Control and Cybernetics

vol. **39** (2010) No. 4

An improved recommender system to avoid the persistent information overload in a university digital library*

by

C. Porcel[1], J.M. Morales del Castillo[2], M.J. Cobo[2], A.A. Ruíz[3], E. Herrera-Viedma[2]

[1]University of Jaén
Department of Computer Science, Jaén, Spain
[2]University of Granada
Department of Computer Science and A.I., Granada, Spain
[3]University of Granada
Department of Library Sciences, Granada, Spain
e-mail: cporcel@ujaen.es, josemdc@ugr.es, mjcobo@decsai.ugr.es, aangel@ugr.es, viedma@decsai.ugr.es

Abstract: Nowadays we are continuously bombarded with a lot of information, and because of it we have serious problems with accessing the relevant information, that is, we suffer from the information overload problems. Recommender systems have been applied successfully to avoid the information overload in different domains, but the number of electronic resources daily generated keeps growing and the problem rises again. Therefore, we find a persistent problem of information overload. In this paper we propose an improved recommender system to avoid the persistent information overload found in a University Digital Library. The idea is to include a memory to remember selected resources but not recommended to the user, and in such a way, the system could incorporate them in future recommendations to complete the set of filtered resources, for example, if there are a few resources to be recommended or if the user wishes output obtained by combination of resources selected in different recommendation rounds.

Keywords: recommender systems, information overload, university digital libraries, fuzzy linguistic modeling.

1. Introduction

Nowadays we live in the so called Information Society, in which we are bombarded with a huge amount of information in all fields of our lives. In this sense,

*Submitted: July 2010; Accepted: October 2010.

the problem of *information overload* is well known for all of us (Meghabghab and Kandel, 2008). However, in the literature we can find a lot of references indicating that the problem is not new, but that the information overload has existed for many years. In the recent years, though, the problem has become more widely recognized and experienced because of the rapid advances made in Information and Communication Technologies (Edmunds and Morris, 2000).

Although it is a very well known and analyzed problem, there is no consensus about the precise definition of information overload (Butcher, 1998). Information overload can be defined as *"The inability to extract needed knowledge from an immense quantity of information for one of many reasons"* (Nelson, 1994). Other sentence, which defines this topics is *"The volume of information on the Internet creates more problems than just trying to search an immense collection of data for a small and specific set of knowledge"* (Nelson, 1994). It can mean several things, such as having more relevant information than one can assimilate or it might mean being burdened with a large supply of unsolicited information, some of which may be relevant (Butcher, 1998; Meghabghab and Kandel, 2008). This great amount of information introduces noise in our information access processes and it makes finding relevant information difficult, affecting also our capacity of making decisions. For example, every day we receive in our accounts a huge amount of emails. Most of them are qualified as spam, but we also receive too a big number of emails containing useful or relevant information. The problem is that due to this fact we may pay inadequate attention to what we think is of minor importance and so misinterpret the message, or we could lose some information thinking that it is not important.

This problem appears specially in the environment of digital libraries where the information is generated much faster than the users can process it. *Digital libraries* are information collections that have associated services delivered to user communities using a variety of technologies. The information collections can be scientific, business or personal data, and can be represented as digital text, image, audio, video, or other media. This information can be digitalized paper or born digital material and the services offered based upon such information can be varied and can be offered to individuals or user communities (Callan et al., 2003; Gonçalves, Fox, Watson and Kipp, 2004; Renda and Straccia, 2005). Libraries offer different types of reference and referral services (e.g., ready reference, exhaustive search, selective dissemination of information), instructional services (e.g., bibliographic instruction, database searching), value added services (e.g., bibliography preparation, language translation) and promotional services (e.g., literacy, freedom of expression). Digital libraries have been applied in a lot of contexts but in this paper we focus on an academic environment. University Digital Libraries (UDL) provide information resources and services to students, faculty and staff in an environment that supports learning, teaching and research (Chao, 2002). A service that is particularly important is the selective dissemination of information or filtering (Morales del Castillo, Pedraza-Jiménez, Ruíz, Peis and Herrera-Viedma, 2009). Users develop interest profiles and as new materials are added to the collection, they are compared to the profiles and the UDL alerts the users with relevant items (Marchionini, 2000).

As we have seen, because of information overload problem, although there is an abundance of information available, it is often difficult to obtain useful or relevant information when it is necessary. When the users of a UDL try to receive useful information, they often obtain irrelevant information or information which is not currently necessary for them. With the expansion of Web, users need easier access to the thousands of resources that are available but yet hard to find (Meghabghab and Kandel, 2008). One solution, very widespread in many environments, meant to reduce information overload, is the use of *recommender systems* (Hanani, Shapira and Shoval, 2001; Reisnick and Varian, 1997; Symeonidis, Nanopoulos, Papadopoulos and Manolopoulos, 2008). A recommender system is an information filtering system that attempts to present information items (movies, music, books, news, images, web pages, papers etc.) that are likely of interest to a user. Recommender systems are especially useful when they identify information a person was previously unaware of. From the theoretical point of view, recommender systems fall into two main categories (Cornelis, Lu, Guo and Zhang, 2007; Hanani, Shapira and Shoval, 2001; Reisnick and Varian, 1997; Symeonidis, Nanopoulos, Papadopoulos and Manolopoulos, 2008):

1. *Content-based recommender systems* recommend information items to a user by means of a process based on the content of the information item and the user's past experience in dealing with similar items, and therefore, ignoring data from other users.
2. *Collaborative recommender systems* recommend information items to a user by means of a process based on the user's social environment and ignoring the content of item, that is, the recommendations to a user are based on other user recommendations with similar user profiles.

In a digital library the collaborative filtering approach is very useful because it allows users to share their experiences, that is, users can rate or add value to information objects and these ratings can be shared with the community, so that popular items can be easily located or people may receive information items found useful by others with similar profiles (Marchionini, 2000; Ross and Sennyey, 2008).

Specifically, we have applied successfully recommender systems to UDL to disseminate relevant information in the systems previously proposed in Porcel, Moreno and Herrera-Viedma (2009) and Porcel and Herrera-Viedma (2010). The experiments and evaluation performed in such systems reveals us that both proposals are satisfactory for the users. However, despite that the use of these techniques to avoid the information overload problem was successful, the number of electronic resources daily generated grows continuously and the problem appears again. Therefore, we are faced with a persistent problem of information overload.

In this paper we propose an improved recommender system based on memory to avoid the persistent information overload found in those systems proposed in Porcel, Moreno and Herrera-Viedma (2009) and Porcel and Herrera-Viedma (2010). We define this recommender system in a multi-granular fuzzy linguistic context (Chang, Wang and Wang, 2007; Herrera and Martínez, 2001; Herrera-Viedma, Cordón, Luque, López and Muñoz, 2003; Herrera-Viedma, Martínez, Mata and Chiclana, 2005; Mata, Martnez and Herrera-Viedma, 2009). In such a way, we incorporate in the recommender system flexible tools to handle the information by allowing for representation of different concepts of the system with different linguistic label sets. It works according to the recommendation approach defined in (Porcel and Herrera-Viedma, 2010), but in this case, it presents a memory to remember selected items but not recommended previously, and in such a way, the system could incorporate them in future recommendations to complete the set of recommendations. This may take place, for example, if there are a few items to be recommended or if the user wishes outputs obtained by combination of items selected in different recommendation rounds. To do so, we ask users to express restrictions on the quantity of items to receive in each recommendation round and about the novelty of such items. As in Porcel, Moreno and Herrera-Viedma (2009) and Porcel and Herrera-Viedma (2010) the recommender system is able to recommend both research resources and collaboration possibilities to aid UDL users to meet other researchers of related areas which could be potential collaborators in research projects.

