Chapter 3

Using a Recommender System to Help the Technology Transfer Office Staff to Disseminate Selective Information

C. Porcel[1], A. Tejeda-Lorente[2], M.A. Martínez[2] and E. Herrera-Viedma[2]

[1] *Department of Computer Science, University of Jaen23071 Jaén, Spain*
[2] *Department of Computer Science and Artificial Intelligence, University of Granada 18071 Granada, Spain*

Email: cporcel@ujaen.es, atejeda@decsai.ugr.es, mnglsmartinez@gmail.com, viedma@decsai.ugr.es

Recommender systems evaluate and filter the great amount of information available on the Web, so they could be used to help users in their access processes to relevant information. In the literature we can find a lot of approaches for generating personalized recommendations. Hybrid recommender systems combine in different ways several approaches, so these recommendation strategies represent a promising solution for multiple applications. In this paper we propose a hybrid fuzzy linguistic recommender system to help the Technology Transfer Office staff in the dissemination of research resources interesting for the users. The system recommends users both specialized and complementary research resources and additionally, it discovers potential collaboration possibilities in order to form multidisciplinary working groups.

3.1 Introduction

Theoretical and empirical works in innovation economics suggest that the use of scientific knowledge by setting up and maintaining good industry/science relations positively affects innovation performance [43]. In terms of organizational structure, creating a specialized Technology Transfer Office Technology Transfer Office (TTO) within a university can be instrumental in developing relations with the industry [45]. The TTOs were established to facilitate commercial knowledge transfers from universities to practitioners or university/industry technology transfer [59]. They are responsible for managing and putting into action the activities which generate knowledge and technical and scientific

C. Kahraman (ed.), *Computational Intelligence Systems in Industrial Engineering*, Atlantis Computational Intelligence Systems 6, DOI: 10.2991/978-94-91216-77-0_3,

collaboration, thus enhancing the interrelation between researchers at the university and the entrepreneurial world and their participation in various support programmes designed to carry out research, development and innovation activities. A service that is particularly important to fulfill this objective is the selective dissemination of information about research resources. But the TTO staff finds difficulties in achieving an effective selective dissemination of information. To solve this problem, automatic techniques are needed in the TTO to facilitate users to selectively access to research resources. In this sense, we consider interesting two different tools to facilitate the access to the information: Information Retrieval Systems [33, 39, 42] and Recommender Systems [4, 17, 40, 54, 60]. The former are focused on information search in a known content repository while the later are focused on information discovery in partially known frameworks.

Recommender systems attempt to discover information items that are likely of interest to a user. They are especially useful when they identify information that a person was previously unaware of. They are becoming popular tools for reducing information overload and to improve the sales in e-commerce web sites [7, 36, 54]. The provision of personalized recommendations requires that the system knows something about every user, such as the ratings provided by the users about the explored items. This knowledge implies that the system must maintain users' profile containing the users' preferences or needs.

From a theoretical point of view, recommender systems have fallen into two main categories: [16, 17, 19, 47, 52, 54, 57, 60]. *Content-based recommender systems* and *Collaborative recommender systems* (see Section 3.2). If we analyze the TTO scope, we find that the collaborative filtering approach is very useful because it allows users to share their experiences, so that popular resources can be easily located or people can receive information items found useful by others with similar profiles. But the collaborative approaches tend to fail when little is known about items, i.e., the system has few ratings. For this reason, we propose to combine the content-based and collaborative approaches to obtain a hybrid recommendation scheme.

The aim of this paper is to present a hybrid fuzzy linguistic recommender system which is applied in the TTO in the University of Granada. In such a way, it allows to help the TTO staff to selectively disseminate research knowledge and the researchers to discover information. The most important novelties of this fuzzy linguistic recommender system are:

- The system implements a hybrid recommendation strategy based in a switching hybrid approach [6], which switches between a content-based recommendation approach and a collaborative one to share the user individual experience and social wisdom.
- The system implements a personalization tool that allows to recommend users three types of items:
- Specialized resources of the own user research area to contribute to his/her specialization.
- Other resources as complementary formation.
- Research collaborators. In this case, it allows researchers to discover new members with complementary profiles, which could provide them real collaboration possibilities to form multidisciplinary working groups and develop common projects.
- The system implements a richer feedback process: when researchers analyze a recommended resource, they provide a satisfaction degree. In such a way, we guarantee that user experiences are taken into account to generate the recommendations done by the system.

The paper is structured as follows. Section 3.2 presents the basic concepts and aspects about the recommender systems. Section 3.3 revises the multi-granular fuzzy linguistic modelling. In Section 3.4 we present the new recommender system to selectively advice research resources in a TTO. Section 3.5 reports the system evaluation and the experimental results. Finally, our concluding remarks are pointed out in Section 3.6.

3.2 Basis of Recommender Systems

The Recommender systems have the effect of guiding the users in a personalized way to relevant or useful objects in a large space of possible options [6]. 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 [7, 10, 14, 15, 36, 41, 54]. The construction of accurate profiles is a key task and the system's success will depend on a large extent on the ability of the learned profiles to represent the user's preferences. Then, 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 which were previously interesting for the user, click-stream data, demographic information, and so on.

Another key aspect to consider when designing the system is the approach used to generate the recommendations. Taking into account the knowledge source, four different approaches can be distinguished: [7, 16, 17, 52, 54, 60]

- *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 [5, 11]. These recommender systems tend to fail when little is known about the user information needs. This is called the new user cold-starting problem [38].
- *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-starting problem [7].
- *Demographic systems*: These systems provide recommendations based on a demographic profile of the user. Recommended products can be produced for different demographic niches, by combining the ratings of users in those niches [51].
- *Knowledge-based systems*: This kind of recommender systems suggest items based on inferences about a users' preferences. This knowledge will sometimes contain explicit knowledge about how the items meet the users' preferences [6].

Each approach has certain advantages and, of course, disadvantages, depending on the scope settings. One solution is to combine different approaches to reduce the disadvantages of each one of them and to exploit their benefits. Using a hybrid strategy, users are provided with more accurate recommendations than those offered by each strategy individually [5, 11, 16]. For this reason, in this paper we propose the use of a hybrid approach.

Moreover, 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 about the recommended items and the system uses these evaluations to automatically update user profiles in order to generate new recommendations [17, 54].

3.3 Multi-Granular Fuzzy Linguistic Modeling

The Fuzzy linguistic modelling of Fuzzy Sets Theory has given very good results to model qualitative information [62] and it has been proven to be useful in many problems, e.g., decision making [2, 9, 21, 23, 37, 64], quality evaluation [8, 28, 35, 48], information retrieval [26, 27, 29–32], political analysis [3], estimation of student performances [50], etc. It is a tool based on the concept of *linguistic variable* proposed by Zadeh [62].

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. According to the uncertainty degree that an expert qualifying a phenomenon has on it, the linguistic term set chosen to provide his knowledge will have more or less terms. When different experts have different uncertainty degrees on the phenomenon, then several linguistic term sets with a different granularity of uncertainty are necessary [22]. The use of different label sets to assess information is also necessary when an expert has to evaluate different concepts, as it happens in information retrieval problems when users have to evaluate the importance of the query terms and the relevance of the retrieved documents [28]. In such situations, we need tools to manage multi-granular linguistic information [25, 34, 46].

