Using Qualitative Reasoning to Model Users Profiles: An Approach to Movie Recommendations

Llorenç Roselló[1], **David Garcia**[2], **Núria Agell**[2], **Mónica Sánchez**[1] and **Francesc Prats** [1]
[1] BarcelonaTech (UPC), Barcelona, Spain
{llorenc.rosello, francesc.prats, monica.sanchez}@upc.edu
[2] ESADE Business School, Ramon Llull University, Barcelona, Spain
{david.garcia19, nuria.agell}@esade.edu

Abstract

This paper presents a methodology for a collaborative recommender system (RS). The methodology is based on the compatibility of groups of users defining their profiles via a qualitative order-of-magnitude model. The distributive lattice structure of the space of qualitative descriptions is considered in defining the distance between existing users and the RSs new users. An application to movie recommendations is presented to show and compare the efficiency of the proposed methodology.

1 Introduction

The RS proposed is a collaborative memory based system where the user is recommended items based on users with similar profiles and preferences. Several different approaches have been discussed in the literature to address the problem of finding user similarities, such as: correlation based [Resnik *et al.*, 1994; Schardanand and Maes, 1995], cosine-based [Breese *et al.*, 1998; Sarwar *et al.*, 2001], and graph theoretic [Aggarwal *et al.*, 1999].

This RS differs from others because it uses a heuristic that allows different levels of precision to be considered simultaneously. This ability of the proposed RS is crucial for recommendation of products whose main features are addressed to users' sensorial perceptions. In this context, users often do not know how to express their preferences with precision.

We present recommendations that find user similarities in terms of profile compatibility with other users. Rather than using classical methods, we put forward an approach to recommending by searching through the most similar neighbors, using a degree of consensus directly or through a dive function that permits consensus based on underlying common values. The degree of consensus allows us to measure the compatibility of a group of users. The previous search of compatible groups with the user makes the recommendation easier and the cost of calculating a minimum distance lower.

This proposed methodology incorporates incomplete or partial knowledge into the recommendation process using qualitative reasoning techniques to assess affinity of its users for recommendations.

Two main advances of this paper with respect to previous works can be highlighted. On the one hand, from a theoretical point of view, the definition of a distance among users presented in Subsection 2.2. On the other hand, the application of the degree of consensus previously defined together with the presented distance to build a recommender system.

This paper is structured as follows. Section 2 introduces the theoretical framework. In Section 3, the recommender system algorithm is presented. In Section 4, an experimental case in movie recommendation is introduced, and its results are compared with two non-personalized models. Conclusions and future research are drawn in Section 5.

2 Theoretical Framework

Qualitative Reasoning (QR) is a sub-area of Artificial Intelligence that seeks to understand and explain human beings' ability to reason without having precise information [Forbus, 1996]. In recommendation processes, it is not unusual for a situation to arise in which different levels of precision have to be worked with simultaneously, depending on the information available to each user.

The RS proposed in this paper requires measuring compatibility among users. The concept of compatibility is based on group consensus theory. In particular, in this paper, we use the definition of consensus introduced in ([Roselló *et al.*, 2010]), which is based on a qualitatively-described system in terms of absolute order-of-magnitude ([Travé-Massuyès and Dague, 2003]).

Order-of-magnitude models are essential among the theoretical tools available for qualitative reasoning [Dague, 1993; Kalagnanam *et al.*, 1991]. They aim to capture order-of-magnitude commonsense inferences [Travé-Massuyès and Dague, 2003]. The *classic order-of-magnitude qualitative spaces* are built from a set of ordered basic qualitative labels. A general algebraic structure, called Qualitative Algebra or Q-algebra, was defined based on this framework, providing a mathematical structure to unify sign algebra and interval algebra through a continuum of qualitative structures built from the roughest to the finest partition of the real line. Q-algebras and their algebraic properties have been extensively studied [Travé-Massuyès and Dague, 2003]

Let us consider a finite set of *basic* labels, $\mathbb{S}_* = \{B_1, \ldots, B_n\}$, which is totally ordered as a chain: $B_1 < \ldots < B_n$. Usually, each basic label corresponds to a linguistic term, for instance "extremely bad" $<$ "very bad" $<$

"bad" $<$ "acceptable" $<$ "good" $<$ "very good"$<$ "extremely good".

The complete description universe for the Order-of-Magnitude Space OM(n) with granularity n, is the set $\mathbb{S}_n$:

$$\mathbb{S}_n = \mathbb{S}_* \cup \{[B_i, B_j] \,|B_i, B_j \in \mathbb{S}_*, i < j\},$$

where the label $[B_i, B_j]$ with $i < j$ is defined as the set $\{B_i, B_{i+1}, \ldots, B_j\}$.

