A Consensus-Driven Group Recommender System

Jorge Castro,[1,*] Francisco J. Quesada,[2,†] Iván Palomares,[3,‡] Luis Martínez[2,§]
[1]Department of Computer Science and Artificial Intelligence, University of Granada, 18071, Granada, Spain
[2]Computer Science Department, University of Jaén, Jaén, Spain
[3]Built Environment Research Institute, University of Ulster, Londonderry BT52 1SA, United Kingdom

Recommender systems aim at filtering large amounts of information for users, providing them with those pieces of information which better meet their preferences or needs. Such systems have been traditionally used in diverse areas, such as e-commerce or tourism. Within this context, group recommender systems address the problem of generating recommendations for groups of users who might have different interests. Although different aggregation processes have been extensively utilized in real-life applications to generate group recommendations, such processes do not guarantee that the list of products recommended to the group reflect a high agreement level among its members' individual preferences. Given the need for considering the added value of obtaining group recommendations under a high agreement level, this paper presents a novel group recommender system methodology that attempts to reach a high level of consensus among individual recommendations of group members. To do this, and inspired by existing group decision-making approaches in the literature, a consensus reaching process is carried out to bring such individual recommendations closer to each other before delivering the group recommendations.

1. INTRODUCTION

In the contemporary context, there is an overwhelming amount of information that leads users into the difficult task of filtering information that meets their actual needs. To address this problem, recommender systems were proposed[1] to filter information, thus delivering to users only the information that meets their preferences or needs.

Traditional recommender systems address the problem of providing recommendations targeted to individual users, but there exist certain products or services,

*Author to whom all correspondence should be addressed; e-mail: jcastro@decsai.ugr.es
†e-mail: fqreal@ujaen.es.
‡e-mail: i.palomares-carrascosa@ulster.ac.uk.
§e-mail: martin@ujaen.es.

INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS, VOL. 00, 1–20 (2015)

View this article online at wileyonlinelibrary.com. • DOI 10.1002/int.21730

such as movies,[2] music,[3] and tourist points of interest,[4,5] that have certain social features, therefore they are meant to be enjoyed by a group of users instead of individually. In this context, traditional recommender systems were limited, for this reason it became necessary to extend such systems to overcome this limitation.

Group recommender systems (GRS)[6] are one of the most challenging, yet necessary, aspects of research in the field of recommender systems: the necessity of generating recommendations targeted to a group of users with individual interests that might be different from each other.[7] In group recommendations, as stated by Jameson and Smyth in Ref. 8, there exist four basic recommending subtasks: (i) acquiring member preferences, (ii) generating recommendations, (iii) explaining group recommendations, and (iv) aiding to make the final choice. In this paper, we focus on improving recommendations by applying techniques from group decision making (GDM) and consensus reaching. Regarding the process to generate group recommendations, two extensions of individual recommender systems have been proposed[9] : *rating aggregation* and *recommendation aggregation*.

(i) In rating aggregation, individual ratings are combined to obtain a group profile that represents the group preferences.
(ii) In recommendation aggregation, each member's individual recommendations are obtained, and these recommendation lists are aggregated to obtain a suitable recommendation list for the group.

The recommendation process in GRSs has been explored for both rating and recommendation aggregation. In this paper, our aim is to meet individual users' needs in a direct way; therefore, we focus on recommendation aggregation. A desirable feature of these predictions would be to minimize the misery of members, regarding their possible disagreement with the best recommended products. Thus the minimum operator has been used in some works for the recommendation aggregation process.[2] However, applying this aggregation process solely does not guarantee that there will exist a high level of agreement among the group users over the recommendations received, but rather a minimum level of agreement.

To overcome this situation, our goal consists not only in seeking a group recommendation that mainly suits a satisfied majority of users in the group, but also in giving current GRS the added value of reaching certain agreement level among users regarding such recommendations. To do this, we consider the use of consensus approaches for GDM, by integrating them into the group recommendation process. In a GDM problem, several individuals or experts attempt to find a common solution for a decision problem composed by a set of alternatives or possible solutions to such a problem.[10,11] Thus, each expert expresses his/her preferences on each alternative. Traditional selection processes for the resolution of GDM problems[12] do not regard the fact that some experts might disagree with the decision made; hence, a number of consensus-based approaches[13,14] have been proposed to overcome this limitation by applying consensus reaching processes (CRPs) to achieve a high level of agreement before making group decisions. In a CRP, experts iteratively bring their preferences closer to each other, until a sufficient level of agreement is reached among them.[15]

Based on the previous goals, in this paper we propose a methodology that, given the recommendations for a group of users, attempts to reach a high level of consensus on the recommendations provided to them. Straightaway, a GRS model that implements such a methodology to deliver agreed recommendations to the group is presented.

This paper is set out as follows: in Section, 2 some basic concepts and preliminaries on GRSs and CRPs for the resolution of GDM problems are reviewed. Section 3 presents the consensus-driven GRS for agreed recommendations proposed in this paper, describing in detail its different phases. Section 4 shows a case study to evaluate the proposed GRS technique and compare it with baseline techniques. Finally, some concluding remarks are pointed out in Section 5.

2. PRELIMINARIES

This section first reviews some basic concepts on GRSs, followed by a brief overview of CRP in GDM.

2.1. Group Recommender Systems

In this section, basic concepts in recommender systems (RS) and group recommendation are explained, describing the inputs and basic techniques for group recommendation.

Traditional RSs (single user) use three sources of information:

- Users: $U = \{u_1, \ldots, u_q\}$ is the set of users of the system, which may provide information about themselves such as age, gender, or zip code.
- Items: $I = \{i_1, \ldots, i_t\}$ is the set of items of the system, which may have content information such as metadata or textual description.
- Ratings: $R \subseteq U \times I \rightarrow D$ is the set of users' ratings over the products, to describe how satisfied a user is regarding a particular item in the rating domain D.

RSs attempt to rank and filter items, as well as predicting ratings for unseen items by the user, to perform recommendations using these data sources. Some examples of existing RS techniques include demographic recommendation,[16] content-based recommendation,[17,18] and collaborative filtering approaches,[19,20] which rely on users' data, items, or ratings, respectively. Formally, recommender systems try to recommend the item or set of items that maximizes a given utility function.

$$Recommendation(I, u) = \arg\max_{i_l \in I}[Utility(i_l, u)] \tag{1}$$

Content-based recommender systems use similarity metrics between user and item profiles as the utility function. On the other hand, in collaborative and demographic approaches, the utility function applied consists in rating prediction.

