



A quality based recommender system to disseminate information in a university digital library



Álvaro Tejeda-Lorente^{a,*}, Carlos Porcel^{b,*}, Eduardo Peis^c, Rosa Sanz^c, Enrique Herrera-Viedma^{a,*}

^a University of Granada, Department of Computer Science and Artificial Intelligence, Granada, Spain

^b University of Jaén, Department of Computer Science, Jaén, Spain

^c Department of Librarianship and Information Science, Granada, Spain

ARTICLE INFO

Article history:

Received 29 October 2012

Received in revised form 7 October 2013

Accepted 27 October 2013

Available online 8 November 2013

Keywords:

Recommender system

Item quality

Fuzzy linguistic modeling

Digital library

ABSTRACT

Recommender systems evaluate and filter the great amount of information available on the Web, so they could be used in an academic environment to help users in their searches of relevant information. In the literature, a lot of approaches for generating personalized recommendations of information items in such environment can be found. Usually, these approaches use user profiles and/or features of items to predict those relevant items, but they do not take into account the quality of these items. To overcome this problem, in this paper we propose a new recommender system based on quality. This system uses the quality of the items to estimate their relevance. The system measures the item quality and takes into account this measure like a new factor to be considered in the recommendation process. In such a way, we present a recommender system based on items' quality, to help users to access relevant research resources. This recommender systems is developed by using a fuzzy linguistic approach and it has been tested satisfactorily in a university digital library.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Nowadays we live in the Information Society and we are constantly bombarded with many information in all scopes of our lives. This problem has become more widely recognized and experienced because of the rapid advances made in Information and Communication Technologies [19].

World Wide Web (called Web) is a popular and interactive medium to collect, disseminate and access an increasingly huge amount of information. Due to its spectacular growth, related to both resources (pages, sites, and services) and visitors, the Web is nowadays the main information repository. This great amount of information introduces noise in our information accesses and this makes difficult to find relevant information and affects too to our decisions. For example, everyday we receive in our accounts a huge amount of emails. Most of them are qualified as spam, but we also receive a big number of emails containing useful or relevant information. The problem is that this fact could provoke 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 of key importance [44].

This explosive growth of the Web stimulates the development of fast and effective automated systems that support an easy and effective access to the information relevant [38]. *Digital libraries*, where the information is generated much faster

* Corresponding authors. Tel.: +34 953213017 (C. Porcel).

E-mail addresses: atejeda@decsai.ugr.es (Á. Tejeda-Lorente), cporcel@ujaen.es (C. Porcel), epais@ugr.es (E. Peis), rosasanz@ugr.es (R. Sanz), viedma@decsai.ugr.es (E. Herrera-Viedma).

than the users can process it, are one of these automated systems [20,31,42,53]. 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 on such information can be varied and can be offered to individuals or user communities [9,20,52]. 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), added value 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 [12].

A service that is particularly important in a UDL is the selective dissemination of information or filtering. 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 [42]. Because of information overload problem it is often difficult to obtain useful or relevant information when it is necessary. Therefore, many times when the users of a UDL try to receive useful information, however they obtain irrelevant and unnecessary information. So, users need easier access to the thousands of resources that are available but yet hard to find [44].

Mainly, in the Web we can find two different tools to facilitate the access to the information: Information Retrieval Systems [37] and Recommender Systems [4,22,51]. The former are focused on information search in a known content repository, while the later are focused on information discovery in partially known frameworks. A *recommender system* attempts to discover information items (movies, music, books, news, images, web pages, papers and so on) that are likely of interest to a user. Recommender systems are especially useful when they identify information that a person was previously unaware of. Furthermore, recommender systems are personalized services because they may treat each user in a different way. They are becoming popular tools for reducing information overload and to improve the sales in e-commerce web sites [7,34,51]. These recommender systems play an important role in highly rated Web sites, such as Amazon,¹ YouTube,² Netflix,³ Tripadvisor,⁴ Last.fm⁵ or IMDb⁶ [13,14].

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. But the way in which this information is acquired and exploited depends on the particular recommendation approach. The system could acquired implicit information about the users analyzing the users behavior, or the system might request the users insert explicitly their preferences. Another question to consider is what additional information is required by the system, and how this information is processed and managed to generate a list of personalized recommendations.