3.3.1 *The 2-Tuple Fuzzy Linguistic Approach*

The 2-tuple fuzzy linguistic modeling [24] is a continuous model of information representation that allows to reduce the loss of information that typically arise when using other fuzzy linguistic approaches (classical and ordinal [20, 62]). To define it both the 2-tuple representation model and the 2-tuple computational model to represent and aggregate the linguistic information have to be established.

Let $S = \{s_0, \ldots, s_g\}$ be a linguistic term set with odd cardinality, where the mid term represents an indifference value and the rest of the terms are symmetric related to it. We assume that the semantics of labels is given by means of fuzzy subsets defined in the $[0,1]$ interval, which are described by their membership functions $\mu_{si} : [0,1] \rightarrow [0,1]$, and we consider all terms distributed on a scale on which a total order is defined, that is, $s_i = s_j \Longleftrightarrow i = j$.

In this fuzzy linguistic context, if a symbolic method [20, 23] aggregating linguistic information obtains a value $\beta \in [0,g]$, and $\beta \notin \{0,\ldots,g\}$, then an approximation function is used to express the result in S.

Definition 3.1 ([24])**.** Let β be the result of an aggregation of the indexes of a set of labels assessed in a linguistic term set S, i.e., the result of a symbolic aggregation operation, $\beta \in [0,g]$. Let $i = \text{round}(\beta)$ and $\alpha = \beta - i$ be two values, such that, $i \in [0,g]$ and $\alpha \in [-0.5, 0.5)$ then s_i represents the linguistic label of the information, and α_i is a numerical value expressing the value of the symbolic translation from the original result β to the closest index label, i, in the linguistic term set ($s_i \in S$).

This model defines a set of transformation functions between numeric values and 2-tuples.

Definition 3.2 ([24]). Let $S = \{s_0, \ldots, s_g\}$ be a linguistic term set and $\beta \in [0, g]$ a value representing the result of a symbolic aggregation operation, then the 2-tuple that expresses the equivalent information to β is obtained with the following function:

$$\Delta : [0, g] \longrightarrow S \times [-0.5, 0.5)$$
$$\Delta(\beta) = (s_i, \alpha), \ \text{with} \ \begin{cases} s_i & i = \text{round}(\beta) \\ \alpha = \beta - i, & \alpha \in [-0.5, 0.5) \end{cases} \tag{3.1}$$

where $round(\cdot)$ is the usual round operation, s_i has the closest index label to β and α is the value of the symbolic translation.

For all Δ there exists Δ^{-1}, defined as $\Delta^{-1}(s_i, \alpha) = i + \alpha$. On the other hand, it is obvious that the conversion of a linguistic term into a linguistic 2-tuple consists of adding a symbolic translation value of 0: $s_i \in S \Rightarrow (s_i, 0)$.

The computational model is defined by presenting the Negation operator, Comparison of 2-tuples and Aggregation operators. Using functions Δ and Δ^{-1} any of the existing aggregation operators can be easily extended for dealing with linguistic 2-tuples [61, 63].

Definition 3.3 (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:

$$\overline{x}^e[(r_1, \alpha_1), \ldots, (r_n, \alpha_n)] = \Delta\left(\sum_{i=1}^{n} \frac{1}{n} \Delta^{-1}(r_i, \alpha_i)\right) = \Delta\left(\frac{1}{n} \sum_{i=1}^{n} \beta_i\right) \tag{3.2}$$

Definition 3.4 (Weighted Average Operator). Let $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ be a set of linguistic 2-tuples, and $W = \{w_1, \ldots, w_n\}$ be their associated weights. The 2-tuple weighted average $\overline{x}^w$ is computed as:

$$\overline{x}^w[(r_1, \alpha_1), \ldots, (r_n, \alpha_n)] = \Delta\left(\frac{\sum_{i=1}^{n} \Delta^{-1}(r_i, \alpha_i) \cdot w_i}{\sum_{i=1}^{n} w_i}\right) = \Delta\left(\frac{\sum_{i=1}^{n} \beta_i \cdot w_i}{\sum_{i=1}^{n} w_i}\right) \tag{3.3}$$

Definition 3.5 (Linguistic Weighted Average Operator). Let $x = \{(r_1, \alpha_1), \ldots, (r_n, \alpha_n)\}$ be a set of linguistic 2-tuples, and $W = \{(w_1, \alpha_1^w), \ldots, (r_n, \alpha_n^w)\}$ be their linguistic 2-tuple associated weights. The 2-tuple linguistic weighted average $\overline{x}_l^w$ is computed as:

$$\overline{x}_l^w[((r_1, \alpha_1), (w_1, \alpha_1^w)), \ldots, ((r_n, \alpha_n), (w_n, \alpha_n^w))] = \Delta\left(\frac{\sum_{i=1}^{n} \beta_i \cdot \beta w_i}{\sum_{i=1}^{n} \beta w_i}\right) \tag{3.4}$$

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

3.3.2 *Linguistic Hierarchy to Model Multi-Granular Linguistic Information*

A Linguistic Hierarchy, LH, is a set of levels $l(t,n(t))$, i.e., $LH = \bigcup_t l(t,n(t))$, where each level t is a linguistic term set with different granularity $n(t)$ from 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 from its predecessor level as: $l(t,n(t)) \rightarrow l(t+1, 2 \cdot n(t) - 1)$.

Definition 3.6 ([25]). Let $LH = \bigcup_t l(t,n(t))$ be a linguistic hierarchy whose linguistic term sets are denoted as $S^{n(t)} = \{s_0^{n(t)}, \ldots, s_{n(t)-1}^{n(t)}\}$. The transformation function between a 2-tuple that belongs to level t and another 2-tuple in level $t' \neq t$ is defined as:

$$
\begin{aligned}
&TF_{t'}^{t} \,:\, l(t,n(t)) \longrightarrow l(t',n(t')) \\
&TF_{t'}^{t}(s_i^{n(t)}, \alpha^{n(t)}) = \Delta\left(\frac{\Delta^{-1}\left(s_i^{n(t)}, \alpha^{n(t)}\right) \cdot (n(t') - 1)}{n(t) - 1}\right)
\end{aligned}
\tag{3.5}
$$

As it was pointed out in [25] this family of transformation functions is bijective. This result guarantees the transformations between levels of a linguistic hierarchy are carried out without loss of information.

3.4 A Recommender System for the Selective Dissemination of Research Resources in a TTO

In this section, we present a fuzzy linguistic hybrid recommender system to discover both researchers information about research resources and collaboration possibilities. The TTO staff manages and spreads knowledge about research resources such as R&D bulletins, R&D&I, calls, notices, research projects and so on [43, 45]. Nowadays, this amount of information grows continuously and the TTO staff needs automated tools to filter and spread that information to the researchers in a simple and timely manner.