Collaborative filtering techniques rely on rating information, and the more information about all users' preferences is available, the better the performance. A simple and extended technique is collaborative filtering using nearest neighbors

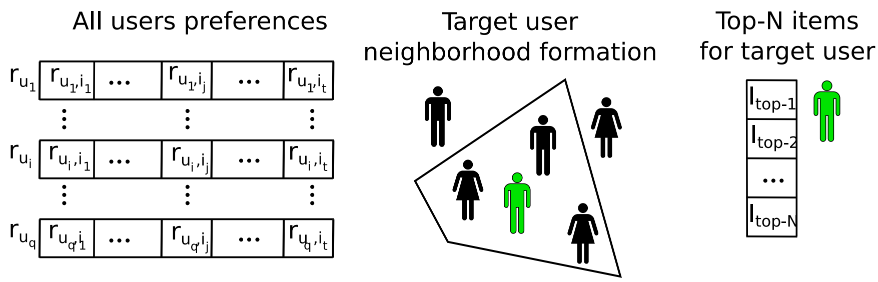

Figure 1. Recommendation aggregation scheme.

approach[21] (see Figure 1). In this technique, users' preferences are represented by a vector with user's ratings, which contains empty cells representing the items that the user has not experienced yet. The idea behind memory-based collaborative filtering is to find the k-most similar users (neighbors) to the target one, computing the similarity between user's ratings. The Pearson correlation coefficient has been proved to be the most suitable similarity measure between users, since it is not affected by the user's bias when rating items,[22] e.g., users who rate items on a consistently high or low basis. Once the neighborhood is selected, it is used to predict the rating for unseen items, by combining the neighborhood ratings on the item. A number of methods to combine neighbors ratings have been proposed.[23] For the sake of simplicity, in this paper, we consider the weighted similarity aggregation.[24] Finally, the recommendation list is constructed as an ordered list of unseen items or *predictions*, ranked by a decreasing order of their prediction value, i.e., an estimated rating value indicating how useful or interesting would the item be for the user.

Basic approaches for GRSs extends RSs so that, instead of recommending to a single user, recommendations are targeted to groups of users ($G = \{g_1, \ldots, g_m\} \subseteq U$). GRSs can operate in different modes, such as finding the most suitable group of users for a target item,[25] or recommending groups to a user for joining them,[26] In this proposal, we focus on item recommendation targeted to groups. Since we apply collaborative approaches, our utility function is rating prediction. Formally, group recommendation consists in finding the item (or set of items) that maximizes the rating prediction for the group of users:

$$GroupRecommendation(I, G) = \arg\max_{i_l \in I}[Prediction(i_l, G)] \tag{2}$$

Taking advantage of previous research on single user RSs, group recommendations are generated by extending them. Thus, there exist two basic approaches:[9]

- *Rating aggregation*,[3] which consists of aggregating individual ratings of each member to compute an aggregated group rating profile or *pseudouser* and perform individual recommendation using this profile as input.
- *Prediction aggregation*,[2,27,28] which aggregates each member's recommendations list into a list targeted to the group (as illustrated in Figure 2).

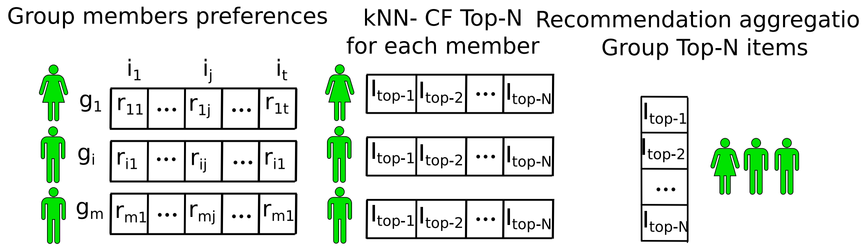

Figure 2. Recommendation aggregation scheme.

A known limitation of GRSs is that recommendations might not meet the individual preferences of all members in the group, e.g., some recommended items to the group might be regarded as satisfactory by some members, and unsatisfactory by several of them. To overcome this, recommendation aggregation minimizing member's misery[2] has been applied, but it does not take into account group dynamics such as group influence on individuals behavior. Our proposal tries to overcome this situation by applying CRPs in recommendation.

2.2. Consensus Reaching Processes in GDM

GDM problems imply the participation of multiple experts, with different knowledge and experience, who have to solve a decision problem making a common decision. A GDM problem is formally characterized by the following elements:[10,11]

- The existence of a decision problem to be solved.
- A set $X = \{x_1, ..., x_n\}(n \geq 2)$, of *alternatives* or possible solutions to the problem.
- A set $E = \{e_1, ..., e_m\}(m \geq 2)$, of participants or *experts*, who express their opinions or preferences over the set of alternatives X.

Decision problems may take place in different environments (certainty, risk, and uncertainty), being most real-life GDM problems usually defined in uncertainty environments. These environments are characterized by existence of vague and imprecise information. To express their opinions in uncertain contexts, experts may express their preferences in different information domains, such as numerical,[29] interval-valued,[30] or linguistic.[31,32]

Usually, experts utilize a preference structure to express their opinions over alternatives. One of the most used preference structures in GDM problems under uncertainty is the *fuzzy preference relation*.[33] A fuzzy preference relation P_i associated with expert e_i is defined by a membership function $\mu_{P_i} : X \times X \to [0, 1]$, and it is represented for X finite as an $n \times n$ matrix:

$$P_i = \begin{pmatrix} - & \cdots & p_i^{1n} \\ \vdots & \ddots & \vdots \\ p_i^{n1} & \cdots & - \end{pmatrix}$$

where each assessment $p_i^{lk} = \mu_{P_i}(x_l, x_k)$ represents the preference degree of the alternative x_l over x_k, $l, k \in \{1, ..., n\}$, $l \neq k$, according to e_i, interpreted as follows:

- $p_i^{lk} > 0.5$ indicates e_i's preference of x_l over x_k.
- $p_i^{lk} < 0.5$ indicates e_i's preference of x_k over x_l.
- $p_i^{lk} = 0.5$ indicates e_i's indifference between x_l over x_k.

Traditionally, the selection process for reaching a solution for a GDM problem only consists of two phases:[34]

1. *Aggregation*: Experts' preferences are combined by using an aggregation operator:[35]
2. *Exploitation*: A selection criterion[12,29] is applied to obtain an alternative or subset of alternatives as the solution to the problem.

The selection process does not guarantee that an agreement level is achieved when making the decision, which would be essential in several real-life situations. To overcome these drawbacks, the so-called CRPs are proposed, in which experts discuss and modify their preferences to bring them closer to each other.

The concept of consensus has been interpreted from different points of view, from a classic view of full agreement (unanimity), usually difficult to achieve in practice, to other softer interpretations. In Ref.15, Saint and Lawson define consensus as *a state of mutual agreement between members of a group, where all legitimate concerns of individuals have been addressed to the satisfaction of the group*. Consensus aims at increasing satisfaction of the group regarding minimizing the misery, which is the main goal pursued in this paper.

The CRP is an iterative and dynamic process aimed at achieving a high degree of agreement before making the decision that solves the GDM problem. Therefore, an essential aspect in such processes is the definition of appropriate *consensus measures* to quantify the level of group agreement from experts' preferences. According to the type of computations and information fusion procedures,[14] the different consensus measures proposed in the literature can be classified as follows (see Figure 3):

- *Consensus measures based on distances to the collective preference*:[36–38] The group opinion, represented by the collective preference, P_c, is calculated by the aggregation of all individual preferences of experts, P_i. Consensus degrees are then obtained by computing the distances between each individual preference and the collective preference, $d(P_i, P_c)$.
- *Consensus measures based on distances between experts*:[39–42] For each different pair of experts in the group, (e_i, e_j), $i < j$, the degrees of similarity between their opinions are computed, based on distance metrics. Similarity values $L(P_i, P_j)$ are then aggregated to obtain consensus degrees.