One of the most used method to generate the recommendations is the *collaborative approach* [26] in which the recommendations to a user are based on other user recommendations with similar user profiles, taking into account the ratings provided by the users. Another widely used approach for its simplicity is the *content-based approach* [40], which generates the recommendations taking into account the items' features and the user past experience dealing with similar items. All approaches have their advantages and disadvantages, so that a widely used solution is the *hybrid approach* [4]. The hybrid approaches combine both previous approaches to reduce the disadvantages of each one of them and exploit their benefits.

In previous proposals we have applied recommender systems in the UDL scope [49,50]. But, despite that the use of these techniques to avoid the information overload problem was successful, we have found different aspects that may limit their performance. They act as an information retrieval system based on matching functions which acts among the resources representation and user profiles. In fact, we do not use a collaborative approach and this limits their performance. Furthermore, as it happens in the Web, the number of electronic resources daily generated grows continuously, so the problem appears again and the system performance is decreased.

Hence, a first step to improve these systems is to adapt the recommendation approach to the new circumstances. If we analyze the UDL scope, we find that the collaborative approach would be 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 could 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., when the system has few ratings, what is known as cold-start problem. For this reason, we propose to combine both the content-based and collaborative approaches to obtain a hybrid recommendation scheme that improves the performance of those systems proposed in [49,50]. On the other hand, in the real life, people usually buy widely known products or products of popular brands. These products are popular because they are considered to have good quality in order to satisfy the users' needs. We could apply a similar

¹ <http://www.amazon.es/>.

² www.youtube.com/.

³ www.netflix.com/.

⁴ www.tripadvisor.es/.

⁵ www.lastfm.es/.

⁶ www.imdb.com/.

idea in the UDL framework. Generated recommendations could be more interesting for the users if we consider the quality of items themselves. That is, to compute the relevance of an item using its quality.

In this paper, we present a new recommender system whose main novelties are the following:

- The system implements a *hybrid recommendation strategy* based on 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 [55].
- The system implements a richer *feedback process*. In [49,50] the user participation in the feedback process is small because it consists in adding or eliminating topics in the user profile, but users could not provide satisfaction degrees. However, to improve the recommendations in this new system, 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.
- Now we face the recommendations generation process about research resources as a task with two distinct elements: On the one hand, finding research resources that are relevant to the users and on the other hand, finding valid research resources from the standpoint of the *quality of the items*.
- The new system incorporates a new module which performs a *re-ranking* process which takes into account the estimated relevance of an item along with the item quality.

But the problem is how to obtain the research resource quality without much interaction from users. We cannot use previous proposals like the presented in [8,12,32,33] because they require intensive user's feedback to evaluate the quality. So, we propose a new way to evaluate the quality of research resources based on the idea that if a research resource is usually preferred to others, indicate us that such item has a certain quality. To do that, we work from the method presented in [49], where we proposed an alternative way to obtain accurate and useful knowledge about the user preferences: The system allows users to provide their preferences by means of incomplete fuzzy linguistic preference relations [1,43], and in such a way, it facilitates users the expression of their preferences and, consequently, the determination of user profiles. The system completes the incomplete preference relations using the tools proposed in [1,3], and it calculates the resources quality using this preference relation. To aggregate the estimated relevance of a research resource along with its quality score in a single score, fuzzy linguistic operators are used [27]. Then, in this paper we present a hybrid fuzzy linguistic recommender system based on items' quality and we apply it in a UDL to help the users to access relevant research resources. The system measures the items' quality and it takes into account this measure like a new factor to be considered in the recommendation process (specifically in a re-ranking module). This system improves the generated recommendations, by including more useful and accurate recommendations, and by increasing its information discovering properties in an academic environment.

The rest of this paper is set out as follows. In Section 2, the background is presented, that is the basis of recommender systems, the fuzzy linguistic modeling to represent information and some aspects on quality evaluation in digital libraries are discussed. Section 3 describes the new recommendation approach based on quality of the items. Section 4 presents the evaluation of the system and the experimental results. Finally, some concluding remarks are pointed out in Section 5.

2. Background

2.1. Basis of recommender systems

Recommender systems help users in the effective identification of items suiting their wishes, needs or preferences. They have the effect of guiding the users in a personalized way to access 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 improving the sales in e-commerce web sites [7,17,18,34,39,51].

