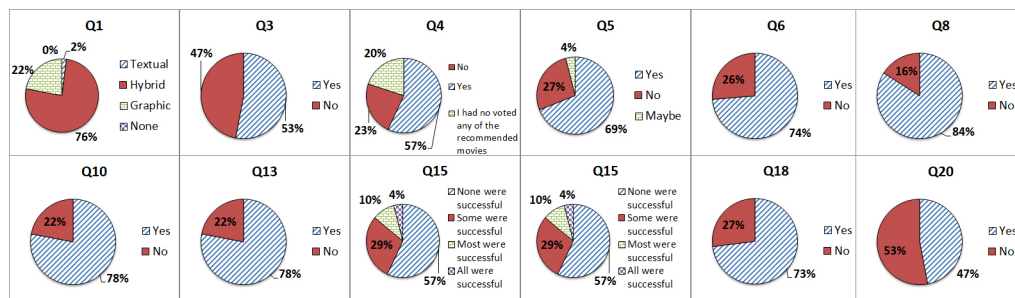


Figure 3. Results of most relevant questions

Question *Q4* showed that 71.8% of the participants confirmed that the explanations matched with their opinion about a movie that they had already watched. The difference between the percentage of *Q3* and *Q4* is due to the fact that *Q4* only took into account the movies already watched by the users. For this reason, we consider that explanations failed to explain movies that users did not know due to the fact that users could be not motivated to see a movie just by reading the title and the textual description. In addition, question *Q5* showed that 69.4% of the participants would accept to see a movie recommended taking into account the information given by the explanations.

Finally, in response to *Q6* the majority of the participants (around 73% of them) indicated that they considered this kind of explanations useful. Some of the reasons stated were the following: i) they help them to have a better understanding about why the item was recommended; ii) they give them information about the interests of the group members; iii) they enable them to consider movies that they would not consider otherwise; and iv) they provide contextual knowledge. The users additionally stated that these explanations were specially important if they did not know the item recommended. Those users that did not found the explanations useful indicated that: they preferred an explanation saying that “we recommend movie X because you liked movie Y”; they wanted information about other members’ ratings; they were sometimes redundant.

The second part covered questions *Q8* to *Q20*, and was used to evaluate the explanations generated by our approach for the recommendations (as a whole). This part targeted the following types of explanations: satisfaction explanations (scope: recommendation, granularity: both); affinity explanations; explanations with respect to the amount of items recommended; and explanations with respect to the profiles of the group members. With respect to the satisfaction explanations (questions *Q8* to *Q11*), 83.7% of the users involved in the experiment stated that the satisfaction explanations with *group* granularity were very useful (*Q8*), especially because: i) the explanations helped them to have a preliminary idea of the quality of the recommendation;

ii) they gave information about the tastes of other users, which can be considered to decide whether or not to see a movie; and iii) if they did not know some of the movies recommended, the explanation encouraged them to ask for reviews of the movies. Related to this, the satisfaction explanations with per-user granularity were also very well received, as the 77.6% of the participants considered these explanations were useful (Q10).

Regarding to the explanations extracted directly from the negotiation information, question Q13 showed that 78.4% of the participants considered the group affinity explanations as accurate. Moreover, 95.9% of the users indicated that at least some of the user-user affinity explanations were accurate (Q15). In addition, 55.1% of the users preferred explanations generated from negotiation information to explanations generated using information from the users profiles (Q16). The remaining explanation types were considered as useful (Q18) by most of the participants (approximately 65% of them). This was expected because of these types of explanations are generated only when our approach must inform the group about an extreme situation that conditioned the operation of the multi-agent negotiation.

Finally, we asked the participants if, after seeing the explanations, they would decide to modify the feedback they had given to the items recommended (Q20): only 46.9% of the participants indicated that they indeed had modified the feedback given. Because of this aspect, and considering that most explanations were considered useful, we believe that it is necessary to adjust some explanation types, if we want to increase their effect on how the group perceives the recommendation.

5. CONCLUSION

In this article, we propose a novel approach that is capable of producing explanations of group recommendations based on negotiation information and without requiring the users to provide additional data. Our approach provides explanations for recommendations to groups that take advantage of the dynamics of the negotiation process. These explanations can i) help members from the group to understand the reason why each item was recommended; ii) persuade them so that they take into account recommendations that otherwise would have ignored; iii) inform them about possible affinities between their interests; and iv) notify them about particularities that occurred during the process of recommendation that could have affected the recommendations generated. To evaluate the proposed explanations, we performed experiments with human subjects in the movies domain. Taking into consideration the results obtained, we argue that the explanations provided by our GRS system are useful and can help users during the decision-making process (i.e. when they evaluate the recommendations and decide whether to accept them or not). The experimental results showed that a significant number of users considered changing their evaluation of the recommended movie based on the explanation offered by our approach. For this reason, one of our findings is that, considering that users could change the feedback given to recommended items, GRS that provide explanations have an advantage over those that do not provide them. In addition, users indicated that among the different types of explanations, they preferred those generated from negotiation information. One of the limitations of our proposal is linked to the evaluation process. Given the reduced number of subjects (with a similar career profile) that participated in our user study, the results obtained, although they are promising and interesting, cannot be directly extrapolated to a larger user population. As a future work, we plan to enhance the explanations provided taking into account the results obtained and mainly the information provided by users in the questionnaire. Particularly, we plan to adapt the explanations according to the different personality characteristics (e.g., the conflict resolution styles). Moreover, future work will aim to model i) different personalities and ii) relationships between the users in the negotiation strategy used by the agents in the GRS. We will also evaluate the proposed approach with a larger set of subjects.