In the past years, a large number of consensus models have been proposed, having each one different features attending to diverse criteria.[14,43] Figure 4 depicts a general scheme, which encompasses most of these existing approaches, having the following phases:

1. *Consensus Measurement*: All experts' preferences, $P_i, i \in \{1, ..., m\}$, are gathered to calculate the current group agreement level, by means of consensus measures.

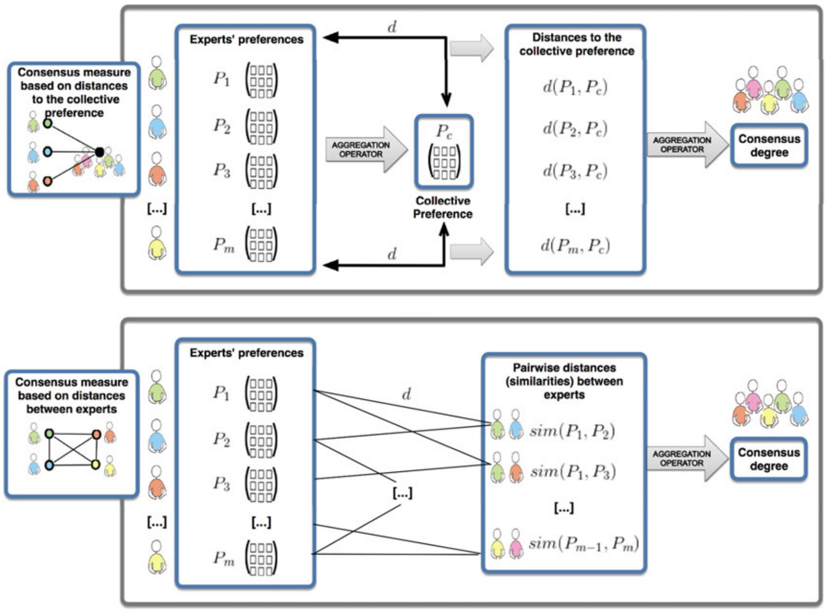

Figure 3. Types of consensus measures.

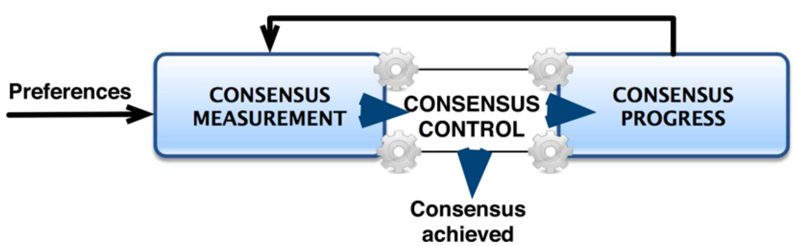

Figure 4. General CRP scheme.

2. *Consensus Control*: Compare the current consensus degree (CCD) with the consensus threshold (CT) defined at the beginning of the CRP. If the CCD is greater than the CT, then the CRP finishes, having achieved consensus. On the other hand, if the CCD is lower than the CT, the CRP continues with another round, until consensus is achieved or the number of consensus rounds exceeds the maximum number of rounds allowed.
3. *Consensus Progress*: To increase the agreement level in the following rounds of the CRP, different procedures can be applied, depending on whether the consensus model considers experts' sovereignty to let them modify their preferences based on feedback received, or such preferences are updated automatically:
   - *Feedback Generation*: In consensus models that implement this mechanism,[44–47] the moderator suggests by means of a feedback mechanism how experts should modify their preferences to bring them closer to the group opinion.
   - *Automatic Updates*: These consensus models[36,48,49] do not incorporate any feedback mechanism. Instead, they implement approaches in which experts provide their initial preferences, and automatic changes on preference values are applied across the CRP to increase the agreement level.

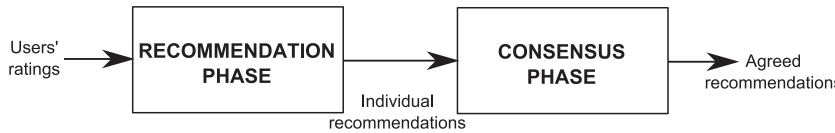

Figure 5. General scheme of the proposal.

REMARK 1. *The consensus model that will be utilized in our proposal of GRS is based on an automatic preferences updating mechanism.*

Further detail on the different types of existing consensus models can be found in Ref.14, where a taxonomy of consensus approaches in a fuzzy context was defined, being models categorized according to the use of feedback generation or automatic updates, as well as the type of consensus measure utilized in each one.

3. CONSENSUS-DRIVEN GROUP RECOMMENDER SYSTEM

This section introduces a novel consensus-based GRS model that delivers recommendations to group of users under a high level of consensus, based on their individual recommendations. In the following subsections, we describe in further details the phases of the underlying methodology in our proposal (see Figure 5):

1. *Recommendation Phase*: In this phase, a collaborative filtering algorithm is applied to obtain individual recommendations for each group member. The resulting recommendations over the top-n commonly predicted items are represented as preference orderings at the end of this phase, and they are used as input for the consensus phase.
2. *Consensus Phase*: The ordered recommendations for group members are transformed into fuzzy preference relations, and a CRP is then applied to bring such preferences closer to each other and obtain a collective preference under a high level of consensus, delivering the recommendations list upon it.

3.1. Notation

Let a GRS be such as

- $U = \{u_1, \ldots, u_q\}$ is the set of all q existing individuals or *users* in the GRS.
- $I = \{i_1, \ldots, i_t\}$ is the set of all t available products or *items* to be recommended.
- $G = \{g_1, \ldots, g_m\}$, $G \subset U$, is a group of m users of the GRS, $m << q$, to whom a list of products shall be recommended.
- D is the rating domain, i.e. the set of possible values that users can utilize to rate items. In this paper we consider $D = \{1, \ldots, 5\}$.
- $r_{il} \in D$ is the rating of user $u_i \in U$, over item $i_l \in I$.
- $\tilde{D}$ is the prediction domain, i.e. the set of possible prediction values that the GRS can assign to the pair formed by a user and an item in I not yet rated by him/her. In our case, $\tilde{D} = [1, 5]$.

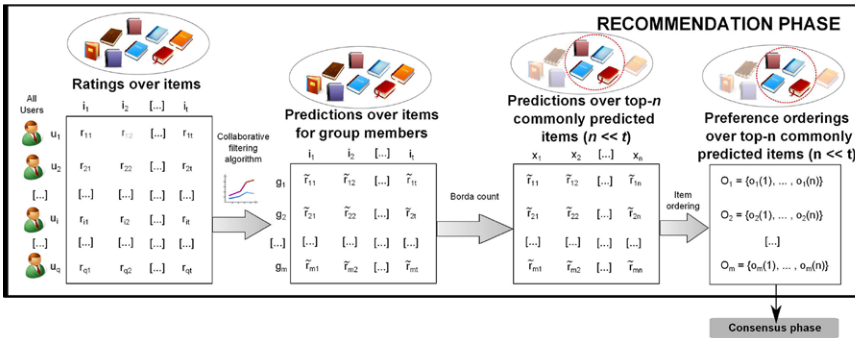

Figure 6. Scheme of computations carried out in the recommendation phase.