Automatic filtering services differ from retrieval services. In filtering the corpus changes continuously, the users have long time information needs (described by mean of user profiles), and the objective is to remove irrelevant data from incoming streams of data items [17,22,42,51]. On the contrary, retrieval services use queries which are introduced by the users into the system to retrieve relevant items. Thus, a result from a recommender system is understood as a recommendation, an option worthy or consideration, while a result from an information retrieval system is interpreted as a match to the user's query [7].

In a recommender system, the users' preferences about research resources 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'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. Two different ways to obtain information about user preferences are distinguished [22], although many systems adopt a hybrid approach:

- The *implicit approach* is implemented by inference from some kind of observation. The observation is applied to the user behavior, such as the 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 and ratings about the explored items. This approach is currently the most common one.

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 [7,21,22,48,51]:

- *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 [5,16]. These recommender systems tend to fail when little is known about the user information needs. This is called the new user cold-starting problem [36].
- *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 [46].
- *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 have 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. This idea is the base of *hybrid recommender approaches*, which combine several algorithms or recommendation approaches [6]. Using a hybrid strategy, users are provided with more accurate recommendations than those offered by each strategy individually [5,16,21]. For this reason, in this paper we propose the use of a hybrid approach. In [6], Burke developed an interesting survey about the possible strategies used to obtain hybrid recommender systems. He proposed seven hybridization strategies to combine the recommendation approaches. One possible strategy consists in incorporate aspects of several recommendation approaches in one approach. Other strategy is that two or more recommendation approaches work in parallel and then the results are combined, or that the recommendations approaches be connected so that the output of one approach serves as input for the next one. Of all, for our purpose we find it interesting that knowns as *switching hybrid strategy* [6], which chooses among different recommendation components and applies the selected one following some criterion. That is, it switches from one recommendation approach to another according to the situation of the system at a given moment. For instance this situation could be given for the number of ratings available to generate recommendations. We could use the collaborative approach, but if there are not enough ratings, the cold-start problem arises. Therefore given this situation it is better to switch to the content-based approach.

Finally, 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 [22,51].

2.2. Fuzzy linguistic approach

In this subsection we present the fuzzy linguistic approach used in our recommender system. The use of Fuzzy Sets Theory has given very good results to model qualitative information [57] and it has been proven to be useful in many problems. It is a tool based on the concept of *linguistic variable* proposed by Zadeh [57].

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/her knowledge will have more or less terms [47].

2.2.1. The 2-tuple fuzzy linguistic approach

The 2-tuple fuzzy linguistic modeling [29] 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 [27,57]). 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 $S = \{s_0, \dots, s_g\}$ 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 $\mu_{s_i} : [0, 1] \rightarrow [0, 1]$, and we consider all terms

distributed on a scale on which a total order is defined, that is, $s_i \leq s_j \iff i \leq j$. We consider that linear triangular membership functions are good enough to capture the vagueness of those linguistic assessments, since it may be impossible or unnecessary to obtain more accurate values. This representation is achieved by the 3-tuple (a, b, c) , where a is the point where the membership is 1 and b and c are the left and right limits of the definition domain of the triangular membership function. For example, the following semantics, represented in Fig. 1, can be assigned to a set of seven terms via triangular membership functions:

$$\begin{aligned}
 P = \text{Perfect} &= (1, 0.83, 1) & VH = \text{Very_High} &= (0.83, 0.67, 1) \\
 H = \text{High} &= (0.67, 0.5, 0.83) & M = \text{Medium} &= (0.5, 0.33, 0.67) \\
 L = \text{Low} &= (0.33, 0.17, 0.5) & VL = \text{Very_Low} &= (0.17, 0, 0.33) \\
 N = \text{None} &= (0, 0, 0.17)
 \end{aligned}$$

In this fuzzy linguistic context, if a symbolic method [27] aggregating linguistic information obtains a value $\beta \in [0, g]$, and $\beta \notin \{0, \dots, g\}$, then an approximation function is used to express the result in S .