REFERENCES

- [1] L. Ardissono, A. Goy, G. Petrone, M. Segnan, and P. Torasso, "Intrigue: Personalized recommendation of tourist attractions for desktop and hand held devices," *Applied Artificial Intelligence*, vol. 17, no. 8-9, pp. 687-714, 2003, doi: <https://doi.org/10.1080/713827254>.
- [2] W. Wang, G. Zhang, and J. Lu, "Hierarchy visualization for group recommender systems," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2017, doi: 10.1109/TSMC.2017.2699690.
- [3] R. B. Nozari and H. Koohi, "A novel group recommender system based on members influence and leader impact," *Knowledge-Based Systems*, vol. 205, p. 106296, 2020, doi: 10.1016/j.knosys.2020.106296.
- [4] Y.-D. Seo, Y.-G. Kim, E. Lee, K.-S. Seol, and D.-K. Baik, "An enhanced aggregation method considering deviations for a group recommendation," *Expert Systems with Applications*, vol. 93, pp. 299-312, 2018, doi: 10.1016/j.eswa.2017.10.027.





- [5] T. N. Nguyen and F. Ricci, "Dynamic elicitation of user preferences in a chat-based group recommender system," in *Proceedings of the Symposium on Applied Computing*, 2017, pp. 1685–1692, doi: 10.1145/3019612.3019764.
- [6] L. Baltrunas, T. Makcinskas, and F. Ricci, "Group recommendations with rank aggregation and collaborative filtering," in *Proceedings of the fourth ACM conference on Recommender systems*, 2010, pp. 119–126, doi: 10.1145/1864708.1864733.
- [7] J. Masthoff, "Group Recommender Systems: Combining Individual Models," *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. Kantor, Eds., Boston, MA, USA: Springer, 2011, pp. 677–702, doi: 10.1007/978-0-387-85820-3_21.
- [8] C. Villavicencio, S. Schiaffino, J. A. Diaz-Pace, and A. Monteserin, "Group recommender systems: A multi-agent solution," *Knowledge-Based Systems*, vol. 164, pp. 436–458, 2019, doi: 10.1016/j.knosys.2018.11.013.
- [9] Y. Zhang and X. Chen, "Explainable recommendation: A survey and new perspectives," *Foundations and Trends in Information Retrieval*, vol. 14, no. 1, pp. 1–101, 2020.
- [10] S. Najafian and N. Tintarev, "Generating consensus explanations for group recommendations: An exploratory study," in *Adjunct Publication of the 26th Conference on User Modeling Adaptation and Personalization*, 2018, pp. 245–250, doi: 10.1145/3213586.3225231.
- [11] E. Ntoutsi, K. Stefanidis, K. Nørnvåg, and H.-P. Kriegel, "Fast group recommendations by applying user clustering," in *Conceptual Modeling*, P. Atzeni, D. Cheung, and S. Ram, Eds., Berlin, Heidelberg: Springer, 2012, doi: 10.1007/978-3-642-34002-4_10.
- [12] M. O'Connor, D. Cosley, J. A. Konstan, and J. Riedl, "PolyLens: a recommender system for groups of users," in *Proceedings of the Seventh European conference on computer supported cooperative work*, Bonn, Germany: Springer, 2001, pp. 199–218, doi: 10.1007/0-306-48019-0_11.
- [13] L. Quijano-Sanchez, C. Sauer, J. A. Recio-Garcia, and B. Diaz-Agudo, "Make it personal: A social explanation system applied to group recommendations," *Expert Systems with Applications*, vol. 76, pp. 36–48, 2017, doi: 10.1016/j.eswa.2017.01.045.
- [14] S. Yoo and N. Kang, "Explainable artificial intelligence for manufacturing cost estimation and machining feature visualization," *Expert Systems with Applications*, vol. 183, p. 115430, 2021, doi: 10.1016/j.eswa.2021.115430.
- [15] J. L. Herlocker, J. A. Konstan, and J. Riedl, "Explaining collaborative filtering recommendations," in *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, 2000, pp. 241–250, doi: 10.1145/358916.358995.
- [16] N. Tintarev and J. Masthoff, *Designing and Evaluating Explanations for Recommender Systems*. Boston, MA, USA: Springer, 2011, pp. 479–510.
- [17] N. Liu, Y. Ge, L. Li, X. Hu, R. Chen, and S.-H. Choi, "Explainable recommender systems via resolving learning representations," in *Proceedings of the 29th ACM international conference on information & knowledge management*, 2020, pp. 895–904, doi: 10.1145/3340531.3411919.