- $\tilde{r}_{il} \in \tilde{D}$ represents a prediction for group member $g_i \in G$ over an item i_l that he/she has not rated yet.
- $X = \{x_1, \ldots, x_n\}$, $X \subset I$, is the set of the top-n items commonly predicted to all members in a group G, with $n << t$.

REMARK 2. *Our proposal aims at seeking consensus among individual recommendations of group members over items in X. To do this, a CRP will be applied by using an automatic consensus model, hence our framework must be also modeled as a GDM problem, in which X represents the set of alternatives of such a problem and G represents the group of individuals taking part in it.*

3.2. Recommendation Phase

This phase includes the necessary computations to generate, for each member in the group, $g_i \in G$, a set of recommendations or predictions over items. Such predictions are ordered from the best to the worst one. A ranking or preference ordering, O_i, is generated as a result of this phase for each group member, g_i, over the top-n items $x_l \in X$ that have been commonly predicted to *all* the members in the group. The steps included in this phase are depicted in Figure 6, and they are described in detail below:

REMARK 3. *In Figure 6, all ratings r_{il} are depicted in the scheme for all the existing users and items, but only some of them may have a defined value, i.e. those corresponding to the items that have been rated by each user. The same situation applies with predictions for group members over all the items in I.*

(A) *Applying single user collaborative filtering algorithm*: First, given *all* users' ratings over some of the existing items in the GRS database, it is necessary to generate, for each group member in G, the individual predictions over items $i_l \in I$ not yet rated by him/her. To do this, first a collaborative filtering algorithm is applied to generate prediction values over items, $\tilde{r}_{il}$, for all the existing users. For this step, the collaborative filtering algorithm applied is user-based k-nearest neighbors algorithm (UBCF).[50]
Once applied the UBCF algorithm, only those predictions $\tilde{r}_{il}$ corresponding to group members $g_i \in G$ are taken into account in the following computations carried out in this phase.

(B) *Filtering of group member predictions over top-n commonly predicted items*: Group members' predictions computed after applying the collaborative filtering algorithm do not cover all the existing items in the GRS, i.e. two different members g_i and g_j may receive predictions over different subsets of items from I; hence, only some of the existing items in the RS database would be recommended to both of them simultaneously. In this step, the subset of items that have been commonly predicted for all the members in group G is determined. Moreover, if we consider a small group size and a large item database, the subset of commonly predicted items to group members might be too large; therefore, from this step onwards we will only consider the n items predicted as the best ones for all group members as a whole. Such items will be referred to as the top-n commonly predicted items for the group, $X = \{x_1, \ldots, x_n\}$, $X \subset I$, with $n << t$.

The following computations are carried out to obtain the set X of top-n commonly predicted items:

(i) Construct a set I^G, such that $X \subseteq I^G \subset I$, of all items commonly predicted to every member in G:

$$I^G = \{i_l \in I : \forall g_i \in G \exists \tilde{r}_{il} \in \tilde{D}\} \tag{3}$$

with $\tilde{D}$ being the prediction domain, e.g., the continuous interval $\tilde{D} = [1, 5]$ if $D = \{1, \ldots, 5\}$.

(ii) For each group member $g_i \in G$, rank commonly predicted items in I^G in a descending order of the prediction values assigned to him/her, $\tilde{r}_{il}$.

(iii) Given $n \in \mathbb{N}, n \leq |I^G|$, fixed a priori, a selection over all individual recommendation lists is applied, since GDM and CRP techniques are only capable of working with a small set of items. A number of rank aggregation techniques are applicable to this prefilter such as Borda or cumulative voting.[51] Our proposal applies single transferable voting[52] over the members recommendations to determine X.

(iv) Once the item set X has been determined, it is necessary to filter predictions, keeping those $\tilde{r}_{il}$ accomplishing $g_i \in G$ and $x_l \in X$, and leaving out those accomplishing $x_l \notin X$ or $u_i \notin G$.

(C) *Constructing preference orderings from top-n common predictions*: Let $\tilde{r}_i = [\tilde{r}_{i1}\tilde{r}_{i2} \ldots \tilde{r}_{in}]$ be each group member's vector of predictions over the top-n commonly predicted items. We then obtain its corresponding preference ordering $O_i = \{o_i(1), \ldots, o_i(n)\}$, being $o_i(\cdot)$ a permutation over the index set $\{1, \ldots, n\}$, such that $o_i(l) < o_i(k) iff \tilde{r}_{il} > \tilde{r}_{ik}$.[53] Therefore, items in X recommended to a group member g_i are ordered from the item having the highest prediction value, to the item having the lowest one.

3.3. Consensus Phase

Once generated, the individual predictions expressed as preference orderings over the top-n items recommended to the group members, this phase aims at using them to obtain a collective list of predictions for the group, which reflects a high level of consensus among individual predictions. For this, a CRP is applied by using an automatic consensus model for GDM problems with fuzzy preference relations.[54] Therefore, individual predictions of group members must be first expressed as fuzzy preference relations so that they can be managed by the consensus model.

The following two steps are carried out in this phase (see Figure 7):

(A) *Representing individual recommendations as fuzzy preference relations*: At the end of the recommendation phase (see Section 3.2), filtered predictions of group members

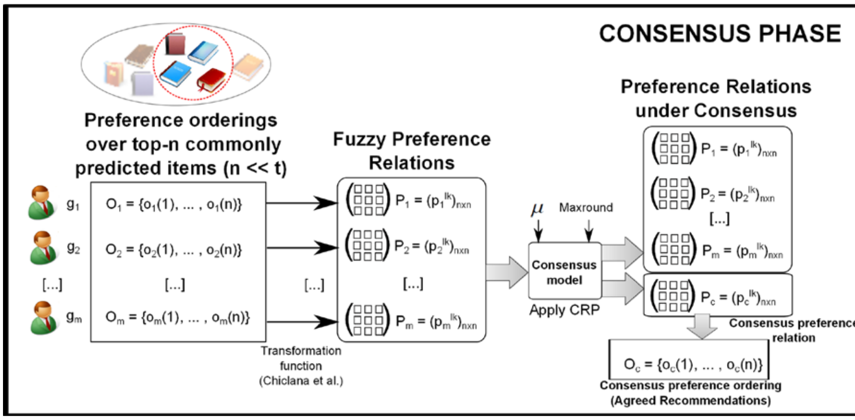

Figure 7. Scheme of computations carried out in the consensus phase.