As aforementioned, we present a hybrid recommender system which can be used in a real TTO environment to achieve an effective selective dissemination of research resources. This system works according to a hybrid recommendation strategy based in a switching hybrid approach [6], which switches between a content-based recommendation approach and a collaborative one to share user experiences by generating social recommendations. Basically, the former is applied when a new item is inserted and the latter is applied when a new researcher is registered. Furthermore, we include another feature when suggesting resources to researchers, because the system recommends both specialized and complementary research resources. It improves the recommendation process, allowing researchers to discover real collaboration possibilities in order to form multidisciplinary working groups. In such a way, the system improves the services of a TTO, selectively disseminating research resources, and allowing to share knowledge in an academic context.

We present a multi-granular fuzzy linguistic recommender system that provides high flexibility in the communication processes between users and the system. We use different label sets ($S_1, S_2, \ldots$) 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 a 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 four concepts that can be assessed in the activity of the recommender system:

- **Importance degree** of a discipline with respect to a resource scope or user interest topic, which is assessed in S_1.
- **Similarity degree** among resources or among users, which is assessed in S_2.
- **Relevance degree** of a resource for a user, which is assessed in S_3.
- **Satisfaction degree** expressed by a user to evaluate a recommended resource, which is assessed in S_4.

We follow a linguistic hierarchy composed by 2 levels, the level 2 (5 labels) to represent importance degrees ($S_1 = S_5$), and the level 3 (9 labels) to represent similarity degrees ($S_2 = S_9$), relevance degrees ($S_3 = S_9$) and satisfaction degrees ($S_4 = S_9$). As the importance degrees are provided by TTO staff, we use a set of five labels to facilitate them the characterization of resource scopes or user interest topics. On the other hand, as the similarity and relevance degrees are computed automatically by the system, we use the set of 9 labels which presents an adequate granularity level to represent the results. Similarly, to provide users with a label set with an adequate granularity level we use the set of 9 labels to express the satisfaction degrees. Using this LH, the linguistic terms in each level are the following:

- $S^5 = \{b_0 = \text{None} = N, b_1 = \text{Low} = L, b_2 = \text{Medium} = M, b_3 = \text{High} = H, b_4 = \text{Total} = T\}$

- $S^9 = \{c_0 = \text{None} = N,\ c_1 = \text{Very\_Low} = VL,\ c_2 = \text{Low} = L,\ c_3 = \text{More\_Less\_Low} = MLL,\ c_4 = \text{Medium} = M,\ c_5 = \text{More\_Less\_High} = MLH,\ c_6 = \text{High} = H,\ c_7 = \text{Very\_High} = VH,\ c_8 = \text{Total} = T\}$.

In Fig. 3.1 we show the basic operating scheme of the recommender system which is based on four main components, which we explain now.

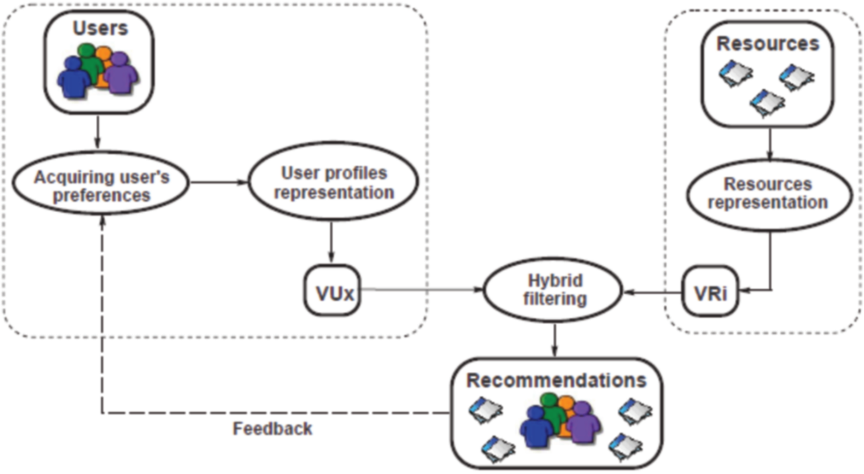

Fig. 3.1 Basic operating scheme

3.4.1 *Resource Representation*

The resources we consider in our system are the research resources such as R&D bulletins, R&D&I, calls, notices or research projects. Once the TTO staff inserts 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 [39] to represent the resource scope and a classification composed by 25 disciplines (see Fig. 3.2), i.e., a research resource i is represented as $VR_i = (VR_{i1}, VR_{i2}, \ldots, VR_{i25})$ 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 TTO staff when they add new resources.

3.4.2 *User Profiles Representation*

We consider that the users of our system are the researchers of the university and the environment companies. To characterize a researcher the system stores the personal information (login, password, name, phone, email, etc.), research group (it is a string composed by 6 digits, 3 characters indicating the research area and 3 numbers identifying the group) and his/her topics of interest. Similarly, we use the vector model [39] to represent the topics of interest. Then, for a researcher e, we have a vector $VU_e = (VU_{e1}, VU_{e2}, \ldots, VU_{e25})$ where each component $VU_{ej} \in S_1$ is a linguistic assessment that represents the importance degree of the discipline j in the topics of interest of researcher e. Similarly these importance degrees are assigned by the TTO staff when they add a new researcher.

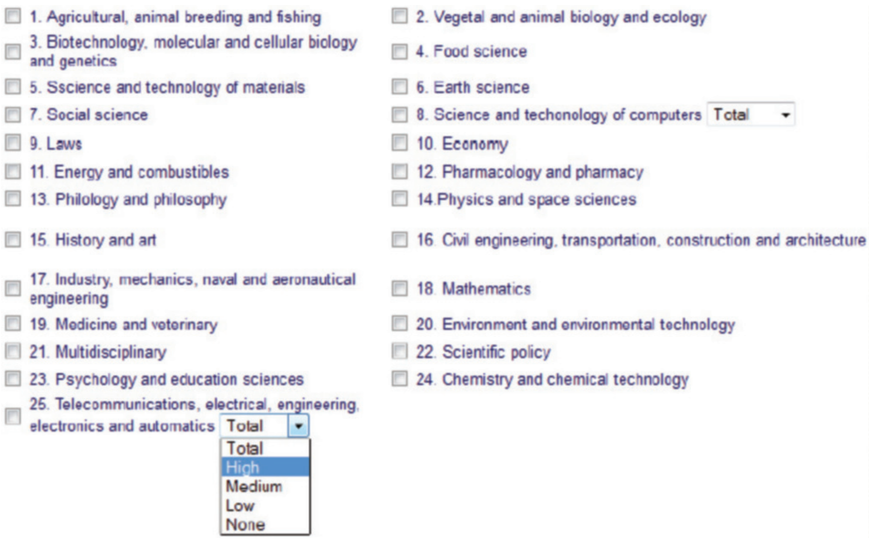

Fig. 3.2 Interface to define the disciplines of the resource scope or user preferences

Furthermore, to avoid the cold-starting problem to handle new items or new users [7, 38], when a new user is inserted, to confirm his/her register it is necessary that he/she assesses some of the resources stored in the system. To do this, the system shows the items randomly and the user assesses what he/she wants.