The paper is structured as follows. Section 2 presents the background necessary for the development of the paper. In Section 3 we present the new improved recommender system based on memory. Section 4 reports the system evaluation and the experimental results. Finally, our concluding remarks are provided.

2. Background

2.1. Basics of recommender systems

Recommender systems help online users in the effective identification of items suiting their wishes, needs or preferences. They have the effect of guiding the user in a personalized way to relevant or useful objects in a large space of possible options (Burke, 2007). These applications improve the information access processes for users not having a detailed product domain knowledge. They are becoming popular tools for reducing information overload and to improve the sales in e-commerce web sites (Burke, 2007; Cao and Li, 2007; Duen-Ren, Chin-Hui and Wang-Jung, 2009; Reisnick and Varian, 1997).

Automatic filtering services differ from retrieval services in that in filtering the corpus changes continuously, the users have long time information needs (described by mean of user profiles) instead of introducing an ad hoc a query into the system, and their objective is to remove irrelevant data from incoming streams of data items (Hanani, Shapira and Shoval, 2001; Marchionini, 2000;

Reisnick and Varian, 1997). A result from a recommender system is understood as a recommendation, an option worth consideration; a result from an information retrieval system is interpreted as a match to the user's query (Burke, 2007).

In order to generate personalized recommendations that are tailored to the user's preferences or needs, recommender systems must collect personal preference information, such as user's history of purchase, items previously interesting for the user, click-stream data, demographic information, and so on. Two different ways for obtaining information about user preferences are distinguished (Hanani, Shapira and Shoval, 2001), although many systems adopt a hybrid approach:

- The *implicit approach* is implemented by inference from some kind of observation. The observation is applied to user behavior or to detecting a user's environment (such as bookmarks or visited URL). The user preferences are updated by detecting changes while observing the user.
- The *explicit approach*, interacts with the users by acquiring feedback on information that is filtered, that is, the users express some specifications of what they desire. This approach is the most common.

There are two main approaches that have been proposed for the implementation of recommender applications (Cornelis, Lu, Guo and Zhang, 2007; Hanani, Shapira and Shoval, 2001; Reisnick and Varian, 1997; Symeonidis, Nanopoulos, Papadopoulos and Manolopoulos, 2008; Yang and Li, 2009):

- *Content-based systems*: They generate the recommendations taking into account the characteristics used to represent the items and the ratings that a user has given to them (Hanani, Shapira and Shoval, 2001; Reisnick and Varian, 1997). These recommender systems tend to fail at the beginning, when the users are provided few ratings.
- *Collaborative systems*: The system generates recommendations using explicit or implicit preferences from many users, ignoring the items representation. Collaborative systems locate peer users with a rating history similar to the current user and they generate recommendations using this neighborhood. These recommender systems tend to fail when little is known about items, i.e., when new items appear, this is called the new item cold-start problem (Bobadilla, Serradilla, Hernando, and MoviLens, 2009; Bobadilla, Serradilla, and Bernal, 2010; Burke, 2007; Leung, Chan, and Chung, 2008).

On the other hand, in Iskold (2007) different real recommendation engines were analyzed and, from a practical point of view, four different types of recommendations were identified:

1. *Personalized recommendation:* Recommend items based on the individual's past behavior, as in the content-based filtering.

2. *Social recommendation:* Recommend items based on the past behavior of similar users, as in the collaborative filtering.
3. *Item recommendation:* Recommend items based on the item itself, as it happens in the information retrieval systems (Korfhage, 1997) but assuming long time queries.
4. A combination of the three approaches above.

In this paper we propose the use of a hybrid approach to alleviate the disadvantages of each one of them and to exploit their benefits; using a hybrid strategy users are provided with recommendations more accurate than those offered by each strategy individually (Burke, 2007; Hanani, Shapira and Shoval, 2001). We focus on content-based and collaborative recommender systems. In these kind of systems, the users' information on preferences can be used to define user profiles that are applied as filters to streams of documents. The construction of accurate profiles is a key task and the system success will depend to a large extent on the ability of the learned profiles to represent the user's preferences (Quiroga and Mostafa, 2002).

The recommendation activity is followed by a relevance feedback phase. Relevance feedback is a cyclic process whereby the users provide the system with their satisfaction evaluations as to the recommended items and the system uses these evaluations to automatically update user profiles in order to generate new recommendations (Hanani, Shapira and Shoval, 2001; Reisnick and Varian, 1997).

2.2. The 2-tuple fuzzy linguistic approach

The fuzzy linguistic modeling (FLM) is a tool based on the concept of *linguistic variable* (Zadeh, 1975) which has given very good results for modeling qualitative information in many problems, e.g., in decision making (Cabrerizo, Alonso and Herrera-Viedma, 2009; Herrera and Herrera-Viedma, 2000; Herrera, Herrera-Viedma and Verdegay, 1997), quality evaluation (Herrera-Viedma, Pasi, López-Herrera and Porcel, 2006; Herrera-Viedma and Peis, 2003) or models of information retrieval (Herrera-Viedma, 2001; Herrera-Viedma and López-Herrera, 2007; Herrera-Viedma, López-Herrera, Luque and Porcel 2007).

The 2-tuple FLM (Herrera and Martínez, 2000) is a continuous model of representation of information allowing for a reduction of the loss of information typical of other fuzzy linguistic approaches (classical and ordinal, see Zadeh, 1975).

Let $S = \{s_0, ..., s_g\}$ be a linguistic term set with odd cardinality, where the mid term represents an indifference value and the rest of the terms are symmetrically related to it. We assume that the semantics of labels is given by means of triangular membership functions and consider all terms distributed on a scale on which a total order is defined, $s_i \leq s_j \iff i \leq j$. In this fuzzy linguistic context, if a symbolic method aggregating linguistic information obtains a value $\beta \in [0, g]$, and $\beta \notin \{0, ..., g\}$, then an approximation function is used to express the result in S. β is represented by means of 2-tuples (s_i, α_i), $s_i \in S$ and $\alpha_i \in [-.5, .5)$ where s_i represents the linguistic label of the information, and α_i is a numerical value expressing the value of the translation from the original result β to the closest index label, i, in the linguistic term set ($s_i \in S$).

For example, let $\mathcal{S} = \{s_0, s_1, s_2, s_3, s_4, s_5, s_6\}$ be our linguistic term set, we could represent the value $\beta = 2.8$ as the 2-tuple $\Delta(\beta) = (s_3, -0.2)$.

This 2-tuple representation model defines a set of transformation functions between numeric values and 2-tuples $\Delta(\beta) = (s_i, \alpha)$ and $\Delta^{-1}(s_i, \alpha) = \beta \in [0, g]$ (Herrera and Martínez, 2000).