were conducted into preference orderings in which, for each member, items were ranked from the best to the worst one according to their prediction values. Here, individual recommendations are first expressed as fuzzy preference relations before applying a CRP on them, since a consensus model for GDM with preference relations is utilized. Therefore, a transformation function is used to express each preference ordering O_i as a fuzzy preference relation P_i. In Ref.55, Chiclana et al. developed a number of transformation functions to deal with multiple preference representation formats in decision-making problems under uncertainty. In particular, they proposed the following method to construct the equivalent fuzzy preference relation $P_i = (p_i^{lk})_{n\times n}$ to a preference ordering $O_i = \{o_i(1), \ldots, o_i(n)\}$:

$$p_i^{lk} = \frac{1}{2}\left(1 + \frac{o_i(k) - o_i(l)}{n-1}\right) \tag{4}$$

EXAMPLE 1. *Let $O_1 = \{3, 2, 4, 1\}$ be a set of recommendations for member g_1 over $n = 4$ items expressed as a preference ordering, in which $o(4) = 1$ indicates that item x_4 is the item with the highest prediction value for g_1. Then, its corresponding fuzzy preference relation is obtained as follows:*

$$P_1 = \begin{pmatrix} - & 0.33 & 0.67 & 0.17 \\ 0.67 & - & 0.83 & 0.33 \\ 0.33 & 0.17 & - & 0 \\ 0.83 & 0.67 & 1 & - \end{pmatrix}$$

where, for instance, p_i^{12} has been computed by using Equation 4 as

$$p_i^{12} = \frac{1}{2}\left(1 + \frac{2-3}{3}\right) = 0.33 \tag{5}$$

(B) *Conducting the CRP*: Once each group member's preference ordering O_i has been conducted into a fuzzy preference relation P_i over the top-n commonly predicted items, a CRP is applied to such preferences for bringing them closer to each other gradually, until a high level of agreement among them is achieved. The consensus model utilized is an automatic model in which the values of preferences are automatically updated to bring them closer to each other (instead of implementing a feedback mechanism for individuals[14]) and it considers the following phases (see Figure 4), which are sequentially carried out at each round until consensus is achieved:

(1) *Consensus Measurement*: Fuzzy preference relations of group members, $P_i = (p_i^{lk})_{n\times n}$, $i = 1, \ldots, m$, are gathered and utilized to determine the level of agreement in the group. To do this, the following computations are carried out:
   (i) For each pair of members in the group g_i, g_j, $(i < j)$, a similarity matrix $SM_{ij} = (sm_{ij}^{lk})_{n\times n}$ is computed, with $sm_{ij}^{lk} \in [0, 1]$ being the degree of similarity between g_i and g_j's assessment on the pair of items (x_l, x_k).[56]
   (ii) A consensus matrix, $CM = (cm^{lk})_{n\times n}$ is obtained by aggregating pairwise similarity matrices. Each element $cm^{lk} \in [0, 1]$, $l \neq k$, is computed by applying an aggregation operator (e.g., the arithmetic mean or the OWA operator[57]) on similarity values.
   (iii) *Computing Consensus Degrees*: Once obtained CM, its values are successively aggregated[47] to obtain an overall consensus degree in the group, $cr \in [0, 1]$.

(2) *Consensus Control*: Once the overall consensus degree for the group is computed in the consensus measurement phase, here it is checked to determine whether it indicates enough agreement or not. If the consensus degree is enough, the CRP finishes having reached consensus among members' preferences. Otherwise, preference values are updated in the consensus progress phase to bring them closer to each other. A consensus threshold $\mu \in [0, 1]$ whose value fixed a priori indicates the minimum degree of agreement required among group members is used in this phase. The larger μ, the higher the consensus degree required. Furthermore, a parameter $Maxround$ can be used to limit the number of consensus rounds carried out without having reached μ.

(3) *Consensus Progress*: In this phase, group members' assessments p_i^{lk} which are farthest from consensus, are identified. A set of updates on the values of such identified assessments are then applied, with the aim of increasing consensus in the following rounds. The following steps are carried out in this phase:
   (i) A collective preference $P_c = (p_c^{lk})_{n\times n}$ is obtained, by aggregating individual assessments on each pair of items.
   (ii) A proximity matrix $PP_i = (pp_i^{lk})_{n\times n}$ whose values indicate the closeness degree between each member's preference relation and the collective preference, P_c, is computed for each $g_i \in G$.[57]
   (iii) Based on proximity matrices, some identification rules[47] are applied to identify group members g_i and pairs of items (x_l, x_k) whose assessments p_i^{lk} are not close enough to consensus. A set of direction rules are then applied to automatically update such identified assessments,[54] to increase the level of consensus in the group.

Once members' preferences have been updated, another CRP round starts by applying again the consensus measurement phase on them. After reaching consensus, the collective preference P_c, which reflects a high agreement level in the group, is used in the following step to deliver a list of agreed recommendations to the group.

(C) *Obtaining Agreed Recommendations List*: A consensus preference ordering is determined from the consensus preference relation P_c by applying a selection criterion,[12,29] to rank items from the best one to the worst one. More specifically, we consider the nondominance criterion[29] that assigns each item $x_l \in X$ a nondominance degree $ND(x_l) \in [0, 1]$ as follows:

(a) Construct a strict consensus preference relation $\tilde{P}_c = (\tilde{p}_c^{lk})_{n\times n}$, where:

$$\tilde{p}_c^{lk} = \begin{cases} \tilde{p}_c^{lk} - \tilde{p}_c^{kl} & \text{if } \tilde{p}_c^{lk} > \tilde{p}_c^{kl}, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

(b) Compute $ND(x_l)$ as

$$ND(x_l) = 1 - \max_k\{\tilde{p}_i^{kl}\} \tag{7}$$

The consensus preference ordering, denoted by

$$O_C = \{o_C(1), \ldots, o_C(n)\}$$

is then obtained by arranging items from their highest nondominance degree to their lowest one, i.e., if $o_C(l) < o_C(k)$ then $ND(x_l) \geq ND_{x_k}$.

As a result of conducting the consensus phase, members of the group are provided with a set of highly agreed recommendations (given by O_C) over the top-n commonly predicted items to all of them.

4. CASE STUDY

The case study to validate the proposal previously presented is deeply described in this section. To do this, the data set utilized in the study is first introduced, followed by the details about the experiment performed and the GRSs technique to be compared with our proposal. Afterward, the experiment results are shown and discussed. Finally, an example to visually demonstrate how the proposal is aiming to improve agreement on group recommendations, is shown.

4.1. Data Set

The data set used in this case study is the MovieLens data set, collected by the GroupLens Research Project[a] at the University of Minnesota. Specifically, ml-100k data set is used, and it consists of a hundred thousand ratings statements given by 943 users over 1682 movies in {1,2,3,4,5} domain.

4.2. Experiment Description

The data set considered does not contain information about possible groups of users; therefore, the group formation technique utilized is a random selection. Thirty different groups of five members each are formed to perform the case study.

The validation technique applied is 20% item holdout, which has been adjusted to be used on group recommendation by selecting the 20% of items rated by

[a]http://grouplens.org/

each group as a test set. Multiple executions have been performed to obtain more significant results.