Consistent with the former example of linguistic labels, the label "moderately good" can be represented by ["acceptable", "good"], i.e., $[B_4, B_5]$. The label "don't know" is represented by ["extremely bad", "extremely good"], i.e., $[B_1, B_7]$. This least precise label is denoted by the symbol ?, i.e., $[B_1, B_n] \equiv$?.

There is a partial order relation $\leq_P$ in $\mathbb{S}_n$, "to be more precise than", given by $L_1 \leq_P L_2 \Longleftrightarrow L_1 \subset L_2$.

The structure $OM(n)$ permits to work with all different levels of precision from the basic labels to the ? label.

The next subsections are the theoretical foundation of the RS presented. This theoretical foundation is based on [Roselló *et al.*, 2010].

2.1 Groups of Compatible Users

Let Λ be the set features that is qualitatively described by means of the $\mathbb{S}_n$ labels, $\Lambda = \{a_1, \ldots, a_N\}$.

The qualitative description is carried out by each user/evaluator and is represented by the function: $Q : \Lambda \rightarrow \mathbb{S}_n$, where $a_l \mapsto Q(a_l) = \mathcal{E}_l$ is the qualitative label with which the evaluator describes a_l.

Let $\mathcal{Q} = \{Q \mid Q : \Lambda \rightarrow \mathbb{S}_n\}$ be the set of all possible qualitative descriptions of Λ over $\mathbb{S}_n$; a group of users determine a subset of $\mathcal{Q}$. Given $Q, Q' \in \mathcal{Q}$, two different operations are defined between them.

1. The mix operation: The operation $Q \sqcup Q'$ leads to a new qualitative description function $Q \sqcup Q' : \Lambda \rightarrow \mathbb{S}_n$ such that, for any $a_t \in \Lambda$,

$$(Q \sqcup Q')(a_l) = Q(a_l) \sqcup Q'(a_l),$$

where $\sqcup$ is the connex union of labels, i.e. the minimum label that contains $Q(a_l)$ and $Q'(a_l)$: $[B_i, B_j] \sqcup [B_h, B_k] = [B_{\min\{i,h\}}, B_{\max\{j,k\}}]$, using the convention $[B_i, B_i] = B_i$.

2. The common operation: The concept of consensus between two qualitative descriptions, Q and Q', is required in order to introduce the common operation:

Two qualitative descriptions Q, Q' are *in consensus*, $Q \rightleftarrows Q'$, iff

$$Q(a_l) \cap Q'(a_l) \neq \emptyset \quad \forall a_l \in \Lambda. \tag{1}$$

Given Q and Q' where $Q \rightleftarrows Q'$, the *common* $Q \cap Q'$ operation produces a new qualitative description function $Q \cap Q' : \Lambda \rightarrow \mathbb{S}_n$ such that

$$(Q \cap Q')(a_l) = Q(a_l) \cap Q'(a_l) \;\; \forall a_l \in \Lambda.$$

In general, a set $\{Q_i\}_{i \in I} \subset \mathcal{Q}$ of qualitative descriptions of Λ over $\mathbb{S}_n$ is *in consensus* iff $\cap_{i \in I} Q_i(a_t) \neq \emptyset \;\; \forall a_t \in \Lambda$. The consensus concept of a set of qualitative descriptions is used in the RS to model compatibility among profiles of the corresponding group of users.

The algebraic structure of the set $\mathcal{Q}$ and the $\sqcup$ and $\cap$ operations are given by the next results (the proofs can be found in [Roselló *et al.*, 2010]): $(\mathcal{Q}, \sqcup, \cap)$ is a weak partial lattice, and if $\mathcal{Q}_L$ is a subset of $\mathcal{Q}$ which is in consensus, then $(\mathcal{Q}_L, \sqcup, \cap)$ is a *distributive lattice* [Birkhoff, 1967].

2.2 A Distance among Users

Let us suppose that there exists a subset $\mathcal{Q}_L$ of $\mathcal{Q}$ which is in consensus (if this situation does not hold, in Section 2.3 a process to obtain consensus is presented). This section is devoted to define a distance between two qualitative descriptions $Q, Q' \in \mathcal{Q}_L$.