3.4.3 *Recommendation Strategy*

In this phase the system filters the incoming information to deliver it to the fitting users. This process is based on a matching process developed by similarity measures, such as Euclidean Distance or Cosine Measure [39]. In particular, we use the standard cosine measure but defined in a linguistic framework:

$$\sigma_l(V_1,V_2) = \Delta\left(g \times \frac{\sum_{k=1}^{n}\left(\Delta^{-1}(v_{1k},\alpha_{v1k}) \times \Delta^{-1}(v_{2k},\alpha_{v2k})\right)}{\sqrt{\sum_{k=1}^{n}\left(\Delta^{-1}(v_{1k},\alpha_{v1k})\right)^2} \times \sqrt{\sum_{k=1}^{n}\left(\Delta^{-1}(v_{2k},\alpha_{v2k})\right)}}\right)^2 \quad (3.6)$$

with $\sigma_l(V_1,V_2) \in S_2 \times [-0.5, 0.5)$, and where g is the granularity of the term set used to express the relevance degree, i.e. S_2, n is the number of disciplines and (v_{ik}, α_{vik}) is the 2-tuple linguistic value of discipline k in the vector V_i representing the resource scope or user interest topics, depending of the used filtering strategy.

This recommender system works according to a hybrid recommendation strategy. Our proposal is based in a *switching hybrid approach* [6], which uses one technique or another, depending on some criterion. In our system, a content-based recommendation approach is applied when a new item is inserted and a collaborative one is applied when a new researcher is registered. In both cases, the recommender system could send three types of recommendations to a researcher research resources of his/her same area (specialized), research resources of complementary areas, and collaboration possibilities with other researchers.

3.4.3.1 *Content-Based Recommendations*

When a new resource i arrives to the system, the system calculates the content-based recommendations to be sent to a researcher e as follows:

- Compute the linguistic similarity degree between VR_i and VU_e.
- Establish if the resource i could contribute to specialize or complement the researcher's profile. Assuming that $S_2 = S^9$, we consider that a resource i is related with the researcher's profile e if $\sigma_l(VR_i, VU_e) > (s_4^9, 0)$, i.e., if the linguistic similarity degree is higher than the mid linguistic label. We consider that the resource i could contribute to specialize the researcher's profile e when $\sigma_l(VR_i, VU_e) \geqslant (s_6^9, 0)$. On the other hand, we consider that the resource i could contribute to complement the researcher's profile e when $(s_2^9, 0) \leqslant \sigma_l(VR_i, VU_e) < (s_6^9, 0)$.

- If i is considered a specialization resource for e, then the system recommends this resource i to e with a relevance degree $i(e) \in S_3 \times [-0.5, 0.5)$ which is obtained as follows:
- Look for all specialized research resources stored in the system that were previously assessed by e, i.e., the set of resources $K = \{1, \ldots, k\}$ such that there exists the linguistic satisfaction assessment $e(j)$, $j \in K$ and $\sigma_l(VR_i, VU_e) \geqslant (s_6^9, 0)$.
- Then,

$$i(e) = \overline{x}_l^w(((e(1),0), \sigma_l(VR_i, VR_1)), \ldots, ((e(k),0), \sigma_l(VR_i, VR_k))) \tag{3.7}$$

where $\overline{x}_l^w$ is the linguistic weighted average operator (see Definition 3.5).
- If i is considered a complementary resource for e, then the system recommends this resource i and its authors (community members that could be potential collaborators) to e with a relevance degree $i(e) \in S_3 \times [-0.5, 0.5)$ which is obtained as follows:
- Look for all complementary research resources stored in the system that were previously assessed by e, i.e., the set of resources $K = \{1, \ldots, k\}$ such that there exists the linguistic satisfaction assessment $e(j)$, $j \in K$ and $(s_2^9, 0) \leqslant \sigma_l(VR_i, VU_e) < (s_6^9, 0)$. The latter defines a complementary linguistic interval around mid label that is considered the maximum complementary level.
- Then,

$$i(e) = \overline{x}_l^w(((e(1),0), f(i,1)), \ldots, ((e(k),0), f(i,k))) \tag{3.8}$$

where f is a triangular multidisciplinary matching function that measures the complementary degree between two resources i and j,

$$f(i,j) = \begin{cases} \Delta(2 \times \Delta^{-1}(\sigma_l(VR_i, VR_j))) & \text{if } 0 \leqslant \Delta^{-1}(\sigma_l(VR_i, VR_j)) \leqslant \frac{1}{2} \\ \Delta(2 \times (1 - \Delta^{-1}(\sigma_l(VR_i, VR_j)))) & \text{if } \frac{1}{2} < \Delta^{-1}(\sigma_l(VR_i, VR_j)) \leqslant 1 \end{cases} \tag{3.9}$$

3.4.3.2 *Collaborative Recommendations*

When new users are inserted into the system, they receive recommendations about already inserted resources which may be interesting for them. Usually, new users provide little information about the items that satisfy their topics of interest, so we use the collaborative approach to generate their recommendations. Exactly, we follow a memory-based algorithm or nearest-neighbor algorithm, which generates the recommendations according to the preferences of nearest neighbors. This algorithm has proven good performance [19, 60]. In the following we describe the process in detail.

Given a new researcher e, the recommendations to be sent to e are obtained in the following steps:

- Identify the set of users $\aleph_e$ most similar to that new user e. To do so, we calculate the linguistic similarity degree between the topics of interest vector of the new user (VU_e) against the vectors of all users already inserted into the system (VU_y, $y = 1,\ldots,n$ where n is the number of users), that is, we calculate $\sigma_l(VU_e, VU_y) \in S_2$. As $S_2 = S^9$, we consider that the user y is near neighbor to e if $\sigma_l(VU_e, VU_y) > (s_4^9, 0)$, i.e., if the linguistic similarity degree is higher than the mid linguistic label.
- Look for the resources stored in the system that were previously well assessed by the near neighbors of e, i.e., the set of resources $K = \{1,\ldots,k\}$ such that there exists a linguistic satisfaction assessment $y(j)$, $y \in \aleph_e$, $j \in K$, and $y(j) \geqslant (s_6^9, 0)$.
- Discover if those resources could contribute with specialized or complementary formation. A resource $j \in K$could contribute to specialize the researcher's formation e when $\sigma_l(VR_j, VU_e) \geqslant (s_6^9, 0)$. On the other hand, we consider that the resource j could contribute to complement the researcher's formation e when $(s_2^9, 0) \leqslant \sigma_l(VR_j, VU_e) < (s_6^9, 0)$.
- If j is considered a specialization resource for e, then the system recommends this resource j to e with a relevance degree $j(e) \in S_3 \times [-0.5, 0.5)$ which is obtained as follows:
- To look for all linguistic satisfaction assessments about resources that were well assessed by the nearest neighbors of e. That is, we recovery $y(j)$, $j \in K$ and $y \in \aleph_e$.
- Then,

$$j(e) = \overline{x}_l^w(((y_1(j), 0), \sigma_l(VU_e, VU_{y1})), \ldots, ((y_n(j), 0), \sigma_l(VU_e, VU_{yn}))) \qquad (3.10)$$

  where $y_1, \ldots, y_n \in \aleph_e$ and $\overline{x}_l^w$ is the linguistic weighted average operator (see Definition 3.5).
- If j is considered a complementary resource for e, then the system recommends this resource j and its authors (community members that could be potential collaborators) to e with a relevance degree $j(e) \in S_3 \times [-0.5, 0.5)$ which is obtained as follows:
- Look for all complementary research resources stored in the system that previously were well assessed by the nearest neighbors of e, i.e., the set of resources $K = \{1,\ldots,k\}$ such that there exists the linguistic satisfaction assessment $y(j)$, with $j \in K$, $y \in \aleph_e$ and $(s_2^9, 0) \leqslant \sigma_l(VU_y, VR_{je}) < (s_6^9, 0)$. The latter defines a complementary linguistic interval around mid label that is considered the maximum complementary level.