The computational model is defined by establishing a negation operator, comparison of 2-tuples and aggregation operators (Herrera and Martínez, 2000; Torra and Narukawa, 2007). Using functions Δ and Δ^{-1} that transform without loss of information numerical values into linguistic 2-tuples and vice versa, any of the existing aggregation operators can be easily extended for dealing with linguistic 2-tuples. Some examples are:

Definition 1 *Arithmetic Mean. Let* $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ *be a set of linguistic 2-tuples, the 2-tuple arithmetic mean* $\overline{x}^e$ *is computed as,*

$$\overline{x}^e[(r_1, \alpha_1), \ldots, (r_n, \alpha_n)] = \Delta(\sum_{i=1}^{n} \frac{1}{n} \Delta^{-1}(r_i, \alpha_i)) = \Delta(\frac{1}{n} \sum_{i=1}^{n} \beta_i). \tag{1}$$

Definition 2 *Weighted Average Operator. Let* $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ *be a set of linguistic 2-tuples and* $W = \{w_1, ..., w_n\}$ *be their associated weights. The 2-tuple weighted average* $\overline{x}^w$ *is:*

$$\overline{x}^w[(r_1, \alpha_1), \ldots, (r_n, \alpha_n)] = \Delta(\frac{\sum_{i=1}^{n} \Delta^{-1}(r_i, \alpha_i) \cdot w_i}{\sum_{i=1}^{n} w_i}) = \Delta(\frac{\sum_{i=1}^{n} \beta_i \cdot w_i}{\sum_{i=1}^{n} w_i}). \tag{2}$$

Definition 3 *Linguistic Weighted Average Operator. Let* $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ *be a set of linguistic 2-tuples and* $W = \{(w_1, \alpha_1^w), ..., (w_n, \alpha_n^w)\}$ *be their linguistic 2-tuple associated weights. The 2-tuple linguistic weighted average* $\overline{x}_l^w$ *is:*

$$\overline{x}_l^w[((r_1, \alpha_1), (w_1, \alpha_1^w))...((r_n, \alpha_n), (w_n, \alpha_n^w))] = \Delta(\frac{\sum_{i=1}^{n} \beta_i \cdot \beta_{W_i}}{\sum_{i=1}^{n} \beta_{W_i}}), \tag{3}$$

with $\beta_i = \Delta^{-1}(r_i, \alpha_i)$ *and* $\beta_{W_i} = \Delta^{-1}(w_i, \alpha_i^w)$.

In any fuzzy linguistic approach, an important parameter to determine is the "granularity of uncertainty", i.e., the cardinality of the linguistic term set S. When different experts have different uncertainty degrees with respect to a phenomenon or when an expert has to assess different concepts, then several linguistic term sets with different granularity of uncertainty are necessary (Herrera and Martínez, 2001; Herrera-Viedma, Martínez, Mata and Chiclana, 2005;

Mata, Martínez and Herrera-Viedma, 2009). In Herrera and Martínez (2001) a multi-granular 2-tuple FLM based on the concept of linguistic hierarchy was proposed.

A *Linguistic Hierarchy, LH,* is a set of levels *l(t,n(t))*, where each level t is a linguistic term set with granularity $n(t)$ different from that of the remaining of levels of the hierarchy. The levels are ordered according to their granularity, i.e., a level $t+1$ provides a linguistic refinement of the previous level t. We can define a level based on its predecessor level as: $l(t, n(t)) \rightarrow l(t+1, 2 \cdot n(t) - 1)$.

A graphical example of a linguistic hierarchy is shown in Fig. 1.

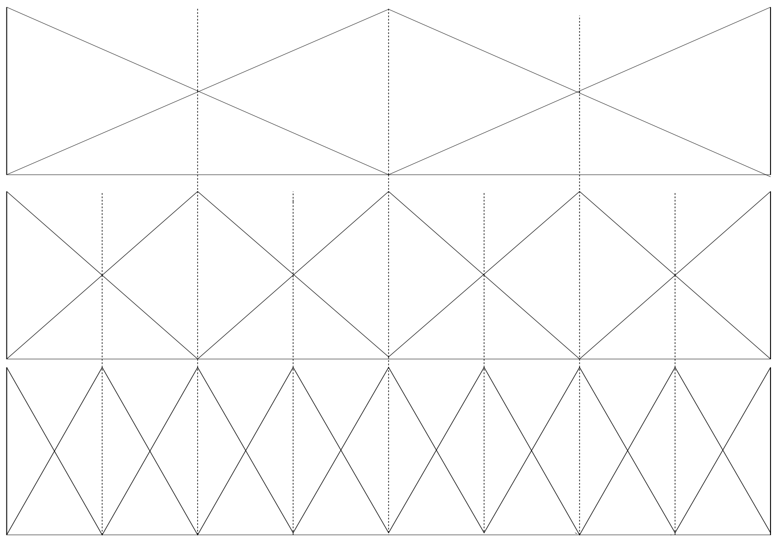

Figure 1. Linguistic hierarchy of 3, 5 and 9 labels

In Herrera and Martínez (2001) a family of transformation functions between labels from different levels was introduced:

Definition 4 *Let* $LH = \bigcup_t l(t, n(t))$ *be a linguistic hierarchy whose linguistic term sets are denoted as* $S^{n(t)} = \{s_0^{n(t)}, ..., s_{n(t)-1}^{n(t)}\}$. *The transformation function between a 2-tuple that belongs to level t and another 2-tuple in level* $t' \neq t$ *is defined as:*

$$TF_{t'}^{t} : l(t, n(t)) \longrightarrow l(t', n(t'))$$

$$TF_{t'}^{t}(s_i^{n(t)}, \alpha^{n(t)}) = \Delta\left(\frac{\Delta^{-1}(s_i^{n(t)}, \alpha^{n(t)}) \cdot (n(t') - 1)}{n(t) - 1}\right).$$

As it was pointed out in Herrera and Martínez (2001) this family of transformation functions is bijective. This result guarantees that the transformations between levels of a linguistic hierarchy are carried out without loss of information.

2.3. Incomplete fuzzy preference relations

DEFINITION 5 *A fuzzy preference relation P on a set of alternatives $X = \{x_1, .., x_n\}$ is a fuzzy set on the product set $X \times X$, i.e., it is characterized by a membership function $\mu_P \colon X \times X \longrightarrow [0, 1]$.*

When cardinality of X is small, the preference relation may be conveniently represented by the $n \times n$ matrix $P = (p_{ij})$, being $p_{ij} = \mu_P(x_i, x_j)$ $(\forall i, j \in \{1, \dots, n\})$ interpreted as the preference degree or intensity of the alternative x_i over x_j, where:

- $p_{ij} = 1/2$ indicates indifference between x_i and x_j,
- $p_{ij} = 1$ indicates that x_i is absolutely preferred to x_j,
- and $p_{ij} > 1/2$ indicates that x_i is preferred to x_j.

However, as we have mentioned, our system integrates the multi-granular FLM based on 2-tuples, so we must define a linguistic preference relation as follows:

DEFINITION 6 *Let $X = \{x_1, .., x_n\}$ be a set of alternatives and S a linguistic term set. A linguistic preference relation $P = p_{ij}(\forall i, j \in \{1, \dots, n\})$ on X is:*

$$\mu_P : X \times X \longrightarrow S \times [0.5, 0.5) \tag{4}$$

where $p_{ij} = \mu_P(x_i, x_j)$ is a 2-tuple which denotes the preference degree of alternative x_i regarding to x_j.

As aforementioned, in many real world Group Decision Making (GDM) problems the experts are often not able to provide all the preference values that are required. In order to model these situations, we use incomplete fuzzy preference relations (Alonso, Chiclana, Herrera, Herrera-Viedma, Alcalá-Fdez and Porcel, 2008; Herrera-Viedma, Alonso, Chiclana and Herrera, 2007; Herrera-Viedma, Chiclana, Herrera and Alonso, 2007; Martínez, Pérez, Barranco and Espinilla, 2008).