Finally, the evaluation measures considered for this experiment are Area Under receiver operating characteristic Curve (AUC) and precision.[58] AUC is used to evaluate classifiers by computing ratio between the true-positive and false-positive rate as a threshold variable is adjusted, which builds a curve. In the case of recommender systems, the threshold variable is the size of the recommendation list. Regarding precision, this measure shows the ratio between the true-positive rate and the number of recommended items. For both measures, the higher the value, the better the GRS performance.

4.3. Group Recommender Systems Comparison Analysis

The aim of this case study is to compare our proposal of consensus-driven GRS model with different group recommendation techniques focused on delivering satisfactory recommendations for all members. Therefore, the use of average aggregation as a baseline technique would not be correct since it does not take into account individual satisfaction, delivering recommendations that members might see as deviated toward other members' preferences.

Therefore, we take group recommendation with recommendation aggregation using minimum as the baseline technique, taking UBCF as the single user recommender system. As UBCF has a number of improvements,[21] Pearson correlation coefficient is taken as similarity measure. Data sparsity can lead to poor values for similarity measure due to a small number of co-rated items; therefore, a correction with relevance factor of 20 is applied to penalize similarities computed with less than 20 corated items. Weighted sum is taken as a prediction technique to aggregate neighborhood ratings. For the sake of comparability of results, the selection of top-n items over which the CRP is carried out is applied as described in Section 3.

Our proposal uses the above mentioned UBCF recommender system, but it performs the recommendation aggregation based on a CRP in which the consensus degree desired (given by the consensus threshold μ) needs to be set; hence, the consensus degree values considered are $\{0.8, 0.85, 0.9\}$.

4.4. Results

The result of AUC (see Figure 8) indicates that applying the CRP clearly improves the baseline results for the three consensus threshold values tested and states that the optimal value of consensus degree in this data set is 0.8, decreasing steadily as consensus degree increments.

Based on the precision measure, the results are now examined in a more detailed way. As shown in Figure 9, different configurations for the CRP module produce variations in recommendation lists delivered to groups. comparing consensus techniques, the results show that applying a consensus degree of 0.8 benefits the recommendation results in comparison with recommendations subject to a higher consensus level. If the results are analyzed regarding the list sizes considered, the results show that for recommendation lists of size lower than four elements the

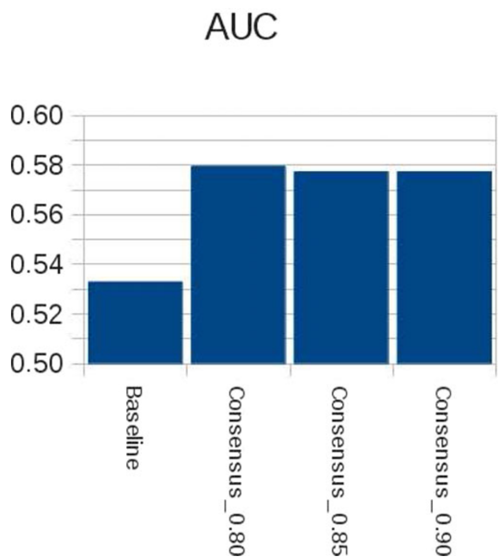

Figure 8. AUC for evaluated configurations.

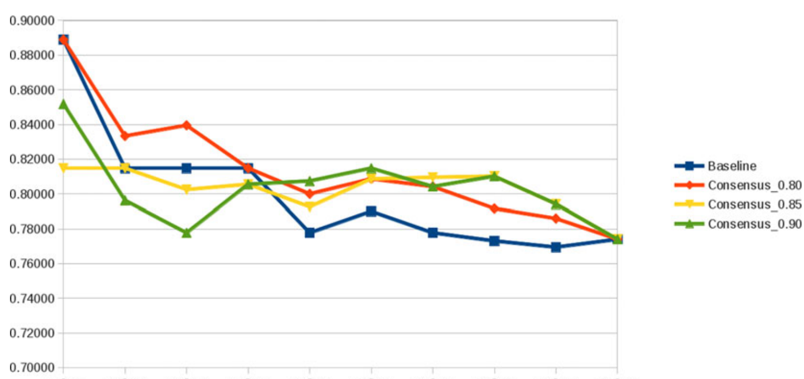

Figure 9. Precision at certain recommendation list sizes for evaluated configurations.

optimal consensus degree value is 0.8. For recommendation lists of size 4 or more, the consensus configurations shown in the precision diagram present similar values for precision, revealing a similar behavior in these cases.

REMARK 4. *In Figure 9, the output of the GRSs depicted are given in the same set of 10 items. Thus, the only difference between configurations is the ordering of those 10 items that the GRSs provided as the output. Precision does not take into account whether a good item is ranked in the first or last position within the ordering. This causes that all the GRSs evaluated in our case study present the same precision at list size of 10 (see Pr@10 in Figure 9).*

4.5. Graphical Visualization

Once the case study has been presented, we consider necessary to graphically depict the results on an intuitive manner to visualize users' preferences with respect to the baseline GRS recommendations and the agreed recommendations (AR). For this reason, in this section we describe an specific example with a group of users extracted from the MovieLens data set.

There are previous approaches on recommendation visualization techniques for individual recommender systems, such as the one presented by Kagie et al. in Ref. 59. These proposals represent the items in the recommendation set to show similarities between products, e.g., the closer two items are depicted with each other, the more similar they are. Following similar ideas, we represent group members' preferences to visualize where their opinions are regarding the group preference. As this problem has already been addressed in GDM, we use a Self-Organizing Maps-based[60] tool for visualizing preferences, so-called MENTOR.[61]

To apply a CRP before comparing graphically the differences between the baseline GRS output and the AR, we use a multiagent-based consensus support system[54] to compute the consensus ranking (see consensus phase, Section 3.3).

A group of five users from MovieLens data set, $U = u_1, ..., u_5$, is taken, and each one has a different prediction ranking. The recommendation lists are reduced taking the first 10 common films of the user rankings, selected using top-*n* member items, which are used as alternatives of a GDM problem, $X = \{x_1, ..., x_{10}\}$:

- x_1: Three Colors: Red.
- x_2: The Fugitive.
- x_3: A Space Odyssey.
- x_4: Manon of the Sprius.
- x_5: Delicatessen.
- x_6: Nikita.
- x_7: The Princess Bride.
- x_8: Raiders of the Lost Ark.
- x_9: Return of the Jedi.
- x_{10}: The Godfather II.

Once performed the initial recommendation phase, a first ranking is obtained from individual recommendations. A CRP is then applied to rerank the recommendations to improve agreement. A comparison among the baseline GRS rank and the agreed recommendations is shown in Table I, showing the differences between both rankings. For instance, "Return of the Jedi," which is the first one in the GRS ranking, changes to the second position, whereas "The princess bride," which is the second one in the GRS ranking, shifts toward the top of the AR. We can see the most drastic change in "The Fugitive," which change from the fourth to the ninth position.

Regarding the user's visualization (Figure 10), the consensus recommendation is situated in the center of the graphic, because CRPs take equidistances to bring users' opinions closer to the group opinion. On the other hand, GRSs attempt to minimize users' misery delivering less satisfactory recommendations to the group without taking into account individuals recommendations when computing group recommendations.