Definition 1 *In the lattice* $(\mathcal{Q}_L, \sqcup, \cap)$ *the* null element $0_{\mathcal{Q}_L}$ *is defined as* $0_{\mathcal{Q}_L} = \sqcup_{Q_i \in \mathcal{Q}_L} Q_i$*, and the* universal element $1_{\mathcal{Q}_L}$ *is defined as* $1_{\mathcal{Q}_L} = \cap_{Q_i \in \mathcal{Q}_L} Q_i$*.*

The null element and the universal elements verify for all $Q \in \mathcal{Q}_L$:

$$0_{\mathcal{Q}_L} \sqcup Q = 0_{\mathcal{Q}_L}, 0_{\mathcal{Q}_L} \cap Q = Q,$$

$$1_{\mathcal{Q}_L} \sqcup Q = Q, 1_{\mathcal{Q}_L} \cap Q = 1_{\mathcal{Q}_L},$$

and then, considering the partial order relation defined by $Q \leq Q'$ iff $Q \sqcup Q' = Q$, we have:

$$0_{\mathcal{Q}_L} \leq Q \leq 1_{\mathcal{Q}_L}.$$

Recall the definition of *chain*: a totally ordered set of a poset.

By "*x covers y*" it is meant that $y < x$ and that $y < z < x$ is not satisfied by any z. A finite chain $a_1 < a_2 < \ldots < a_k$ is a *maximal chain* if each a_{i+1} covers a_i for $i = 1, \ldots, k-1$, and it is denoted by $[a_1, a_k]$.

Let us assume that Λ is a finite set. Since $\mathbb{S}_n$ is also finite, then all the chains in $(\mathcal{Q}_L, \sqcup, \cap)$ are finite. Therefore, all finite maximal chains between fixed end points have the same length (Jordan-Dedekind theorem) [Birkhoff, 1967].

Definition 2 *If $Q, Q' \in \mathcal{Q}_L$, the length of a chain with end points Q and Q', $l([Q, Q'])$, is the cardinal of any maximal chain between Q and Q'. The length of $Q \in \mathcal{Q}_L$, $l(Q)$, is the length of $[0_{\mathcal{Q}_L}, Q]$.*

In the distributive lattice $(\mathcal{Q}_L, \sqcup, \cap)$ the following statement is satisfied for all Q and Q' in $\mathcal{Q}_L$:

$$l(Q) + l(Q') = l(Q \sqcup Q') + l(Q \cap Q'). \tag{2}$$

Lemma 1 *Since in $\mathcal{Q}_L$ the operations $\sqcup$ and $\cap$ are the infimum and supremum respectively then:*

$$(Q \cap Q') \sqcup (Q' \cap Q'') \geq Q' \tag{3}$$

$$Q' \geq (Q \sqcup Q') \cap (Q' \sqcup Q''). \tag{4}$$

Proof: It is a simple exercise to check that $((Q \cap Q') \sqcup (Q' \cap Q'')) \cap Q' = (Q \cap Q') \sqcup (Q' \cap Q'')$ and $Q' \cap ((Q \sqcup Q') \cap (Q' \sqcup Q'')) = Q'$. $\square$

The next theorem defines a distance in the lattice $(\mathcal{Q}_L, \sqcup, \cap)$:

Theorem 1 *In the lattice* $(\mathcal{Q}_L, \sqcup, \cap)$, *the function* $d : \mathcal{Q}_L \times \mathcal{Q}_L \to \mathbb{R}$ *defined as*

$$d(Q, Q') = l(Q \cap Q') - l(Q \sqcup Q') \tag{5}$$

is a distance.

Proof:

1. Positive definiteness: Because $Q \sqcup Q' \leq Q \cap Q' \ \forall Q, Q'$ it is trivial to see that $l(Q \sqcup Q') \leq l(Q \cap Q')$, so $d(Q, Q') \geq 0$.
   If $Q = Q'$ then $d(Q, Q') = 0$. Conversely,
   $$d(Q, Q') = 0 \Rightarrow l(Q \sqcup Q') = l(Q \cap Q'),$$
   and this, together with the fact that $Q \sqcup Q' \leq Q \cap Q'$ and the Jordan-Dedekind theorem, leads to $Q \sqcup Q' = Q \cap Q'$.
   By the absorptive laws of lattices:
   $Q \cap (Q \sqcup Q') = Q$ and $Q \sqcup (Q \cap Q') = Q$.
   We have
   $Q = Q \cap (Q \sqcup Q') = Q \cap (Q \cap Q') = Q \cap Q'$,
   $Q' = Q' \cap (Q \sqcup Q') = Q' \cap (Q \cap Q') = Q \cap Q'$,
   so $Q = Q'$.
2. Symmetry: Since $\sqcup$ and $\cap$ are commutative, $d(Q, Q') = d(Q', Q)$.
3. Triangle inequality: For all $Q, Q', Q'' \in \mathcal{Q}_L$ $d(Q, Q') \leq d(Q, Q'') + d(Q'', Q')$.
   We have
   $d(Q, Q'') + d(Q'', Q') = l(Q \cap Q'') + l(Q' \cap Q'') - (l(Q \sqcup Q'') + l(Q' \sqcup Q''))$.
   The two first summands can be expressed using the property (2):
   $l(Q \cap Q'') + l(Q' \cap Q'') = l((Q \cap Q'') \sqcup (Q' \cap Q'')) + l((Q \cap Q'') \cap (Q' \cap Q''))$,
   and then, by (3)
   $$l(Q \cap Q'') + l(Q' \cap Q'') \geq l(Q'') + l((Q \cap Q' \cap Q'')).$$
   Similarly, from (2):
   $l(Q \sqcup Q'') + l(Q' \sqcup Q'') = l((Q \sqcup Q'') \sqcup (Q' \sqcup Q'')) + l((Q \sqcup Q'') \cap (Q'' \sqcup Q'))$,
   and then, by (4):
   $$l(Q \sqcup Q'') + l(Q' \sqcup Q'') \leq l(Q \sqcup Q' \sqcup Q'') + l(Q'').$$
   So,
   $$d(Q, Q'') + d(Q'', Q') \geq l(Q \cap Q' \cap Q'') - l(Q \sqcup Q' \sqcup Q'').$$
   Now, using the fact that
   $$Q \cap Q' \cap Q'' \geq Q \cap Q' \Rightarrow l(Q \cap Q' \cap Q'') \geq l(Q \cap Q')$$
   and
   $$Q \sqcup Q' \sqcup Q'' \leq Q \sqcup Q' \Rightarrow l(Q \sqcup Q' \sqcup Q'') \leq l(Q \sqcup Q'),$$
   we conclude that
   $d(Q, Q'') + d(Q'', Q') \geq l(Q \cap Q') - l(Q \sqcup Q') = d(Q, Q')$.

$\square$

In the next sections the concepts introduced above are used to define the degree of consensus of a group of users.

2.3 Degree of Consensus

Given a space $\mathbb{S}_n$, a finite non empty set $\Lambda = \{a_1, \ldots, a_N\}$ and a group of evaluators $\mathbb{E} = \{\alpha_1, \ldots, \alpha_M\}$, the *group evaluation* of Λ is considered as the pair $(\Lambda, \mathcal{Q}_{\mathbb{E}})$, where $\mathcal{Q}_{\mathbb{E}} = \{Q_i : \Lambda \to \mathbb{S}_n \mid i \in \{1, \cdots M\}\}$, and Q_i is the evaluation of α_i.

Let's suppose that there is consensus among the group, i.e., $\cap_{i=1}^{M} Q_i(a_t) \neq \emptyset \ \forall a_t \in \Lambda$. The next definition regarding the degree of consensus is from [Roselló *et al.*, 2010]:

Definition 3 *Given a group evaluation in consensus* $(\Lambda, \mathcal{Q}_{\mathbb{E}})$, *i.e.,* $\cap_{i=1}^{M} Q_i$ *exists, let* μ *be a normalized measure defined on* $\mathbb{S}_n$, *i.e., a measure such that* $\mu(?) = 1$ *and* $\overline{\mu}$ *a normalized measure defined on the set* Λ. *The* degree of consensus *among the group,* $\kappa(\mathcal{Q}_{\mathbb{E}})$, *is*

$$\kappa(\mathcal{Q}_{\mathbb{E}}) = \frac{H(\sqcup_{i=1}^{M} Q_i)}{H(\cap_{i=1}^{M} Q_i)} \tag{6}$$

where the entropy of a qualitative description Q is

$$H(Q) = \sum_{\mathcal{E} \in \mathbb{S}_n,\, \mu(\mathcal{E}) \neq 0} \overline{\mu}(Q^{-1}(\mathcal{E})) I(\mathcal{E}), \tag{7}$$

with $I(\mathcal{E}) = \log \frac{1}{\mu(\mathcal{E})}$.