DEFINITION 7 *A function $f \colon X \longrightarrow Y$ is* partial *when not every element in the set X necessarily maps onto an element in the set Y. When every element from the set X maps onto one element of the set Y, then we have a* total *function.*

DEFINITION 8 *A two-tuple fuzzy linguistic preference relation P on a set of alternatives X with a partial membership function is an* incomplete two-tuple fuzzy linguistic preference relation*.*

3. An improved recommender system based in memory

In this section, we present our proposal consisting of adding a second selection over the resources selected in a previous recommendation process. The system allows for avoiding the persistent information overload continuously present in a University Digital Library (UDL). The UDL staff manages and spreads a lot of knowledge about research resources such as electronic books, electronic papers, electronic journals, official dailies and so on (Callan et al., 2003; Renda and Straccia, 2005). Nowadays, this amount of information is growing and there is a need of automated tools to filter and spread that information to the users in a simple and timely manner.

Recommender systems have been applied successfully in UDL to disseminate relevant information (Porcel, Moreno and Herrera-Viedma, 2009; Porcel and Herrera-Viedma, 2010). However, despite the fact that the use of these techniques to avoid the information overload problem was successful, a huge amount of electronic resources are daily available in the UDL, so the problem appears again. As it was pointed out at the beginning, we propose a new recommender system to overcome this drawback by implementing a second filter which allows for a major reduction of the overload. This selection is made taking into account the restrictions defined by the users about the amount of resources that they want to receive and about the novelty of these resources. Some users only want to receive information about new items, but other might want to receive information about previous items considered relevant but not recommended to them.

Usually the amount of resources that might be recommended by the system is greater than the amount of resources that users want to receive. The idea is to keep all these resources considered relevant by the system but not recommended to the users, due to the restrictions imposed by themselves. These resources might be remembered for the later recommendation rounds. For example, when the amount of recommended resources at a given moment is not sufficient to satisfy the users' restrictions.

Moreover, the system recommends collaboration possibilities by indicating other researchers of related areas, who could collaborate on projects or own interest works. In such a way, this new recommender system improves the services that a UDL provides to the users, because it allows them to express some restrictions about the resources they want to receive and it is easier to obtain the knowledge about the users. It allows the system to decrease the time cost to establish the user profiles.

This hybrid recommender system is based on the multi-granular fuzzy linguistic approach presented in Section 2.2. So, in order to allow a higher flexibility in the communication processes between users and the recommender system, we use different label sets $(S_1,\ S_2, ...)$ to represent the different concepts to be assessed in its filtering activity. These label sets S_i are chosen from those label sets that compose an LH, i.e., $S_i \in LH$. We should point out that the number of different label sets that we can use is limited by the number of levels of LH, and therefore, in many cases the label sets S_i and S_j can be associated to a same label set of LH but with different interpretations depending on the concept to be modeled. We consider five concepts that can be assessed in the activity of this recommender system:

- **Importance degree** of a discipline with respect to a resource scope or user preferences (S_1).
- **Relevance degree** of a resource for a user (S_2).
- **Compatibility degree** between two users (S_3).
- **Preference degree** of a resource regarding another one (S_4).
- **Qualitative number of resources**. The users can specify a label indicating a rough value about the number of resources the want to receive (S_5).

Following the linguistic hierarchy shown in Fig. 1, we use the level 2 (5 labels) to assign importance and preference degrees ($S_1 = S^5$ and $S_4 = S^5$), the level 3 (9 labels) to assign relevance and compatibility degrees ($S_2 = S^9$ and $S_3 = S^9$), and the level 1 (3 labels) to represent the qualitative number of resources which the users want to receive ($S_5 = S^3$). As the importance degrees are provided by library staff, we use a set of 5 labels to facilitate them the characterization of resource scopes or user interest topics. The preference degrees are provided by the user, so we consider it suitable to use a set of 5 labels. On the other hand, as the relevance and compatibility degrees are computed automatically by the system, we use the set of 9 labels to have an adequate granularity level to represent the results. Finally, as the qualitative number of resources is established by the users we consider it sufficient to use the set of 3 labels. Using this LH, the linguistic terms in each level are the following:

- $S^3 = \{a_0 = Null = N,\ a_1 = Medium = M,\ a_2 = Total = T\}$,
- $S^5 = \{b_0 = Null = N,\ b_1 = Low = L,\ b_2 = Medium = M,\ b_3 = High = H,\ b_4 = Total = T\}$,
- $S^9 = \{c_0 = Null = N,\ c_1 = Very\_Low = VL,\ c_2 = Low = L,\ c_3 = More\_Less\_Low = MLL,\ c_4 = Medium = M,\ c_5 = More\_Less\_High = MLH,\ c_6 = High = H,\ c_7 = Very\_High = VH,\ c_8 = Total = T\}$.

In Fig. 2 we show the basic operating scheme of the recommender system, which is based on four main components:

1. *Resource representation.* The system obtains an internal representation of resources based on their scope.
2. *User profile representation.* The system obtains an internal representation of the users based on their research area, topics of interest and the restrictions imposed by themselves.

3. *Recommendation process.* The system generates the recommendations according to the filtering approach based on memory.
4. *Feedback phase.* The users provide to the system their opinions about the received recommendations.

In the following we explain these components in detail.

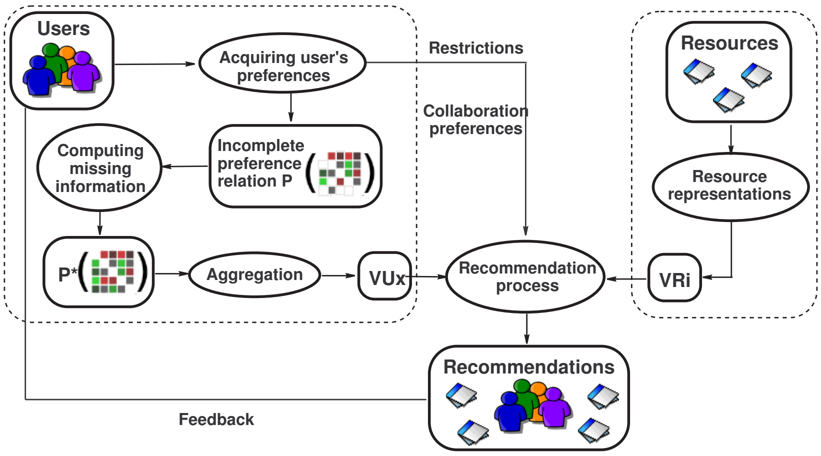

Figure 2. Basic operating scheme

3.1. Resource representation

The resources considered in our system are papers in journals, conference contributions, contributions to book chapters, books or edited books. Once the library staff insert all the available information about a new resource, the system obtains an internal representation mainly based on the resource scope. We use the *vector model* (Korfhage, 1997) to represent the resource scope and a classification composed by 25 disciplines (see Fig. 3), i.e., a research resource i is represented as

$$VR_i = (VR_{i1}, VR_{i2}, ..., VR_{i25}),$$

where each component $VR_{ij} \in S_1$ is a linguistic assessment that represents the importance degree of the discipline j with regard to the scope of i. These importance degrees are assigned by the library staff when they add a new resource.