Definition 1. [29] 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 $\alpha \in [-.5, .5]$ 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 2. [29] Let $S = \{s_0, \dots, s_g\}$ 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:

$$\Delta : [0, g] \longrightarrow S \times [-0.5, 0.5] \tag{1}$$

$$\Delta(\beta) = (s_i, \alpha), \text{ with } \begin{cases} s_i & i = \text{round}(\beta) \\ \alpha = \beta - i & \alpha \in [-.5, .5] \end{cases} \tag{2}$$

where $\text{round}(\cdot)$ 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 following operators:

1. Negation operator: $\text{Neg}((s_i, \alpha)) = \Delta(g - (\Delta^{-1}(s_i, \alpha)))$.
2. Comparison of 2-tuples (s_k, α_1) and (s_l, α_2) :
 - If $k < l$ then (s_k, α_1) is smaller than (s_l, α_2) .
 - If $k = l$ then
 - (a) if $\alpha_1 = \alpha_2$ then (s_k, α_1) and (s_l, α_2) represent the same information,
 - (b) if $\alpha_1 < \alpha_2$ then (s_k, α_1) is smaller than (s_l, α_2) ,
 - (c) if $\alpha_1 > \alpha_2$ then (s_k, α_1) is bigger than (s_l, α_2) .

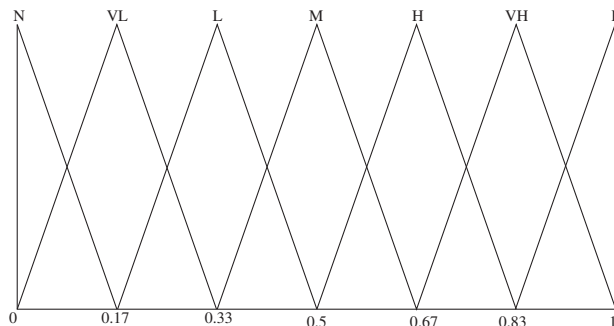


Fig. 1. A set of seven linguistic terms with its semantics.

3. Aggregation operators [58]. The aggregation of information consists of obtaining a value that summarizes a set of values, therefore, the result of the aggregation of a set of 2-tuples must be a 2-tuple. In the literature we can find many aggregation operators which allow us to combine the information according to different criteria. Using functions Δ and Δ^{-1} that transform without loss of information numerical values into linguistic 2-tuples and viceversa, any of the existing aggregation operators can be easily extended to deal with linguistic 2-tuples. Some examples are:

Definition 3. Arithmetic Mean. Let $x = \{(r_1, \alpha_1), \dots, (r_n, \alpha_n)\}$ be a set of linguistic 2-tuples, the 2-tuple arithmetic mean \bar{x}^e is computed as:

$$\bar{x}^e[(r_1, \alpha_1), \dots, (r_n, \alpha_n)] = \Delta\left(\sum_{i=1}^n \frac{1}{n} \Delta^{-1}(r_i, \alpha_i)\right) = \Delta\left(\frac{1}{n} \sum_{i=1}^n \beta_i\right). \tag{3}$$

Definition 4. Weighted Average Operator. Let $x = \{(r_1, \alpha_1), \dots, (r_n, \alpha_n)\}$ be a set of linguistic 2-tuples and $W = \{w_1, \dots, w_n\}$ be their associated weights. The 2-tuple weighted average \bar{x}^w is:

$$\bar{x}^w[(r_1, \alpha_1), \dots, (r_n, \alpha_n)] = \Delta\left(\frac{\sum_{i=1}^n \Delta^{-1}(r_i, \alpha_i) \cdot w_i}{\sum_{i=1}^n w_i}\right) = \Delta\left(\frac{\sum_{i=1}^n \beta_i \cdot w_i}{\sum_{i=1}^n w_i}\right). \tag{4}$$

Definition 5. Linguistic Weighted Average Operator. Let $x = \{(r_1, \alpha_1), \dots, (r_n, \alpha_n)\}$ be a set of linguistic 2-tuples and $W = \{(w_1, \alpha_1^w), \dots, (w_n, \alpha_n^w)\}$ be their linguistic 2-tuple associated weights. The 2-tuple linguistic weighted average \bar{x}_l^w is:

$$\bar{x}_l^w[((r_1, \alpha_1), (w_1, \alpha_1^w)) \dots ((r_n, \alpha_n), (w_n, \alpha_n^w))] = \Delta\left(\frac{\sum_{i=1}^n \beta_i \cdot \beta_{w_i}}{\sum_{i=1}^n \beta_{w_i}}\right), \tag{5}$$

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

2.2.2. Linguistic hierarchy to model multi-granular linguistic information

When different experts have different uncertainty degrees on the phenomenon, then several linguistic term sets with a different granularity of uncertainty are necessary [28]. The use of different label sets to assess information is also necessary when an expert has to evaluate different concepts. In such situations, we need tools to manage multi-granular linguistic information [30].