The necessary and sufficient condition for which there exists consensus is $\cap_{i=1}^{M} Q_i(a_l) \neq \emptyset, \forall a_l \in \Lambda$. If this situation does not hold, then a process has to be initiated to obtain consensus. The algorithm presented in [Roselló *et al.*, 2010] is based on the following idea: If two people disagree on some fact and they want to reach an agreement, i.e., reach consensus, they have to reconsider their positions and find points in common. In this section this idea is formalized by using the concepts already given. It can be understood as a process of automatic negotiation.

Definition 4 *Given a space* $\mathbb{S}_n$ *with basic labels* $\mathcal{S} = \{B_1, \ldots, B_n\}$, *and a space* $\mathbb{S}_{n+1}$ *with basic labels* $\mathcal{S}' = \{B'_1, \ldots, B'_{n+1}\}$, *the* dive function *is the map* $\phi_0 : \mathbb{S}_n \to \mathbb{S}_{n+1}$ *defined as follows:*
For basic labels $B_i \in \mathbb{S}_n$, *then* $\phi_0(B_i) = [B'_i, B'_{i+1}]$, *and, for non-basic labels,* $\phi_0([B_i, B_j]) = \bigcup_{k=i}^{j} \phi_0(B_k) = [B'_i, B'_{j+1}]$.

With this function, each basic label in $\mathbb{S}_n$ is "split" into two new basic labels in $\mathbb{S}_{n+1}$. And in general, for each label $\mathcal{E}$ in $\mathbb{S}_n$, $\Phi_0(\mathcal{E})$ is obtained by adding a new basic label. In this same way, we can define $\phi_i : \mathbb{S}_{n+i} \to \mathbb{S}_{n+i+1}$, for $i \geq 1$, and the following chain can be considered:

$$\mathbb{S}_n \overset{\phi_0}{\hookrightarrow} \mathbb{S}_{n+1} \overset{\phi_1}{\hookrightarrow} \mathbb{S}_{n+2} \hookrightarrow \cdots \hookrightarrow \mathbb{S}_{n+m} \overset{\phi_m}{\hookrightarrow} \mathbb{S}_{n+m+1}$$

Then, given $\mathcal{E}, \mathcal{F} \in \mathbb{S}_n$ such that $\mathcal{E} \cap \mathcal{F} = \emptyset$, we can see that there exists a natural number $k \geq 1$ such that :

$$(\phi_{k-1} \circ \cdots \circ \phi_0)(\mathcal{E}) \cap (\phi_{k-1} \circ \cdots \circ \phi_0)(\mathcal{F}) \neq \emptyset.$$

Similarly and given $\mathcal{E}_1, \ldots, \mathcal{E}_M \in \mathbb{S}_n$ such that $\cap_{i=1}^{M} \mathcal{E}_i = \emptyset$, there exists $k \geq 1$ such that

$$\cap_{i=1}^{M} (\phi_{k-1} \circ \cdots \circ \phi_0)(\mathcal{E}_i) \neq \emptyset.$$

The next result allows us to extend the measure defined in $\mathbb{S}_n$ to the new space $\mathbb{S}_{n+1}$ (for a proof see [Roselló *et al.*, 2010]):
Let μ be a normalized measure defined on $\mathbb{S}_n$ and let us suppose that $\mathbb{S}_n$ is "dived" in $\mathbb{S}_{n+1}$. Then the measure μ can be extended to a normalized measure μ' in $\mathbb{S}_{n+1}$ defined, taking weights $0 < \lambda_1, \ldots, \lambda_n < 1$, in the following way:

$$\begin{aligned} \mu'(B'_1) &= (1-\lambda_1)\mu(B_1) \\ \mu'(B'_2) &= \lambda_1\mu(B_1) + (1-\lambda_2)\mu(B_2) \\ \mu'(B'_i) &= \lambda_{i-1}\mu(B_{i-1}) + (1-\lambda_i)\mu(B_i) \\ \mu'(B'_{n+1}) &= \lambda_n\mu(B_n) \end{aligned}$$

And for a non-basic label $\mathcal{E}' = [B'_i, B'_j] \in \mathbb{S}_{n+1}$,

$$\mu'(\mathcal{E}') = \sum_{k=i}^{j} \mu'(B'_k).$$

With the defined dive function and this extension of the measure, we can thus enact a process to reach consensus in a group evaluation $(\Lambda, \mathcal{Q}_{\mathbb{E}})$. The fully detailed process can be found in [Roselló *et al.*, 2010].

Now, we can calculate the degree of consensus κ within the group evaluation in which consensus has been obtained.

3 The Recommender System Algorithm

So far, we have introduced all the concepts needed to explain the system based on compatibility between users' profiles through the concept of consensus. This section is devoted to the explanation of the proposed RS.