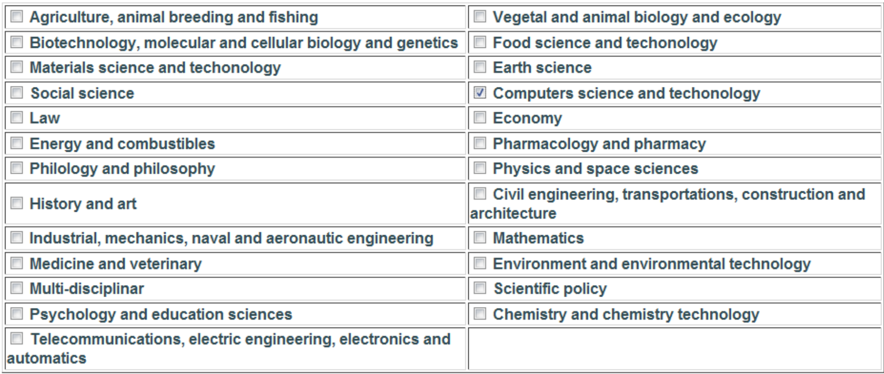

Figure 3. Interface for defining the disciplines of the resource scope

3.2. User profiles

User profiles are composed of three kinds of user preferences:

1. restrictions about the recommendations they want to receive,
2. user preferences on topics of interest, and
3. user preferences on collaboration possibility with other users.

The users can specify two kinds of **restrictions** about their preferences on the recommendations they want to receive:

1. Restrictions, which indicate the number of recommended items that they want to receive. The system enables specifiying this number in two manners:
   - *Quantitative.* The users indicate the number of resources using an exact value. For example, if they want to receive only X resources, and the number of selected resources to be recommended is bigger than X, the system only recommends the X more relevant and it remembers the rest of resources (not recommended) for later recommendation rounds.
   - *Qualitative.* The users specify a label indicating a rough value about the number of resources they want to receive. They select this value using a label of a term set S_5. Then, the users can specify if they want few, a medium number, or a lot of resources.
2. Restrictions, which indicate the kind of resources that they want to receive. There are users who only want to receive recommendations about resources just inserted into the system. But, also there are users who do care about receiving recommendations on resources previously inserted in the system, but still with validity. These resources could be more interesting than a new resource. In the second case, the users choose to receive recommendations about a set of resources composed by the new and remembered ones.

Now we explain how the users provide their **preferences on topics of interest** using the scheme proposed in Porcel and Herrera-Viedma (2010). We represent such user preferences using the vector model (Korfhage, 1997). We ask users to provide their preferences on some research resources, usually a limited number of resources, four or five. The choice of research resources is made by the staff taking into account the relevance supplied by the users. As in Martínez, Pérez, Barranco and Espinilla (2008) we suggest that users represent their preferences by means of incomplete fuzzy linguistic preference relations. Then, the system presents to users only a selection of the most representative resources, and the users provide their preferences about these resources by means of an incomplete fuzzy preference relation. Furthermore, according to the results presented in Alonso, Chiclana, Herrera, Herrera-Viedma, Alcalá-Fdez and Porcel (2008), it is enough that the users provide only a row of the preference relation. Then, we use the method proposed in Alonso, Chiclana, Herrera, Herrera-Viedma, Alcalá-Fdez and Porcel (2008) to complete the relations. Other possibilities to deal with missing information are studied in Alonso, Herrera-Viedma, Chiclana and Herrera (2009).

This method is an iterative procedure to estimate the missing values of incomplete preference relations, using only the preference values provided by a particular expert. By doing this, we assure that the reconstruction of the incomplete preference relation is compatible with the rest of the information provided by that expert. The procedure is guided by the consistency property and only uses known preference values. In particular, it is based on the additive consistency property of a fuzzy preference relation. Once the system completes the fuzzy linguistic preference relation provided by the user, it is possible to obtain a vector representing the user preferences on the topics of interest. Next, we explain this process in detail:

1. **Acquiring the user preferences on a limited number of research resources**: At the beginning, the main goal is to help the users to provide their preferences, while ensuring that these preferences are as consistent as possible. The system shows to the users the five most representative resources, $R = \{r_1, .., r_5\}$, and asks them to express their preferences by means of an incomplete fuzzy linguistic preference relation (see Fig. 4).

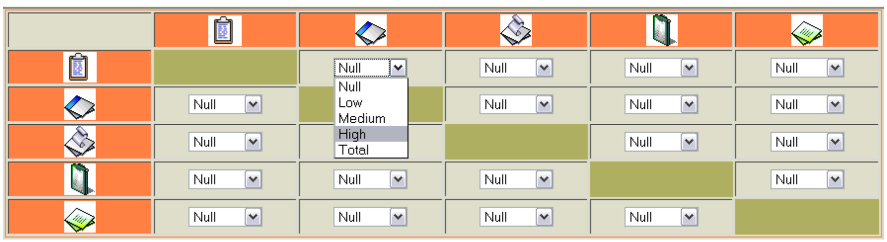

Figure 4. Interface for defining user preferences

The users only fill those preferences that they wish, assigning labels of S_4. In the preference relation, each preference value p_{ij} represents the linguistic preference degree of resource i over the resource j according to the user feeling. As aforementioned, the simplest case would be to provide a relation with only one row of preference values:

$$P = \begin{pmatrix} - & p_{12} & p_{13} & p_{14} & p_{15} \\ x & - & x & x & x \\ x & x & - & x & x \\ x & x & x & - & x \\ x & x & x & x & - \end{pmatrix}. \tag{5}$$

Then, the system completes the preference relation P using the method proposed in Alonso, Chiclana, Herrera, Herrera-Viedma, Alcalá-Fdez and Porcel (2008), and obtains the relation P^*:

$$P^* = \begin{pmatrix} - & p_{12} & p_{13} & p_{14} & p_{15} \\ p^*_{21} & - & p^*_{23} & p^*_{24} & p^*_{25} \\ p^*_{31} & p^*_{32} & - & p^*_{34} & p^*_{35} \\ p^*_{41} & p^*_{42} & p^*_{43} & - & p^*_{45} \\ p^*_{51} & p^*_{52} & p^*_{53} & p^*_{54} & - \end{pmatrix} \tag{6}$$

where $p_{1j} \in S_4$ are the degrees inserted by the user concerning the preferences of the resource x_1 with respect to x_j, p_{ii} represents indifference, and each p^*_{ij} is the estimated degree for the user about his/her preference of the resource x_i with respect to x_j.

2. In order to obtain **user preferences on topic of interest, i.e., the user preference vector**, firstly we calculate the user preference degrees on each considered resource according to the preference relation P^*, and secondly, we use this preference degrees together with the vectors that represent each research resource to obtain the user preference vector. To obtain them we propose the application of the arithmetic mean $\overline{x}^e$ (Definition 1) because in this way we can aggregate all the degrees with similar importance. Then, the preference degree of the resource i for the expert called DG_i, is computed as follows:

$$DG_i = \overline{x}^e[p^*_{i1}, \ldots, p^*_{i5}]. \tag{7}$$

Then, to obtain the user preference vector x, i.e. $VU_x = (VU_{x1}, VU_{x2}, ..., VU_{x25})$, we aggregate the vectors that represent the characteristics of the chosen research resources, i.e. $\{VR_1, ..., VR_5\}$, weighted by means of the user preference degrees $\{DG_1, \ldots, DG_5\}$. To do that, we use the linguistic weighted average operator defined in Definition 3, and then each position $k = \{1, \ldots, 25\}$ of the vector VU_x, is computed as follows:

$$VU_{xk} = \overline{x}_l^w[(VR_{1k}, DG_1), \ldots, (VR_{5k}, DG_5)]. \tag{8}$$

On the other hand, to complete the user profile, the system asks every user to express his/her **collaboration preferences**, i.e. if he/she wants to receive recommendations on collaboration possibilities with other users. This could help users to develop multi-disciplinary studies or participate in collaborative research projects (Porcel, Moreno and Herrera-Viedma, 2009). Users should respond to this question with "Yes" or "No".