- Then,

$$j(e) = \overline{x}_l^w(((y_1(j),0),h(e,y_1)),\ldots,((y_n(j),0),h(e,y_n))) \tag{3.11}$$

where f is a triangular multidisciplinary matching function that measures the complementary degree between two resources i and j,

$$h(i,j) = \begin{cases} \Delta(2\times\Delta^{-1}(\sigma_l(VU_i,VU_j))) & \text{if } 0 \leqslant \Delta^{-1}(\sigma_l(VU_i,VU_j)) \leqslant \dfrac{1}{2} \\ \Delta(2\times(1-\Delta^{-1}(\sigma_l(VU_i,VU_j)))) & \text{if } \dfrac{1}{2} < \Delta^{-1}(\sigma_l(VU_i,VU_j)) \leqslant 1 \end{cases} \tag{3.12}$$

3.4.4 *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 recommendations, 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 values and vice versa. This feedback activity is developed in the following steps:

- The system recommends the user U a resource R, and then the system asks him/her his/her opinion or evaluation judgements about recommended resource.
- The user communicates his/her linguistic evaluation judgements to the system, $rc_y \in S_2$.
- 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 $\overline{x}^e$ given in Definition 3.3.

3.5 Experiments and Evaluation

In this section we present the evaluation of the proposed recommender system. We propose two kind of experiments, offline and online ones. We begin with an offline setting, where the proposed recommendation approach is compared with other approaches without user interaction, using a standard data set. However, in many applications, accurate predictions are important but insufficient with respect to the user satisfaction. For instance, users may be interested in discovering new items not expected for them, more than getting an exact prediction of their preferences. Consequently, we also propose online experiments, that is, practical studies where a small group of users interact with the system and report us their experiences.

3.5.1 *Evaluation Metrics*

In the scope of recommender systems, precision, recall and F1 are widely used measures to evaluate the quality of the recommendations [10, 12, 56]. To calculate these metrics we need to build 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.

Definition 3.7. *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:

$$P = \frac{N_{rs}}{N_s} \tag{3.13}$$

Definition 3.8. *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:

$$R = \frac{N_{rs}}{N_r} \tag{3.14}$$

Definition 3.9. *F1* is a combination metric that gives equal weight to both precision and recall, and it is calculated as follows: [10, 56]

$$F1 = \frac{2 \times R \times P}{R + P} \tag{3.15}$$

Besides, in order to test the performance of our model and to compare it with other approaches, we also calculate the system accuracy, that is, its capability to predict users' ratings. We propose to use the Mean Absolute Error.

Definition 3.10. *Mean Absolute Error (MAE)* [19, 58] is a commonly used accuracy metric which considers the average absolute deviation between a predicted rating and the user's true rating:

$$MAE = \frac{\sum_{i=1}^{n} abs(p_i - r_i)}{n} \tag{3.16}$$

where n is the number of cases in the test set, p_i the predicted rating for a item, and r_i the true rating.

3.5.2 *Offline Experiments*

3.5.2.1 *Data Set*

We use MovieLens data sets [18, 49] to develop the offline experiments because the data sets are publicly available and have been usually used to evaluate recommender systems, and in such a way, we could compare our system with other models. MovieLens data sets [49] are related with a cinematographic scope and they were collected by the GroupLens Research Project at the University of Minnesota during the seven-month period from September 19$^{\text{th}}$, 1997 through April 22$^{\text{nd}}$, 1998. Specifically, we use the 100K ratings data set which contains 1682 movies, 943 users and a total of 100 000 ratings on a scale of 1-5 (where 1 = Awful, 2 = Fairly bad, 3 = It's OK, 4 = Will enjoy, 5 = Must see). Each user has rated at least 20 movies. However, to apply this data set to our hybrid recommender system, we need to develop a transformation process in order to adapt the data to the features of our approach. In our system we represent both the resources and the user profiles using vectors. So, we need to transform the MovieLens data sets to this representation avoiding the loss of information. Then, we have to build vectors to represent the users' topics of interest and the movies. The idea is to obtain such vectors from the data stored in the MoviLens data sets.

The 1682 movies are classified into the following 19 genres: unknown, action, adventure, animation, children, comedy, crime, documentary, drama, fantasy, film-noir, horror, musical, mystery, romance, sci-fi, thriller, war and western. In fact, the file *u.item* contains information about the movies, with a tab separated list of the fields movie id, movie title, release date, video release date, IMDb URL, and the last 19 fields are the genres: a value of 1 indicates the movie is of that genre and a value of 0 indicates it is not; movies can be in several genres at once. For each movie we build a vector with 19 positions (one for each genre), following the approach pointed in Subsection 3.4.1 $VR_i = (VR_{i1}, VR_{i2}, \ldots, VR_{i19})$ where each component $VR_{ij} \in S_1$ is a linguistic assessment that represents the importance degree of the genre j with regard to the movie i. Therefore, when the value in the file *u.item* is 1 (the movie is of that genre), we assign the maximum label of S_1 $((b_4, 0)$ in this case) and when the value is 0 the assigned label is the minimum of S_1 $((b_0, 0))$.

On the other hand, our system works with the user topics of interest, which are also represented by a vector. So, for each user we need a vector similar to that used to represent the movies. The problem is that MovieLens data sets don't include this information directly, because the file *u.user* only includes demographic information about the users (user id, age, gender, occupation and zip code). However, the information about the topics of interest for each user could be obtained from the available data, aggregating the ratings assigned by the users on each movie with the genre information of the movies. The file *u.data* contains the 100 000 ratings on a scale of 1-5; this is a tab separated list of user id, item id, rating and timestamp. The information about the genres is in the file *u.item*; the movie ids are the ones used in the *u.data* data set. Following the approach pointed in Subsection 3.4.2, for each user e we build a vector with 19 positions $VU_e = (VU_{e1}, VU_{e2}, \ldots, VU_{e19})$ where each component $VU_{ej} \in S_1$ is a linguistic assessment that represents the importance degree of the genre j in the topics of interest related with the user e. These importance degrees are calculated using a weighted average operator:

$$VU_{ej} = \Delta\left(\frac{\sum_{m=1}^{n} r_{em} \cdot g_{mj}}{\sum_{m=1}^{n} r_{em}}\right) \tag{3.17}$$

where r_{em} is the rating assigned by the user e on the movie m and g_{mj} is the value of the genre j for the movie m.