- [18] J. L. Herlocker, "Position statement-explanations in recommender systems," in *CHI'99 Workshop, Interacting with Recommender Systems*, 1999.
- [19] N. Tintarev and J. Masthoff, *Explaining Recommendations: Design and Evaluation*. Boston, MA, USA: Springer, 2015, pp. 353–382.
- [20] A. Felfernig, L. Boratto, M. Stettinger, and M. Tkalcic, *Group Recommender Systems: An Introduction*, Boston, MA, USA: Springer, 2018.
- [21] A. Papadimitriou, P. Symeonidis, and Y. Manolopoulos, "A generalized taxonomy of explanations styles for traditional and social recommender systems," *Data Mining and Knowledge Discovery*, vol. 24, no. 3, 2012, doi: 10.1007/s10618-011-0215-0.
- [22] N. Tintarev and J. Masthoff, "A survey of explanations in recommender systems," in *Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop*, 2007, pp. 801–810, doi: 10.1109/ICDEW.2007.4401070.
- [23] P. Wärnestål, "User evaluation of a conversational recommender system," in *Proceedings of the 4th IJCAI Workshop on Knowledge and Reasoning in Practical Dialogue Systems*, 2005, pp. 32–39.
- [24] K. McCarthy, J. Reilly, L. McGinty, and B. Smyth, "Thinking positively - explanatory feedback for conversational recommender systems," in *Proceedings of the European Conference on Case-Based Reasoning (ECCBR-04) Explanation Workshop*, 2004, pp. 115–124.
- [25] A. Jameson and B. Smyth, "Recommendation to Groups," in *The adaptive web: methods and strategies of web personalization*, Berlin, Heidelberg: Springer, 2007, pp. 596–627.
- [26] U. Endriss, "Monotonic concession protocols for multilateral negotiation," in *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, 2006, pp. 392–399, doi: 10.1145/1160633.1160702.
- [27] M. Zamanian and P. Heydari, "Readability of texts: State of the art," *Theory & Practice in Language Studies*, vol. 2, no. 1, 2012, doi: 10.4304/tpls.2.1.43-53.
- [28] R. Flesch, "Flesch-kincaid readability test," *Retrieved October*, vol. 26, no. 3, 2007.
- [29] "Grouplens research," *GroupLens*, [Online]. Available: <https://grouplens.org/datasets/movielens/> (accessed May 15, 2022).
- [30] "Research questionnaire," [Online]. Available: <https://tinyurl.com/magres-explanations> (accessed Jun. 1, 2022)

BIOGRAPHIES OF AUTHORS







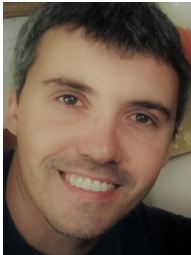
Christian Villavicencio    is a Software Developer (Globant Contractor) at Expedia Group. He received his Ph.D. in Computer Science in 2019 at UNCPBA. He can be contacted at email: christian.villavicencio@isistan.unicen.edu.ar.







Silvia Schiaffino     is a full-time professor in the Computer Science Department at UNCPBA (Tandil, Argentina), member of ISISTAN Research Institute and researcher at CONICET. She received her Magister degree in 2001 and her Ph.D. in 2004. Her main research interests are user profiling, personalization, recommender systems, and machine learning. She can be contacted at email: silvia.schiaffino@isistan.unicen.edu.ar.



Andres Diaz-Pace     is currently a professor at UNICEN University (Tandil, Buenos Aires, Argentina), and also a Principal Research Fellow of CONICET, Argentina. From 2007 to 2010, he was a member of the technical staff at the Software Engineering Institute (SEI, Pittsburgh, USA). His primary research interests are: quality-driven architecture design, AI and NLP techniques applied to software engineering, and architecture-based evolution and conformance. He can be contacted at email: andres.diazpace@isistan.unicen.edu.ar.



Ariel Monteserin     is a researcher at ISISTAN Research Institute at CONICET-UNICEN, Argentina. He is an associated professor in the Computer Science Department at Univ. Nac. del Centro de la Pcia. de Bs. As (UNICEN), Tandil, Argentina. His main interests are negotiation among intelligent agents, influence maximization algorithms, and smart cities. He received his Ph.D. on Computer Science in 2009. He can be contacted at email: ariel.monteserin@isistan.unicen.edu.ar.