3.3. Memory based recommendation strategy

The proposed recommender system based on memory, meant to avoid the persistent information overload works in two phases:

1. To generate the recommendations using a hybrid recommendation approach as in Porcel and Herrera-Viedma (2010).
2. To apply a second filter or selection process according to the user's restrictions.

3.3.1. Phase 1. Recommendation generation process

In this phase the system generates the recommendations to deliver the information resources to the respective users. This process is based on a matching process developed between user profiles and resource representations (Hanani, Shapira and Shoval, 2001; Korfhage, 1997). To do that, we can use different kinds of similarity measures, such as Euclidean Distance or Cosine Measure. Particularly, we use the standard cosine measure (Korfhage, 1997), because we represent both the resources and the users as vectors, and this measure uses the cosine of the angle between them. That is, it focuses on the content of a resource, not on its extension. So, with this measure, a paper about a particular topic is considered similar to a book about the same topic. As the components of the vectors used to represent user profiles and research resources are 2-tuple linguistic values, then we define the cosine measure in a 2-tuple linguistic context. Given two vectors of 2-tuple linguistic values,

$$V_1 = ((v_{11}, \alpha_{v11}), (v_{12}, \alpha_{v12}), \ldots, (v_{125}, \alpha_{v125}))$$

and

$$V_2 = ((v_{21}, \alpha_{v21}), (v_{22}, \alpha_{v22}), \ldots, (v_{225}, \alpha_{v225}))$$

the linguistic similarity between both, called $\sigma_l(V_1, V_2) \in S_1$, is defined as:

$$\sigma_l(V_1, V_2) = \Delta(g \times \frac{\sum_{k=1}^{25}(\Delta^{-1}(v_{1k}, \alpha_{v1k}) \times \Delta^{-1}(v_{2k}, \alpha_{v2k}))}{\sqrt{\sum_{k=1}^{25}(\Delta^{-1}(v_{1k}, \alpha_{v1k}))^2} \times \sqrt{\sum_{k=1}^{25}(\Delta^{-1}(v_{2k}, \alpha_{v2k}))^2}}). \tag{9}$$

where g is the granularity of S_1 and (v_{ik}, α_{vik}) is the 2-tuple linguistic value of term k in the vector(V_i).

When a new resource i is inserted into the system, we calculate the linguistic similarity measures, $\sigma_l(VR_i, VU_j)$, between the representation vector of this new resource (VR_i) and all the user preference vectors, $\{VU_1, \ldots, VU_m\}$, where m is the number of users in the system. These user preference vectors are obtained as shown in Section 3.2.

Then, if $\sigma_l(VR_i, VU_j) \geq \psi$, the user j is selected to receive recommendations about resource i. Previously, we have defined a linguistic threshold value (ψ) to filter the output of the system. Next, the system applies to each $\sigma_l(VR_i, VU_j)$ the transformation function from Definition 4, to obtain the relevance degree of the resource i for the user j, expressed using a label of the set S_2.

The collaboration preferences, provided by the users, are used to classify the selected users in two sets, collaborators $\mathcal{U_C}$ and non-collaborators $\mathcal{U_N}$. For the users of $\mathcal{U_N}$ the system has finished the first phase of the recommendation process.

For the users in $\mathcal{U_C}$ the system calculates the collaboration possibilities. To do it, for each two users $x, y \in \mathcal{U_C}$, the system performs the following steps:

1. Calculate the linguistic similarity measure between both users, $\sigma_l(VU_x, VU_y)$.
2. Obtain the linguistic compatibility degree between both users, which must be expressed in S_3. To do that, we apply the transformation function from Definition 4 to $\sigma_l(VU_x, VU_y)$.

3.3.2. Phase 2. Second filtering

In phase 1, the system has selected a number of resources NRS_u to be recommended to the user U. In this phase the system applies a second filter to the selection made in phase 1. This filter takes into account the user's U restrictions, and we obtain them by recovering the preferences inserted by U: number of resources and kind of resources. The number of resources can be a quantitative amount or a qualitative one. The idea is similar, but in the first case we work with a crisp value and in the second one we must work with fuzzy values. So, in the first case we compare two numerical values with relational operators, but in the second one we compare fuzzy values using the comparison operator defined in Herrera and Martínez (2000). Then, the user U wants to receive a number of recommendations REC_u (exactly or approximately):

1. If there are no enough resources to satisfy the amount of recommended resources specified by the user, $NRS_u < REC_u$, the system remembers the items previously selected but not recommended that now could be recommended. The system then repeats the recommendation process detailed in phase 1, but now incorporating these remembered resources.

2. If the number of selected resources is sufficient, $NRS_u \geq REC_u$, the system checks the user's restrictions as to whether he/she want only new resources or if he/she is also interested in previous resources but still with validity, which could be more interesting than a new resource. If U wants both kinds of resources, the system repeats the recommendation process of the first phase, but now incorporating these remembered resources.

Finally, the system sends to the users of the set $\mathcal{U}_{\mathcal{N}}$ the resource information and its calculated linguistic relevance degree, and for the users of $\mathcal{U}_{\mathcal{C}}$ the system sends the resource information, its calculated linguistic relevance degree and the collaboration possibilities characterized by its linguistic compatibility degrees.

3.4. Feedback phase

In this phase the recommender system recalculates and updates the recommendations of the accessed resources. When the system sends recommendations to the users, then they provide a feedback by assessing the relevance of the recommended resources, i.e., they supply their opinions about the recommendations received from the system. If they are satisfied with the received recommendation, they shall provide high value and vice versa. This feedback activity is developed in the following steps:

1. The system recommends the user U a resource R, and then the system asks for the opinion or evaluation judgements about recommended resource.
2. The user communicates the linguistic evaluation judgements to the system, $rc_y \in S_2$.
3. This evaluation is registered in the system for future recommendations. The system recalculates the linguistic recommendation of R by aggregating the opinions about R provided by all users. In such a way, the opinion supplied by U is considered. This can be done using the 2-tuple aggregation operator as $\overline{x}^e$ from Definition 3.

4. System evaluation

In this section we present the evaluation of the proposed system. The idea is to determine whether it fulfills the proposed innovations, that is, if the recommended items are useful and interesting for the users, while reducing the impact of the information overload. As of now we have implemented a trial version, in which the system works only with 10 researchers. But in the future we think to apply it in a real UDL, with the possibility of including a more exhaustive evaluation study. Moreover, our idea is also to set out some tests to measure the users' satisfaction.

4.1. Evaluation metrics

In the scope of recommender systems, precision, recall and F1 are measures widely used to evaluate the quality of the recommendations (Cao and Li, 2007; Cleverdon and Keen, 1966). We use them to evaluate the new proposal and to compare it with our previous system (Porcel and Herrera-Viedma, 2010). To calculate these measures we need a contingency table to categorize the items with respect to the information needs. The items are classified both as relevant or irrelevant, and selected (recommended to the user) or not selected. The contingency table (see Table 1) is created using these four categories.

Table 1. Contingency table for the resources

	Selected	Not selected	Total
Relevant	Nrs	Nrn	Nr
Irrelevant	Nis	Nin	Ni
Total	Ns	Nn	N

Precision is defined as the ratio of the selected relevant items to the selected items, that is, it measures the probability of a selected item to be relevant:

$$P = \frac{N_{rs}}{N_s}. \tag{10}$$

Recall is calculated as the ratio of the selected relevant items to the relevant items, that is, it represents the probability of a relevant item to be selected:

$$R = \frac{N_{rs}}{N_r}. \tag{11}$$

F1 is a combination measure that gives equal weight to both precision and recall:

$$F1 = \frac{2 \times R \times P}{R + P}. \tag{12}$$

4.2. Experimental results

The purpose of the experiment is to test the performance of the proposed system, so we compared the recommendations made by the system with the information provided by the library staff. When the users receive a recommendation, they provide a feedback to the system, assessing the relevance of the recommended resource, i.e., they provide their opinions about the recommendation supplied by the system. If they are satisfied with the recommendation, they provide a higher value. We use that feedback information to evaluate the system, applying the measures described in the previous section.