Consider the recommendation process of a product described by a set of features Λ, where each feature can be described by an element of the space $\mathbb{S}_n$. Let $\mathbb{A}$ be the set of alternatives to be recommended. Let $\mathbb{E}$ be a set of users, which are the training set for the RS; each of these provide a qualitative description $Q_i : \Lambda \longrightarrow \mathbb{S}_n$, which assign a label of $\mathbb{S}_n$to each feature, and let $\mathcal{Q}_{\mathbb{E}} \subset \mathcal{Q}$ be the set of these qualitative descriptions. Finally, we have to assume that there exists a function $f : \mathcal{Q} \to \mathbb{A}$ that assigns each Q_i to an element of a set $\mathbb{A}$ of alternatives.

In Figure 1, we can see a representation of the training set, where each dot is an element $Q \in \mathcal{Q}_{\mathbb{E}}$ and the dotted closed lines express that these two groups of users are compatible, i.e. these two subsets of $\mathcal{Q}_{\mathbb{E}}$ are in consensus (of course each Q is in consensus with itself).

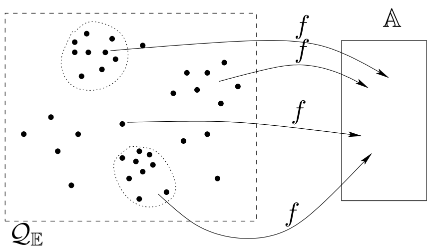

Figure 1: The training set of the recommender system.

The goal of the system is, for a new user with qualitative description $Q' : \Lambda \to \mathbb{S}_n$, to assign an alternative $f(Q') \in \mathbb{A}$.

Let us denote by $\mathcal{C}(Q')$ the set of the subsets of $\mathcal{Q}_{\mathbb{E}} \cup \{Q'\}$ that are in consensus, contain Q', and their cardinal is greater than or equal to 2. Let $i_{Q'}$ be its cardinal:

$$i_{Q'} = |\mathcal{C}(Q')|. \tag{8}$$

The main idea of the algorithm is that the best alternative for the new user with qualitative description Q' is the alternative of the user with nearest qualitative description to Q'. This can be done in the following steps:

1. First of all, the algorithm finds $\mathcal{C}(Q')$ and $i_{Q'}$.
2. If $i_{Q'} \geq 1$ then we choose a subset with highest degree of consensus (6):
$$C(Q')^* = \arg \max_{C_i(Q') \in \mathcal{C}(Q')} \kappa(C_i(Q')).$$
3. The next step is to assign to Q' the alternative corresponding to the user with qualitative description nearest to Q' in the subset $C(Q')^*$:
$$f(Q') = f(\arg \min_{Q \in C(Q')^*} d(Q, Q')),$$
where the distance is the expression in (5).
4. If in (8) $i_{Q'} = 0$, then we have to apply the automatic negotiation process introduced in Section 2.3, in order to find at least one subset in $\mathcal{C}(Q')$ and get an $i_{Q'} \geq 1$ (see Figure 2). Once it is found, the algorithm follows as in step 2.

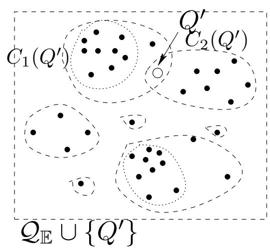

Figure 2: Here the automatic negotiation process has been applied in one step. This has produced two subsets $C_1(Q')$ and $C_2(Q')$.

4 An Experimental Case in Movie Recommendation

An experimental case is presented in this section for top-N movie recommendations. The RS presented, based on the algorithm given in Section 3, can be considered as a hybrid RS. Hybrid recommender systems combine collaborative and content-based methods [Adomavicius and Tuzhilim, 2005]; in our case content-based characteristics to define users' profiles are incorporated into the collaborative approach given. In particular, users' profiles are elaborated from their order-of-magnitude preferences on 18 pre-fixed movie genres (induced from their favorite movies). These content-based profiles are used to define qualitative descriptions, which involve different levels of precision, and allow us to find compatibility among users. The alternative assigned to a new user is a top-N movie recommendation list considering their movie genre preferences. The main goal of this section is to present and assess this movie RS by using a set of offline tests.