Table I. Comparison between GRS ranking and CR.

Baseline Group Recommendation	Agreed Recommendation
Return of the Jedi	The Princess bride
The Princess Bride	Return of the Jedi
Manon of the Sprius	Manon of the Sprius
The Fugitive	Three Colors: Red
Raiders of the Lost Ark	Raiders of the Lost Ark
The Godfather II	Nikita
Three Colors: Red	A Space Odyssey
Nikita	The Godfather II
Delicatessen	The Fugitive
A Space Odyssey	Delicatessen

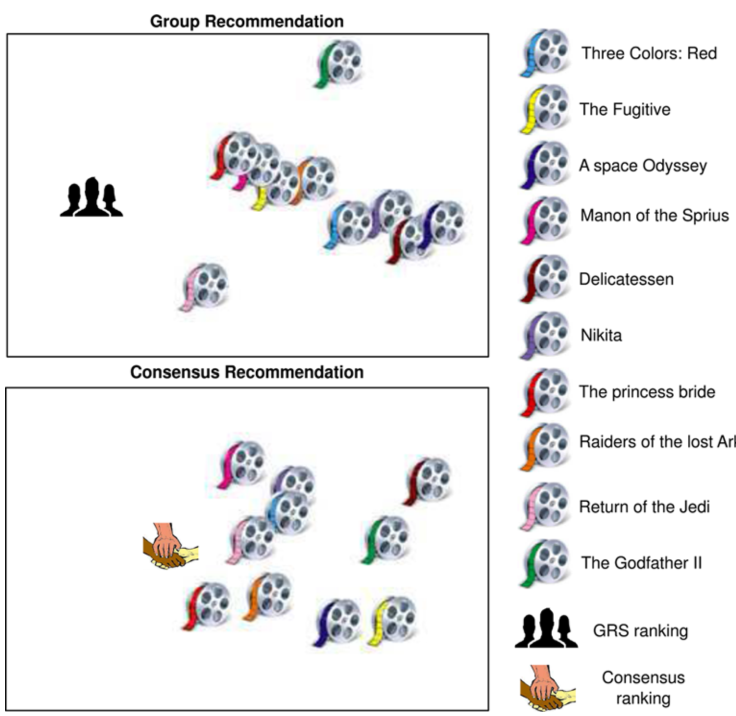

Figure 10. Users' visualization regarding group recommendation and consensus recommendation.

5. CONCLUDING REMARKS

In this paper, a novel technique for group recommendations based on CRPs has been proposed, to provide existing GRSs with the added value of improving group members' satisfaction regarding the items recommended. The case study results showed that applying CRPs in group recommendation clearly improves the results compared with baseline group recommending techniques.

The results suggest that bringing consensus into recommendation processes is a promising future research direction; hence, the exploration of other approaches developed in group decision making area might benefit recommendation.

Future works are mainly focused on developing a graphical tool, which allows users to visualize their position regarding items and other users in GRSs.

ACKNOWLEDGEMENTS

This work is partially supported by the Spanish Ministry of Education through the Research Project TIN-2012-31263, the Spanish Ministry of Education, Culture and Sport FPU fellowship (FPU2013/01151) and ERDF.

References

1. Ricci, F, Rokach, L, Shapira, B, Kantor, PB. Introduction to recommender systems handbook. Berlin: Springer; 2011.
2. O'Connor M, Cosley D, Konstan JA, Riedl J. Polylens: a recommender system for groups of users. In: Proc ECSCW'01;2001. pp 199–218.
3. McCarthy JF, Anagnost TD. Musicfx: an arbiter of group preferences for computer supported collaborative workouts. In: Proc 1998 ACM Conf on Computer-Supported Cooperative Work; Seattle, WA, November 14–18, 1998. pp 363–372.
4. McCarthy K, Salamó M, Coyle L, McGinty L, Smyth B, Nixon P. Cats: a synchronous approach to collaborative group recommendation. In: FLAIRS Conf, Melbourne Beach, FL; 2006. pp 86–91.
5. Noguera JM, Barranco MJ, Segura RJ, Martínez L. A mobile 3d-gis hybrid recommender system for tourism. Inform Sci 2012;215:37–52.
6. De Pessemier T, Dooms S, Martens L. Comparison of group recommendation algorithms. Multimedia Tools Appl 2014;72(3):2497–2541.
7. Masthoff J. Group recommender systems: combining individual models. In: Ricci F, Rokach L, Shapira B, Kantor PB, editors. Recommender systems handbook. Berlin: Springer; 2011. pp 677–702.
8. Jameson A, Smyth B. Recommendation to groups. In: The adaptive web. Berlin: Springer; 2007. pp 596–627.
9. Cantador I, Castells P. Group recommender systems: new perspectives in the social web. In: Recommender systems for the social web, Berlin: Springer; 2012. pp 139–157.
10. Butler CTL, Rothstein A. On conflict and consensus: A handbook on formal consensus decision making. Portland, ME: Food Not Bombs Publishing; 2006.
11. Lu J, Zhang G, Ruan D, Wu F. Multi-objective group decision making. London: Imperial College Press; 2006.
12. Herrera F, Herrera-Viedma E, Verdegay JL. A sequential selection process in group decision making with linguistic assessments. Inform Sci 1995;85(4):223–239.
13. Herrera-Viedma E, García-Lapresta JL, Kacprzyk J, Fedrizzi M, Nurmi H, Zadrozny S, editors. Consensual processes. Stud Fuzziness Soft Comput 2011;267.
14. Palomares I, Estrella FJ, Martínez L, Herrera F. Consensus under a fuzzy context: taxonomy, analysis framework AFRYCA and experimental case of study. Inform Fusion 2014;20:252–271.
15. Saint S, Lawson JR. Rules for reaching consensus. A modern approach to decision making. New York: Jossey-Bass; 1994.
16. Vozalis MG, Margaritis KG. Using SVD and demographic data for the enhancement of generalized collaborative filtering. Inform Sci 2007;177(15):3017–3037.
17. Castro J, Rodríguez RM, Barranco MJ. Weighting of features in content-based filtering with entropy and dependence measures. Int J Comput Intell Syst 2014;7(1):80–89.