We considered a data set with 30 research resources of different areas, collected by the library staff from different information sources. These resources were included into the system following the indications described in Section 3.1. Initially, we limited the experiments to 10 users; all of them completed the registration process and they inserted their preferences about the five most relevant resources presented by the system (like in Fig. 4) and limitations about the number of resources they want to receive, taking into account that each user can select a different number.

From this information provided by the users, the system builds the user profiles. These user profiles, obtained from the provided preferences and the resources previously inserted, constituted our training data set. Then, we added 40 new resources that constituted the test data set. The system filtered these resources and recommended each one to the suitable users. To obtain data to compare, the 40 new resources also were recommended using the advice of the library staff.

For example, user 1 wants to receive 5 items, and the system selected 4 resources as relevant. However, from the information provided by the library staff and the user feedback, we could see that the system selected 1 irrelevant resource for user 1, and it did not select 2 resources that library staff considered relevant for the user 1. Then, to build the contingency table, we compared the recommendations provided by the system with the recommendations provided by the library staff and the relevance degrees inserted by the users. With this information, we build the contingency table for the recommended resources. It is shown in Table 2.

Table 2. Experimental contingency table

	User1	User2	User3	User4	User5	User6	User7	User8	User9	User10
Nrs	4	7	3	5	3	3	6	4	7	4
Nrn	2	4	1	3	2	1	1	2	3	3
Nis	1	3	2	2	2	1	4	3	3	1
Nr	6	11	4	8	5	4	7	6	10	7
Ns	5	10	5	7	5	4	10	7	10	5

From this contingency table we obtain the corresponding precision, recall and F1, which are shown in Table 3. The averages of precision, recall and F1 are 68.06%, 67.13% and 67.95%, respectively. Fig. 5 shows the graph with the precision, recall and F1 values for each user.

These values reveal a good performance of the proposed system, and therefore, high user satisfaction. However, to check it, now we compare the new system with the system proposed in Porcel and Herrera-Viedma (2010), using the precision, recall and F1, that we have analyzed. In Porcel and Herrera-Viedma

Table 3. Detailed experiment results for the recommendations

	Precision (%)	Recall (%)	F1 (%)
User1	80.00	66.67	72.73
User2	70.00	63.64	66,67
User3	60.00	75.00	66.67
User4	71.43	62.50	66.67
User5	60.00	60.00	60.00
User6	75.00	75.00	75.00
User7	60.00	85.71	70.59
User8	57.14	66.67	61.54
User9	70.00	70.00	70.00
User10	80.00	57.14	66.67
Average	68.06	67.13	67.95

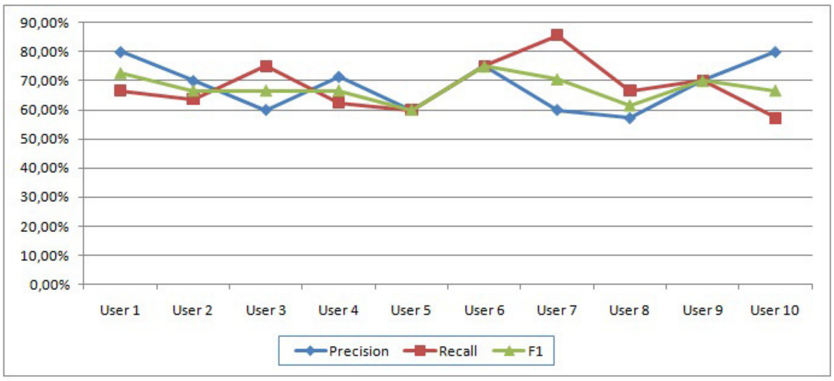

Figure 5. Experiment results

(2010) the results obtained were 67.50%, 61.39% and 63.51%, respectively, but as we have seen, with the new proposal we obtain 68.06%, 67.13% and 67.95%, improving therefore, over the previous approach.

5. Concluding remarks

Digital libraries can serve as powerful tools for universities to reach out and expand their sphere of influence in the society. UDL provide effective channels for the dissemination of knowledge, but we are bombarded continuously with a lot of information, and we face information overload problems. Users of UDL need tools to assist them in their processes of information gathering and the recommender systems have been applied successfully. However the number of electronic resources managed by the UDL staff continuously increases and the problem appears again. Therefore, we find the persistent problem of information overload. In this paper we have proposed an improved recommender system which uses a memory to avoid the information overload found in previous systems (Porcel, Moreno and Herrera-Viedma, 2009; Porcel and Herrera-Viedma, 2010). The proposal is to use the previously selected items to make a new selection. This selection is made taking into account the restrictions defined by the users, concerning the amount of resources that they want to receive and the novelty of these resources. The experimental results reveal a satisfactory performance of this new recommender system.

Acknowledgments

This work has been developed with the financing of FEDER funds in FUZZYLING project (TIN2007-61079), FUZZYLING-II project (TIN2010-17876), PETRI project (PET2007-0460), project of Ministry of Public Works (90/07) and Excellence Andalusian Project (TIC5299).

References

Alonso, S., Chiclana, F., Herrera, F., Herrera-Viedma, E., Alcalá-Fdez, J. and Porcel C. (2008) A Consistency-Based Procedure to Estimating Missing Pairwise Preference Values. *International Journal of Intelligent Systems* **23**, 155-175.

Alonso, S., Herrera-Viedma, E., Chiclana, F. and Herrera, F. (2009) Individual and Social Strategies to Deal With Ignorance Situations in Multi-Person Decision Making. *International Journal of Information Technology and Decision Making* **8** (2), 313-333.

Bobadilla, J., Serradilla, F. and Hernando, A. (2009) Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems* **22** (4), 261-265.

Bobadilla, J., Serradilla, F. and Bernal, J. (2010) A new collaborative filtering metric that improves the behavior of recommender systems. *Knowledge-Based Systems* **23** (6), 520-528.

Burke, R. (2007) Hybrid Web Recommender Systems. In: P. Brusilovsky, A. Kobsa and W. Nejdl, eds., *The Adaptive Web*, **LNCS 4321**, 377-408.

Butcher, H. (1998) *Meeting Managers' Information Needs.* A Managing Information Report. London, Aslib, The Association for Information Management.

Cabrerizo, F.J., Alonso, S. and Herrera-Viedma, E. (2009) A Consensus Model for Group Decision Making Problems with Unbalanced Fuzzy Linguistic Information. *International Journal of Information Technology & Decision Making* **8** (1), 109-131.

Callan, J. et al. (2003) Personalisation and Recommender Systems in Digital Libraries. *Joint NSF-EU DELOS Working Group Report.*

Cao, Y. and Li, Y. (2007) An intelligent fuzzy-based recommendation system for consumer electronic products. *Expert Systems with Applications* **33**, 230-240.

Chang, S.L., Wang, R.C. and Wang, S.Y. (2007) Applying a direct multi-granularity linguistic and strategy-oriented aggregation approach on the assessment of supply performance. *European Journal of Operational Research* **177** (2), 1013-1025.

Chao, H. (2002) Assessing the quality of academic libraries on the Web: The development and testing of criteria. *Library & Information Science Research* **24**, 169-194.