3.5.2.2 *Results of Offline Experiments*

We use the **cross validation** to determine the validity of our model and to analyze the obtained results. Cross validation is typically used to estimate how accurately a predictive model will perform in practice [53]. The data set is divided in complementary subsets, performing the analysis on one subset, called the training set, and validating the analysis on the other subset, called the testing set. To reduce variability, multiple rounds of cross-validation are performed using different partitions, and the validation results are averaged over the different rounds. In k-fold cross validation [53], the original sample is randomly partitioned into k folds. One fold is selected as the testing set, used to estimate the error, and the remaining $k-1$ folds are used as training data set. The cross-validation process is then repeated k times, with each of the k folds used exactly once as the testing set. The k results then can be averaged to produce a single estimation about the deviations between the predictions and the actual ratings.

Values of the folding parameter k commonly assumed are $4, 5, \ldots, 10$. We have chosen a value of $k = 5$. In order to perform 5-fold cross validation, we use the data sets *u1.base* and *u1.test* through *u5.base* and *u5.test* provided by MovieLens which split the collection into 80% for training and 20% for testing, respectively. From the training data sets we build the necessary vectors as we have shown in the previous section. We use them as the input data to predict the unrated ratings. It allows us to measure the system capability in order to predict the users' ratings, calculating the MAE. Besides, to test the effect of the number of neighbors (value of $\aleph_e$ used in the collaborative recommendations) on the accuracy of

Table 3.1 MAE values for our system with MovieLens data sets

$\aleph_e$	u1	u2	u3	u4	u5	Avg MAE
5	0.7405	0.7398	0.7424	0.7432	0.7437	0.7419
10	0.7379	0.7339	0.7353	0.7386	0.7381	0.7368
20	0.7356	0.7351	0.7357	0.7372	0.7370	0.7361
30	0.7434	0.7431	0.7446	0.7452	0.7448	0.7442
50	0.7471	0.7463	0.7468	0.7479	0.7473	0.7471

the system, we have considered the most 5, 10, 20, 30 and 50 similar users. The obtained results are shown in Table 3.1.

As we can see, the performance of the system is quite uniform across the Movielens data set, but considering 10 and 20 similar users we obtain a better average MAE than the rest of configurations. The other three groups (with the most 5, 30 and 50 similar users), present results close to one another. In Table 3.1 we can see that the average MAE increases as the number of neighbors grows or when we consider very few neighbors. When the number of neighbors is between 10 and 20, there is a significant drop in average MAE which indicates a considerable increase in prediction accuracy; in fact, the best average results are obtained considering the most 20 similar users. Therefore, we decide that a number of neighbors between 15 and 20, are the most suitable for our system.

3.5.2.3 *Comparison with other approaches*

In order to compare the results of our system with other, we have implemented several content-based and collaborative models. Firstly, we have implemented a pure content-based approach (CB) [4, 5, 52] in which the similarity between two items is calculated using the cosine measure. We also have implemented the user-based collaborative approach (UBC) [18, 56, 60]. This method uses the ratings of users that are most similar to the target user for predicting the ratings of unrated items; the similarity between users is computed using Pearson's correlation coefficient. Finally, we have implemented the item-based collaborative approach (IBC) [4, 14, 55] in which the similarities of items are used to predict the ratings. The prediction is computed by taking a weighted average of the target user's ratings on similar items. In our experimentation we have used both the cosine and Pearson measure, titled IBC-C and IBC-P, respectively.

To compare the different approaches, we have followed the experimental setting described previously, that is, we perform the 5-fold cross validation, using as training and testing data sets the files *u1.base* and *u1.test* through *u5.base* and *u5.test* provided by

Table 3.2 Average MAE values to compare with other models

	Our system	CB	UBC	IBC-C	IBC-P
Average MAE	0.7412	0.9187	0.7848	0.7705	0.7716
Improvement %		23.94%	5.88%	3.95%	4.10%

MovieLens. With these experiments we calculate the average MAE for all the tests and rounds. To do the comparison, in the case of our system we have used the average MAE for the five values of $\aleph_e$ studied in the previous subsection (see Table 3.1). We prefer to use the average value and not the better MAE, to obtain more significant and realistic results. Table 3.2 presents the MAE results obtained by each approach, where we can see how our system improves the results obtained by the rest of approaches. The row entitled with *Improvement %* presents the improvement percentage obtained with our system over the other approaches.

3.5.3 *Online Experiments*

We have enabled the proposed recommender system for a small group of users, who interact with the system and report us their experience. We test its main features, i.e., its capacities to discover both specialized or complementary resources and collaboration possibilities.

3.5.3.1 *Data Set*

For the online evaluation, we have considered a data set with 200 research resources related with different areas collected by the TTO staff from different information sources. These resources were included into the system following the indications described in Section 3.4.3.1. We assume that the recommender system has to generate recommendations to 15 users and that these users have completed the registration process and evaluated at least 25 resources. From these user assessments, the system is able to build the user profiles.

The resources and the provided user assessments constitute our training data set. Then, we have added 100 new research resources that conform the test data set. The system filtered these 100 resources and it recommended them to the suitable users. To obtain data to compare, these 100 new research resources also were recommended using the advices of the TTO staff.

3.5.3.2 *Results of Online Experiments*

Using the described data set we obtained the contingency table. For example, for user 1, the TTO staff considered 23 relevant resources, of which 16 were specialized and 7 were complementary. Our system selected 26 resources as relevant for user 1, being only 17 really relevant. From these 17 relevant resources, the system classified 15 as specialized and 2 as complementary. Comparing with the recommendations provided by the TTO staff we had 2 resources which are misclassified. So, the success rate for the user 1 was $((17-2)/26) \times 100 = 57.69\%$. Analyzing the contingency table we obtain that the system shows an average precision (success rate) of 61.28%, which is a satisfactory value of face-on system performance.

Similarly, we used the previous scenario to analyze the collaboration possibilities of our recommender system. However, in this case, the items to recommend are not the research resources, but the collaboration opportunities that could appear when the resource is a research project. Thus, we assumed that our system had to recommend research resources to 15 users and a training data set composed by 200 research resources of different areas. Then, we added 100 new resources, of which 30 resources were research projects that constituted the test data set. To compare the collaboration recommendations provided by the system and by the TTO staff we used those 30 projects, and not only the projects considered as relevant by the system or by the TTO staff. So, we can obtain specific measures with regard to collaboration recommendations.

Then, for the 30 projects we compared the collaboration recommendations made by the system with the collaboration recommendations provided by the TTO staff. We classified the collaboration recommendations taking into account the categorization described in Table 3.3. To understand the meaning of this table we provide the following example. Suppose that for project 1, the TTO staff selected user 7 and indicated him/her that he/she could collaborate with users 2, 11 and 12 to develop the project. Our system also selected user 7 for project 1, but in this case it recommended the collaboration with users 2, 3 and 12, and therefore, these recommendations didn't match with the TTO staff recommendations. That is, our system presented 2 hits (for users 2 and 12), 1 failure (user 3) and a non-detected collaboration (user 11). Then, for project 1, $Nchs = 2$, $Nchn = 1$ and $Ncfs = 1$.