4.1 Data Set Description

A selection of a MovieLens data set, provided by the GroupLens Research at the University of Minnesota, is used to test the proposed hybrid system. In particular, the files used were the movies and ratings files, the files being structured as follows:

```
Movies(movieID, title, genres)
Ratings(userID, movieID, rating)
```

Movie ratings are considered in a 1-to-5 ordinal scale while movie genres are represented by a dynamic attribute list. There are 18 different attributes available as genres.

To define the data set for the test, a data subset from the Movielens dataset is selected in the following way. First, we selected films that have received between 10,000 and 15,000 ratings. Then, to avoid movies equally rated for everybody and sparsity-related problems, the next movies and users restrictions are considered:

- Those movies with a rating standard deviation below 1 are discarded;
- Users must have rated at least 85% of the movies.

As a result, a first data subset containing 200 users and 62 movies was considered in this experimental case.

4.2 Obtaining Users' Profiles

A crucial step for using the algorithm with the data of the database MovieLens is obtaining from this a qualitative profile of each user. This profile is obtained using the scores that users have given the films. The way of obtaining this profile is not unique (note that the algorithm presented *begins* with the users' qualitative description of the set Λ, usually using a specific interface with the computer system).

In this context, the set Λ contains the 18 available genres, and each feature in Λ will be described by an element of a space $\mathbb{S}_4$, with basic labels $B_1 = \texttt{"I hate it"}, B_2 = \texttt{"I don't like it"}, B_3 = \texttt{"I like it"}, B_4 = \texttt{"I love it"}$.

The qualitative description Q_{α_i} corresponding to each user $\alpha_i \in \mathbb{E}$ is obtained counting how many times this user has selected a favorite movie with genres $a_{i_1}, \ldots, a_{i_k}$. These 18 numbers are normalized and mapped to a label of the space $\mathbb{S}_4$ using the function $q : \mathbb{R} \to \mathbb{S}_4$:

$$q(x) = \begin{cases} B_1, & \text{if } x \in [0, 1/7) \\ [B_1 B_2], & \text{if } x \in [1/7, 2/7) \\ [B_1 B_3], & \text{if } x \in [2/7, 3/7) \\ ?, & \text{if } x \in [3/7, 4/7) \\ [B_2 B_4], & \text{if } x \in [4/7, 5/7) \\ [B_3 B_4], & \text{if } x \in [5/7, 6/7) \\ B_4, & \text{if } x \in [6/7, 1] \end{cases}$$

Finally, the profile of each user $\alpha_i \in \mathbb{E}$ is obtained as a vector of 18 components corresponding to their qualitative descriptions of the 18 genre preferences.

4.3 Experimental Methodology

To test the RS, we run a set of offline leave-one-out tests where our recommendations are compared to two well-known non-personalized models. Non-personalized recommenders present a predefined list of items to any user, regardless of their preferences. In this test, the models used are the Movie Average, where the top-N items with the highest average rating are recommended and the Top Popular, which recommends top-N items with the highest popularity (largest number of ratings).

For each user u from the data set, our RS, as explained in Sections 3 and 4.2, performs the four following steps:

1. Obtain the profile for user u from his set of preferred top-N movies;
2. Search user v, from the set of users with the highest degree of consensus that includes u, with a minimum-distance profile to that of user u;
3. Obtain the set of preferred top-N movies by user v;
4. Extract the common movies set M between the two sets of preferred movies for u and v.

Then, to test our methodology, we compute the following indicators for different values of N [Cremonesi *et al.*, 2010]:

- Coincidence percentage between preferred top-N movies by users and their recommendation for the three RSs to be compared.
- Rating difference between preferred top-N movies by user u and their recommendation for each movie m in M for Movie Average and the presented RS.

When all users have been tested, an average of the results is calculated, being the final result for an N items recommendation.

Figure 3 reports the performance of the recommender system algorithm presented versus non-personalized methods Movie Average and Top Popular. It compares their average of coincidence percentage for an N items recommendation following a leave-one-out test. For each value of the number of items to be recommended (horizontal axe) the averages of coincidence percentage are represented considering the three recommender systems (vertical axe). N values are natural numbers in the range from 1 to 10, since 10 is usually considered the maximum number of movies to be recommended.

Note that, as expected, results improve when N increases. Our RS performs significantly better than the non-personalized ones.