18. Lops P, De Gemmis M, Semeraro G. Content-based recommender systems: state of the art and trends. In: Recommender systems handbook. Berlin: Springer; 2011. pp 73–105.
19. Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems. Computer 2009;42(8):30–37.
20. Ben Schafer J, Frankowski D, Herlocker J, Sen S. Collaborative filtering recommender systems. In: The adaptive web, Berlin: Springer; 2007. pp 291–324.
21. Breese JS, Heckerman D, Kadie C. Empirical analysis of predictive algorithms for collaborative filtering. In: Proc 14th Conf on Uncertainty in Artificial Intelligence. San Francisco, CA: Morgan Kaufmann; 1998. pp 43–52.
22. Jun Ahn H. A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. Inform Sci 2008;178(1):37–51.
23. Fouss F, Achbany Y, Saerens M. A probabilistic reputation model based on transaction ratings. Inform Sci 2010;180(11):2095–2123.
24. Sarwar B, Karypis G, Konstan J, Riedl J. Item-based collaborative filtering recommendation algorithms. In; Proc 10th Int Conf on World Wide Web. ACM, 2001. pp 285–295.
25. Zheng N, Bao H. Flickr group recommendation based on user-generated tags and social relations via topic model. In: Advances in neural networks. Berlin: Springer; 2013. pp 514–523.
26. Myszkorowski K, Zakrzewska, D. Using fuzzy logic for recommending groups in e-learning systems. In: Computational collective intelligence. technologies and applications. Berlin: Springer; 2013. pp 671–680.
27. Meena R, Bharadwaj K. Group recommender system based on rank aggregation an evolutionary approach. In: Mining intelligence and knowledge exploration. Switzerland: Springer International Publishing; 2013. pp 663–676.
28. Song Y, Hu Z, Liu H, Shi Y, Tian H. A novel group recommendation algorithm with collaborative filtering. In Int Conf on Social Computing; Washington, DC, September 8–14 (SocialCom); 2013. pp 901–904.
29. Orlovsky SA. Decision-making with a fuzzy preference relation. Fuzzy Sets Syst 1978;1(3):155–167.
30. Fu C, Yang SL. An evidential reasoning based consensus model for multiple attribute group decision analysis problems with interval-valued group consensus requirements. Eur J Oper Res 2012;223(1):167–176.
31. Herrera F, Herrera-Viedma E. Linguistic decision analysis: steps for solving decision problems under linguistic information. Fuzzy Sets Syst 2000;115(1):67–82.
32. Rodríguez RM, Martínez L. An analysis of symbolic linguistic computing models in decision making. Int J Gen Syst 2013;42(1):121–136.
33. Parreiras RO, Ekel P, Bernardes F Jr. A dynamic consensus scheme based on a nonreciprocal fuzzy preference relation modeling. Inform Sci 2012;211(1):1–17.
34. Roubens M. Fuzzy sets and decision analysis. Fuzzy Sets Syst 1997;90(2):199–206.
35. Beliakov G, Pradera A, Calvo T. Aggregation functions: A guide for practitioners. Berlin; Springer; 2007.
36. Ben-Arieh D, Chen Z. Linguistic labels aggregation and consensus measure for autocratic decision-making using group recommendations. IEEE Trans Syst Man Cybernet A 2006;36(1):558–568.
37. Herrera F, Herrera-Viedma E, Verdegay JL. A model of consensus in group decision making under linguistic assessments. Fuzzy sets and Syst 1996;78(1):73–87.
38. Spillman B, Bezdek JC, Spillman R. Development of an instrument for the dynamic measurement of consensus. Commun Monogr 1979;46:1–12.
39. Bordogna G, Fedrizzi M, Pasi G. A linguistic modeling of consensus in group decision making based on OWA operators. IEEE Trans Syst Man Cybernetics A 1997;27(1):126–133.
40. Herrera F, Herrera-Viedma E, Verdegay JL. Linguistic measures based on fuzzy coincidence for reaching consensus in group decision making. Int J Approx Reason 1997;16(3-4):309–334.

41. Kacprzyk J, Fedrizzi M. A “soft” measure of consensus in the setting of partial (fuzzy) preferences. Eur J Oper Res 1988;34(1):316–325.
42. Kacprzyk J, Fedrizzi M, Nurmi H. Group decision making and consensus under fuzzy preferences and fuzzy majority. Fuzzy Sets Syst 1992;49(1):21–31.
43. Herrera-Viedma E, Cabrerizo FJ, Kacprzyk J, Pedrycz W. A review of soft consensus models in a fuzzy environment. Inform Fusion2014; 17: 4–13.
44. Bryson N. Group decision-making and the analytic hierarchy process. exploring the consensus-relevant information content. Comput Oper Res 1996;23(1):27–35.
45. Carlsson C, Ehrenberg D, Eklund P, Fedrizzi M, Gustafsson P, Lindholm P, Merkuryeva G, Riissanen T, Ventre AGS. Consensus in distributed soft environments. Eur J Oper Res 1992;61(1–2):165–185.
46. Herrera-Viedma E, Herrera F, Chiclana F. A consensus model for multiperson decision making with different preference structures. IEEE Trans Syst Man Cybernet 2002;32(3): 394–402.
47. Mata F, Martínez L, Herrera-Viedma E. An adaptive consensus support model for group decision-making problems in a multigranular fuzzy linguistic context. IEEE Trans Fuzzy Syst 2009;17(2):279–290.
48. Wu Z, Xu J. A consistency and consensus based decision support model for group decision making with multiplicative preference relations. Decis Support Syst 2012;52(3):757–767.
49. Zhang G, Dong Y, Xu Y, Li H. Minimum-cost consensus models under aggregation operators. IEEE Trans Syst Man Cybernet 2011;41(6):1253–1261.
50. Sarwar B, Karypis G, Konstan J, Riedl J. Incremental singular value decomposition algorithms for highly scalable recommender systems. In: Fifth Int Conf on Computer and Information Science; Dhaka, Bangladesh, December 27–28, 2002. pp 27–32.
51. J Arrow K, Sen A, Suzumura K. Handbook of social choice & welfare, volume 2. Amsterdam: Elsevier; 2010.
52. Dwork C, Kumar R, Naor M, Sivakumar D. Rank aggregation methods for the web. In: Proc 10th Int Conf on World Wide Web; Hong Kong, May 1–5, 2001. pp 613–622.
53. Tanino T. Fuzzy preference orderings in group decision making. Fuzzy Sets Syst 1984;12(2):117–131.
54. Palomares I, Martínez L. A semi-supervised multi-agent system model to support consensus reaching processes. IEEE Trans Fuzzy Syst 2014;22(4):762–777.
55. Chiclana F, Herrera-Viedma E, Alonso S, Marqués-Pereira RA. Preferences and consistency issues in group decision making. In: Bustince H, Herrera F, Montero J, editors. Fuzzy sets and their extensions: Representation, aggregation and models. Berlin: Springer; 2008. pp 219–237.
56. Herrera-Viedma E, Martínez L, Mata F, Chiclana F. A consensus support system model for group decision making problems with multigranular linguistic preference relations. IEEE Trans Fuzzy Syst 2005;13(5):644–658.
57. Palomares I, Liu J, Xu Y, Martínez L. Modelling experts’ attitudes in group decision making. Soft Comput 2012;16(10):1755–1766.
58. Herlocker JL, Konstan JA, Terveen LG, Riedl JT. Evaluating collaborative filtering recommender systems. ACM Trans Inform Syst 2004;22(1):5–53.
59. Kagie M, Van Wezel M, Groenen PJF. Map based visualization of product catalogs. Berlin: Springer; 2011.
60. Kohonen T. Self-organizing maps. Heidelberg, Germany: Springer; 1995.
61. Palomares I, Martínez L, Herrera F. MENTOR: A graphical monitoring tool of preferences evolution in large-scale group decision making. Knowl-Based Syst 2014;58;66–74. Special issue on intelligent decision support making tools and techniques.