Assuming this framework, we obtained the Table 3.4 for the 30 projects, being the average precision of 70.44%, the average recall of 72.50% and an average F1 of 70.40%, which show a satisfactory behavior of our system. The obtained results indicate that the

Table 3.3 Contingency table for the collaboration recommendations

	Selected	Not selected	Total
Considered by TTONot considered by TTO	NchsNcfs	NchnNcfn	NchNcf
Total	Ns	Nn	N

Table 3.4 Contingency table for the collaboration recommendations

Project	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Nchs	2	3	4	2	4	2	3	2	3	3	2	2	3	2	2
Nchn	1	1	2	1	1	0	1	1	1	2	2	1	1	0	1
Ncfs	1	1	1	2	2	1	1	1	1	1	1	1	0	2	1
Nch	3	4	6	3	5	2	4	3	4	5	4	3	4	2	3
Ns	3	4	5	4	6	3	4	3	4	4	3	3	3	4	3
Project	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Nchs	3	2	2	3	4	3	2	3	3	3	2	3	2	3	1
Nchn	2	1	1	1	1	1	1	2	1	2	0	1	0	1	1
Ncfs	1	1	0	1	1	1	1	1	2	1	1	1	1	2	1
Nch	5	3	3	4	5	4	3	5	4	5	2	4	2	4	2
Ns	4	3	2	4	5	4	3	4	5	4	3	4	3	5	2

collaboration recommendations provided by our system are useful to researchers, and quite similar to those provided by the TTO staff.

3.6 Concluding Remarks

The TTO is responsible for putting into action and managing the activities which generate knowledge and technical and scientific collaboration. A service that is particularly important to fulfill this objective is the selective dissemination of information about research resources. The TTO staff and researchers need tools to assist them in their processes of information discovering because of the large amount of information available on these systems.

We have presented a fuzzy linguistic recommender system to spread selectively research resources in a TTO. Particularly, we propose to use a hybrid approach as recommendation engine, that is, integrating a content-based approach with a collaborative one, in order to take the advantages of both strategies and reduce the disadvantages of each one of them. This recommender system recommends specialized resources, complementary resources and collaboration possibilities that allows the researchers to meet other researchers and to form multidisciplinary groups. Besides, the system improves the feedback process using satisfaction degrees. We have applied our research in a real environment provided by the TTO. The system advices researchers and environment companies about resources that could be interesting for them and collaboration possibilities with other researchers. The experimental results show us that the recommendations provided by our system are useful to researchers.

Analyzing our system, we could conclude that its main limitation is the need for interaction with TTO staff to establish the internal representations for the user profiles and the items. With regard to future research, we believe that a promising direction is to study automatic techniques to establish the representation of user profiles and items. Moreover, we want to explore new improvements of the recommendation approach, exploring new methodologies for the generation of recommendations, as for example, bibliometric tools to enrich the information on the researchers and research resources [1, 13].

Acknowledgments

This book chapter has been developed with the financing of Projects 90/07, TIN2007-61079, PET2007-0460, TIN2010-17876, TIC5299 and TIC-5991.

Bibliography

[1] S. Alonso, F. Cabrerizo, E. Herrera-Viedma, F. Herrera, hg-index: A new index to characterize the scientific output of researchers based on the h- and g- indices, *Scientometrics* **82** (2), 391–400 (2010).

[2] S. Alonso, E. Herrera-Viedma, F. Chiclana, F. Herrera, A web based consensus support system for group decision making problems and incomplete preferences, *Information Sciences* **180** (23), 4477–4495 (2010).

[3] B. Arfi, Fuzzy decision making in politics. a linguistic fuzzy-set approach (lfsa), *Political Analysis* **13** (1), 23–56 (2005).

[4] A. Barragáns-Martínez, E. Costa-Montenegro, J. Burguillo, M. Rey-López, F. Mikic-Fonte, A. Peleteiro, A hybrid content-based and item-based collaborative filtering approach to recommend tv programs enhanced with singular value decomposition, Information Sciences **180** (22), 4290–4311 (2010).

[5] C. Basu, H. Hirsh, W. Cohen, Recommendation as classification: Using social and content-based information in recommendation, in: *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, 1998.

[6] R. Burke, Hybrid recommender systems: Survey and experiments, *User Modeling and User-Adapted Interaction* **12**, 331–370 (2002).

[7] R. Burke, Hybrid web recommender systems, P. Brusilovsky, A. Kobsa, and W. Nejdl (Eds.): *The Adaptive Web*, LNCS **4321**, 377–408 (2007).

[8] F. Cabrerizo, J. López-Gijón, A. Ruíz-Rodríguez, E. Herrera-Viedma, A model based on fuzzy linguistic information to evaluate the quality of digital libraries, *International Journal of Information Technology & Decision Making* **9** (3), 455–472 (2010).

[9] F. Cabrerizo, I. Pérez, E. Herrera-Viedma, Managing the consensus in group decision making in an unbalanced fuzzy linguistic context with incomplete information, *Knowledge-Based Systems* **23** (2), 169–181 (2010).

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

[11] M. Claypool, A. Gokhale, T. Miranda, Combining content-based and collaborative filters in an online newpaper, in: *Proceedings of the ACM SIGIR-99 Workshop on Recommender Systems-Implementation and Evaluation*, 1999.

[12] C. Cleverdon, E. Keen, Factors determining the performance of indexing systems, *ASLIB Cranfield Research Project 2*.

[13] M. Cobo, A. López-Herrera, E. Herrera-Viedma, F. Herrera, An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field, *Journal of Informetrics* **5** (1), 146–166 (2011).

[14] M. Deshpande, G. Karypis, Item-based top-n recommendation algorithms, ACM Transactions on Information Systems **22**, 143–177 (2004).

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

[16] N. Good, J. Schafer, J. Konstan, A. Borchers, B. Sarwar, J. Herlocker, J. Riedl, Combining collaborative filtering with personal agents for better recommendations, in: *Proceedings of the Sixteenth National Conference on Artificial Intelligence*, 1999.

[17] U. Hanani, B. Shapira, P. Shoval, Information filtering: Overview of issues, research and systems, *User Modeling and User-Adapted Interaction* **11**, 203–259 (2001).

[18] J. Herlocker, J. Konstan, A. Borchers, J. Riedl, An algorithmic framework for performing collaborative filtering, in: *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1999.

[19] J. Herlocker, J. Konstan, L. Terveen, J. Riedl, Evaluating collaborative filtering recommender systems, *ACM Transactions on Information Systems* **22** (1), 5–53 (2004).

[20] F. Herrera, E. Herrera-Viedma, Aggregation operators for linguistic weighted information, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems* **27**, 646–656 (1997).