Figure 4 shows the average movie rating differences between the real values given by each user being tested and the values of the recommendations. Note that these differences can only be computed for our RS and the recommendations given by the Movie Average model, since Top Popular just select the most popular movies (largest number of ratings) and does not provide their specific ratings. For each value of the number of items to be recommended (horizontal axe) the

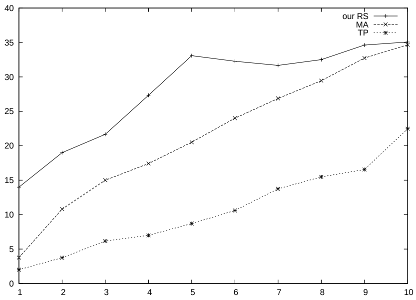

Figure 3: Coincidence percentage for three models.

average movie rating differences are represented considering the two recommender systems (vertical axe).

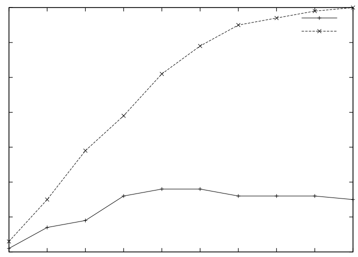

Figure 4: Average movie rating differences.

Figure 4 shows that our RS outperforms the non-personalized Movie Average in terms of rating recommendation difference.

5 Conclusions and Future Research

This paper presents a theoretical framework, which provides a new methodology for recommender systems based on group consensus theory. An experimental case for movie recommendations, based on the algorithm presented, is also described and assessed.

This work shows that the method presented for movie recommendations performs better than non-personalized models. Due to the high flexibility and adaptability of the method presented, we believe that its use in recommendation in environments where stating preferences involving different levels of precision can be very interesting.

Future work will be focused in three directions: First, from a theoretical point of view, defining other distance metrics among users and comparing results. Second, studying the analysis of the system implementation when recommending sensory products, where different levels of precision are required. And finally, comparing its performance with personalized models will complete the evaluation of the movie recommender system.

A web-based software tool for collecting and summarizing users' opinions and for working simultaneously with different levels of precision is being built by using the concepts presented in this paper. The software developed will be adapted to design a recommender system based on the methodology defined in this paper.

Acknowledgments

This work has been partly funded by MICINN (Spanish Ministry of Education and Science) SENSORIAL Research Project (TIN2010-20966-C02).

References

[Adomavicius and Tuzhilim, 2005] G. Adomavicius and A. Tuzhilim. Towards the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–742, June 2005.

[Aggarwal *et al.*, 1999] C. C. Aggarwal, J. L. Wolf, K-L. Wu, and K-L. Yu. Horting hatches an egg: A new graph-theoretic approach to collaborative filtering. In *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,*, August 1999.

[Birkhoff, 1967] G. Birkhoff. *Lattice Theory*. American Mathematical Society (Colloquium Publications), 1967.

[Breese *et al.*, 1998] J.S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *Proc. of the fourteenth Conference on Uncertainty in Artificial Intelligence*, July 1998.

[Cremonesi *et al.*, 2010] Paolo Cremonesi, Yehuda Koren, and Roberto Turrin. Performance of recommender algorithms on top-n recommendation tasks. In *Proceedings of the 4th ACM Conference on Recommender Systems*, September 2010.

[Dague, 1993] P. Dague. Numeric reasoning with relative orders of magnitude. In *AAAI Conference, Washington*, 1993.

[Forbus, 1996] K. Forbus. *Qualitative Reasoning*. CRC Hand-book of Computer Science and Engineering. CRC Press, 1996.

[Kalagnanam *et al.*, 1991] J. Kalagnanam, H.A. Simon, and Y. Iwasaki. The mathematical bases for qualitative reasoning. *IEEE Expert.*, 1991.

[Resnik *et al.*, 1994] P. N. Resnik, M. Iakovou, P. Sushak, and J. Riedl. Grouplens: An open architecture for collaborative filtering of netnews. In *Proc. of the 1994 Computer Supported Cooperative Work Conference*, 1994.

[Roselló *et al.*, 2010] L. Roselló, F. Prats, N. Agell, and M. Sánchez. Measuring consensus in group decisions by means of qualitative reasoning. *International Journal of Approximate Reasoning*, 51:441–452, January 2010.

[Sarwar *et al.*, 2001] B. Sarwar, G. Karypis, J Konstan, and J Riedl. Application of dimensionality reduction in recommender systems – a case study. In *Proc. of the ACM WebKDD Workshop*, 2001.

[Schardanand and Maes, 1995] U. Schardanand and P. Maes. Social information filtering: Algorithms for automating and 'word of mouth'. In *Proc. of the Conf. on Human Factors in Computing Systems*, 1995.

[Travé-Massuyès and Dague, 2003] L. Travé-Massuyès and P. Dague, editors. *Modèles et raisonnements qualitatifs*. 2003.