Cleverdon, C.W. and Keen, E.M. (1966) *Factors Determining the Performance of Indexing Systems*, Vol. 2 - Test Results. ASLIB Cranfield Res. Proj., Cranfield, Bedford, England.

Cornelis, C., Lu, J., Guo, X. and Zhang, G. (2007) One-and-only item recommendation with fuzzy logic techniques. *Information Sciences* **177** (22), 4906-4921.

Duen-Ren, L., Chin-Hui, L. and Wang-Jung, L. (2009) A hybrid of sequential rules and collaborative filtering for product recommendation. *Information Sciences* **179**, 3505-3519.

Edmunds, A. and Morris, A. (2000) The problem of information overload in business organizations: a review of the literature. *International Journal of Information Management* **20**, 17-28.

Gonçalves, M.A., Fox, E.A., Watson, L.T. and Kipp, N.A. (2004) Streams, structures, spaces, scenarios, societies (5s): A formal model for digital libraries. *ACM Trans. on Information Systems* **22** (2), 270-312.

Hanani, U., Shapira, B. and Shoval, P. (2001) Information Filtering: Overview of Issues, Research and Systems. *User Modeling and User-Adapted Interaction* **11**, 203-259.

Herrera, F. and Herrera-Viedma, E. (2000) Linguistic decision analysis: Steps for solving decisions problems under linguistic information. *Fuzzy Sets and Systems* **115**, 67-82.

Herrera, F., Herrera-Viedma, E. and Verdegay, J.L. (1997) Linguistic Measures Based on Fuzzy Coincidence for Reaching Consensus in Group Decision Making. *International Journal of Approximate Reasoning* **16**, 309-334

Herrera, F. and Martínez, L. (2000) A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems* **8** (6), 746-752.

Herrera, F. and Martínez, L. (2001) A model based on linguistic 2-tuples for dealing with multigranularity hierarchical linguistic contexts in multi-expert decision-making. *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics* **31** (2), 227-234.

Herrera-Viedma, E. (2001) An information retrieval system with ordinal linguistic weighted queries based on two weighting elements. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **9**, 77-88.

Herrera-Viedma, E., Alonso, S., Chiclana, F. and Herrera, F. (2007) A consensus model for group decision making with incomplete fuzzy preference relations. *IEEE Transactions on Fuzzy Systems* **15** (5), 863-877.

Herrera-Viedma, E., Chiclana, F., Herrera, F. and Alonso, S. (2007) Group decision making model with incomplete fuzzy preference relations based on additive consistency. *IEEE Transactions on Systems, Man and Cybernetics, Part B* **37** (1), 176-189.

Herrera-Viedma, E., Cordón, O., Luque, M., López, A.G. and Muñoz, A.M. (2003) A Model of Fuzzy Linguistic IRS Based on Multi-Granular Linguistic Information. *International Journal of Approximate Reasoning* **34** (3), 221-239.

Herrera-Viedma, E. and López-Herrera, A.G. (2007) A model of information retrieval system with unbalanced fuzzy linguistic information. *International Journal of Intelligent Systems* **22** (11), 1197-1214.

Herrera-Viedma, E., López-Herrera, A.G., Luque, M. and Porcel, C. (2007) A Fuzzy Linguistic IRS Model Based on a 2-Tuple Fuzzy Linguistic Approach. *International Journal of Uncertainty, Fuzziness and Knowledge-based Systems* **15** (2), 225-250.

Herrera-Viedma, E., Martínez, L., Mata, F. and Chiclana, F. (2005) A Consensus Support System Model for Group Decision-making Problems with Multi-granular Linguistic Preference Relations. *IEEE Transactions on Fuzzy Systems* **13** (5), 644-658.

Herrera-Viedma, E., Pasi, G., López-Herrera, A.G., and Porcel, C. (2006) Evaluating the information quality of web sites: A methodology based on fuzzy computing with words. *Journal of the American Society for Information Science and Technology* **57** (4), 538-549.

Herrera-Viedma, E. and Peis, E. (2003) Evaluating the informative quality of documents in SGML-format using fuzzy linguistic techniques based on computing with words. *Information Processing and Management* **39** (2), 233-249.

Iskold, A. (2007) The Art, Science and Business of Recommendation Engines. `http://www.readwriteweb.com/archives/recommendation_engines.php`.

Korfhage, R.R. (1997) *Information Storage and Retrieval.* Wiley Computer Publishing, New York.

Leung, C.W., Chan, S.C. and Chung, F. (2008) An empirical study of a cross-level association rule mining approach to cold-start recommendations. *Knowledge-Based Systems*, **21** (7), 515-529.

Marchionini, G. (2000) Research and Development in Digital Libraries. `http://ils.unc.edu/~march/digital_library_R_and_D.html`.

Martínez, L., Pérez, L.G., Barranco, M. and Espinilla, M. (2008) Effectiveness of Knowledge Based Recommender Systems Using Incomplete Linguistic Preference Relations. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **16** (2), 33-56.

Mata, F., Martnez, L. and Herrera-Viedma, E. (2009) An Adaptive Consensus Support Model for Group Decision Making Problems in a Multi-Granular Fuzzy Linguistic Context. *IEEE Transactions on Fuzzy Systems* **17** (2), 279-290.

Meghabghab, G. and Kandel, A. (2008) *Search Engines, Link Analysis, and Users Web Behavior. Springer-Verlag, Berlin Heidelberg.*

Morales del Castillo, J.M., Pedraza-Jiménez, R., Ruíz, A.A., Peis, E. and Herrera-Viedma, E. (2009) A Semantic Model of Selective Dissemination of Information for Digital Libraries. *Information Technology and Libraries* **28**(1), 22-31.

Nelson, M.R. (1994) We have the information you want, but getting it will cost you: being held hostage by information overload. *Crossroads: Special issue on the Internet* **1** (1), 11-15.

Porcel, C., Moreno, J.M. and Herrera-Viedma, E. (2009) A multi-disciplinar recommender system to advice research resources in University Digital Libraries. *Expert Systems with Applications* **36** (10), 12520-12528.

Porcel, C. and Herrera-Viedma, E. (2010) Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries. *Knowledge-Based Systems* **23** (1), 32-39.

Quiroga, L.M. and Mostafa, J. (2002) An experiment in building profiles in information filtering: the role of context of user relevance feedback. *Information Processing and Management* **38**, 671-694.

Renda, M.E. and Straccia, U. (2005) A personalized collaborative Digital Library environment: a model and an application. *Information Processing and Management* **41**, 5-21.

Reisnick, P. and Varian, H.R. (1997) Recommender Systems. *Special issue of Communications of the ACM* **40** (3), 56-59.

Ross, L. and Sennyey, P. (2008) The Library is Dead, Long Live the Library! The Practice of Academic Librarianship and the Digital Revolution. *The Journal of Academic Librarianship* **34** (2), 145-152.

Symeonidis, P., Nanopoulos, A., Papadopoulos, A.N. and Manolopoulos, Y. (2008) Collaborative recommender systems: Combining effectiveness and efficiency. *Expert Systems with Applications* **34**, 2995-3013.

Torra, V. and Narukawa, Y. (2007) *Modeling Decisions: Information Fusion and Aggregation Operators.* Springer.

Yang, J.M and Li, K.F. (2009) Recommendation based on rational inferences in collaborative filtering *Knowledge-Based Systems* **22** (1), 105-114.

Zadeh, L.A. (1975) The concept of a linguistic variable and its applications to approximate reasoning. *Information Sciences.* Part I. **8**, 199-249. Part II. **8**, 301-357. Part III. **9**, 43-80.