[21] F. Herrera, E. Herrera-Viedma, Choice functions and mechanisms for linguistic preference relations, *European Journal of Operational Research* **120**, 144–161 (2000).

[22] F. Herrera, E. Herrera-Viedma, L. Martínez, A fusion approach for managing multi-granularity linguistic term sets in decision making, *Fuzzy Sets and Systems* **114**, 43–58 (2000).

[23] F. Herrera, E. Herrera-Viedma, J. Verdegay, Direct approach processes in group decision making using linguistic OWA operators, *Fuzzy Sets and Systems* **79**, 175–190 (1996).

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

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

[26] E. Herrera-Viedma, 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 (2001).

[27] E. Herrera-Viedma, Modeling the retrieval process of an information retrieval system using an ordinal fuzzy linguistic approach, *Journal of the American Society for Information Science and Technology* **52** (6), 460–475 (2001).

[28] E. Herrera-Viedma, O. Cordón, M. Luque, A. López, A. Muñoz, A model of fuzzy linguistic IRS based on multi-granular linguistic information, *International Journal of Approximate Reasoning* **34** (3), 221–239 (2003).

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

[30] E. Herrera-Viedma, A. López-Herrera, A review on information accessing systems based on fuzzy linguistic modelling, *International Journal of Computational Intelligence Systems* **3** (4), 420–437 (2010).

[31] E. Herrera-Viedma, A. López-Herrera, M. Luque, C. Porcel, 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 (2007).

[32] E. Herrera-Viedma, A. López-Herrera, C. Porcel, Tuning the matching function for a threshold weighting semantics in a linguistic information retrieval system, *International Journal of Intelligent Systems* **20** (9), 921–937 (2005).

[33] E. Herrera-Viedma, A. G. López-Herrera, S. Alonso, J. Moreno, F. Cabrerizo, C. Porcel, A computer-supported learning system to help teachers to teach fuzzy information retrieval systems, *Information Retrieval* **12** (2), 179–200 (2009).

[34] E. Herrera-Viedma, L. Martínez, F. Mata, F. Chiclana, 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 (2005).

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

[36] M. Hsu, A personalized english learning recommender system for esl students, *Expert Systems with Applications* **34**, 377–408 (2008).

[37] Y. Jiang, Z. Fan, J. Ma, A method for group decision making with multi-granularity linguistic assessment information, *Information Sciences* **178** (41), 1098–1109 (2008).

[38] A. Jun, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, *Information Sciences* **178** (1), 37–51 (2008).

[39] R. Korfhage, *Information Storage and Retrieval*, New York: Wiley Computer Publishing, 1997.

[40] S. Lee, Y. Cho, S. Kim, Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations, *Information Sciences* **180** (11), 2142–2155 (2010).

[41] C. Long-Sheng, H. Fei-Hao, C. Mu-Chen, H. Yuan-Chia, Developing recommender systems with the consideration of product profitability for sellers, *Information Sciences* **178** (4), 1032–1048 (2008).

[42] A. López-Herrera, E. Herrera-Viedma, F. Herrera, Applying multi-objective evolutionary algorithms to the automatic learning of extended boolean queries in fuzzy ordinal linguistic information retrieval systems, *Fuzzy Sets and Systems* **160**, 2192–2205 (2009).

[43] I. Macho-Stadler, D. Perez-Castrillo, and R. Veugelers, Licensing of university inventions: The role of a technology transfer office, *International Journal of Industrial Organization* **25**, 483–510 (2007).

[44] G. Marchionini, Research and development in digital libraries (2000). URL: `http://ils.unc.edu/~march/digital_library_R_and_D.html`

[45] G. Markman, P. Phan, D. Balkin, and P. Gianiodis, Entrepreneurship and university-based technology transfer, *Journal of Business Venturing* **20**, 241–263 (2005).

[46] F. Mata, L. Martínez, E. Herrera-Viedma, 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 (2009).

[47] J. Morales-del Castillo, E. Peis, A. Ruíz-Rodríguez, E. Herrera-Viedma, Recommending biomedical resources: A fuzzy linguistic approach based on semantic web, *International Journal of Intelligent System* **25**, 1143–1157 (2010).

[48] J. Moreno, J. Morales-del Castillo, C. Porcel, E. Herrera-Viedma, A quality evaluation methodology for health-related websites based on a 2-tuple fuzzy linguistic approach, *Soft Computing* **14** (8), 887–897 (2010).

[49] Movielens – movie recommendations. URL: `http://movielens.umn.edu`

[50] E. Nasibov, A. Kinay, An iterative approach for estimation of student performances based on linguistic evaluations, *Information Sciences* **179** (5), 688–698 (2009).

[51] M. Pazzani, A framework for collaborative, content-based and demographic filtering, *AI Review* **13** (56), 393–408 (1999).

[52] A. Popescul, L. Ungar, D. Pennock, S. Lawrence, Probabilistic models for unified collaborative and content-based recommendation in sparce-data environments, in: *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence* (*UAI*), San Francisco, 2001.

[53] P. Refaeilzadeh, L. Tang, H. Liu, *Cross-validation* (2008). URL: `http://www.public.asu.edu/~ltang9/papers/ency-cross-validation.pdf`

[54] P. Reisnick, H. Varian, Recommender systems, *Special issue of Communications of the ACM* **40** (3), 56–59 (1997).

[55] B. Sarwar, G. Karypis, J. Konstan, R. J., Item-based collaborative filtering recommendation algorithms, in: *Proceedings of the ACM World Wide Web Conference*, 2001.

[56] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Analysis of recommendation algorithms for ecomerce, *Proceedings of ACM E-Commerce 2000 conference*, 158–167 (2000).

[57] J. Serrano-Guerrero, E. Herrera-Viedma, J. Olivas, A. Cerezo, F. Romero, A google wave-based fuzzy recommender system to disseminate information in university digital libraries 2.0, *Information Sciences* **181**, 1503–1516 (2011).

[58] G. Shani, A. Gunawardana, Recommender Systems Handbook, chap. *Evaluating Recommendation Systems*, Ricci, F., Rokach, L., Shapira, B., and Kantor, P.B. Eds. (Springer), 257–298 (2011).

[59] D. Siegel, D.Waldman, L. Atwater, A. Link, Toward a model of the effective transfer of scientific knowledge from academicians to practitioners: qualitative evidence from the commercialization of university technologies, *Journal of Engineering and Technology Management*, 115–142 (2004).

[60] P. Symeonidis, A. Nanopoulos, A. Papadopoulos, Y. Manolopoulos, Collaborative recommender systems: Combining effectiveness and efficiency, *Expert Systems with Applications* **34** (4), 2995–3013 (2008).

[61] I. Truck, H. Akdag, A tool for aggregation with words, *Information Sciences* **179** (14), 2317–2324 (2009).

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

[63] A. Zenebe, A. Nocio, Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems, *Fuzzy Sets and Systems* **160** (1), 76–94 (2009).

[64] G. Zhang, J. Lu, A linguistic intelligent user guide for method selection in multi-objective decision support systems, *Information Sciences* **179** (14), 2299–2308 (2009).