In [30] a multi-granular fuzzy linguistic modeling based on a 2-tuple fuzzy linguistic approach and the concept of linguistic hierarchy was proposed. A *Linguistic Hierarchy, LH*, is a set of levels $l(t, n(t))$, i.e., $LH = \cup_t l(t, n(t))$, 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: $l(t, n(t)) \rightarrow l(t + 1, 2 \cdot n(t) - 1)$. Table 1 shows the granularity needed in each linguistic term set of the level t depending on the value $n(t)$ defined in the first level (3 and 7 respectively). A graphical example of a linguistic hierarchy is shown in Fig. 2.

In [30] was demonstrated that linguistic hierarchies are useful to represent multi-granular linguistic information and that they allow to combine multi-granular linguistic information without loss of information. To do this, a family of transformation functions between labels from different levels was defined:

Definition 6. Let $LH = \cup_t l(t, n(t))$ be a linguistic hierarchy whose linguistic term sets are denoted as $S^{n(t)} = \{s_0^{n(t)}, \dots, s_{n(t)-1}^{n(t)}\}$. The transformation function between a 2-tuple that belongs to level t and another 2-tuple in level $t' \neq t$ is defined as:

$$TF_{t'}^t : l(t, n(t)) \rightarrow l(t', n(t')) \tag{6}$$

$$TF_{t'}^t(s_i^{n(t)}, \alpha^{n(t)}) = \Delta\left(\frac{\Delta^{-1}(s_i^{n(t)}, \alpha^{n(t)}) \cdot (n(t') - 1)}{n(t) - 1}\right). \tag{7}$$

Table 1
Linguistic hierarchies.

	Level 1	Level 2	Level 3
$l(t, n(t))$	$l(1, 3)$	$l(2, 5)$	$l(3, 9)$
$l(t, n(t))$	$l(1, 7)$	$l(2, 13)$	

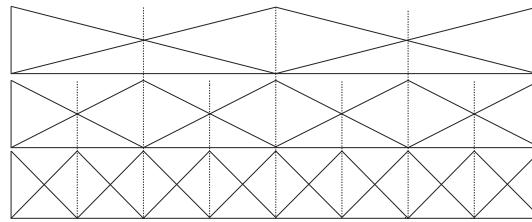


Fig. 2. Linguistic hierarchy of 3, 5 and 9 labels.

As it was pointed out in [30] 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. To define the computational model, we select a level to make uniform the information (for instance, the highest granularity level) and then we can use the operators defined in the 2-tuple fuzzy linguistic approach.

2.2.3. Incomplete fuzzy preference relations

Definition 7. A fuzzy preference relation P on a set of alternatives $X = \{x_1, \dots, x_n\}$ is a fuzzy set on the product set $X \times X$, i.e., it is characterized by a membership function $\mu_p : X \times X \rightarrow [0, 1]$.

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 .

As our system integrates the multi-granular FLM based on 2-tuples, so we must define a linguistic preference relation as follows.

Definition 8. Let $X = \{x_1, \dots, x_n\}$ 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:

$$\mu_p : X \times X \rightarrow S \times [0.5, 0.5] \quad (8)$$

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 .

However, in many 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 [1,3,43].

Definition 9. A function $f : X \rightarrow Y$ 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 10. 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.

2.3. On evaluation of quality in digital libraries

Digital libraries are relatively new and innovative information systems, under constant development and change, and, therefore, evaluation processes are of critical importance to ensure not only their correct evolution, but also their acceptance by the users and application communities. In the digital libraries scope, the objective of the evaluation is to assess to what extent a digital library meets its objectives and offer suggestions for improvements [15]. Digital libraries evaluation has many facets depending on the characteristics and the perspective of the evaluating agent.

Different approaches to evaluate the success of a digital library have been studied [35,24,23,41] involving users, collections, and systems, aimed at identifying generalizable metrics or context specific methods. The most recognized digital libraries evaluation criteria are derived from evaluation criteria for traditional libraries, information retrieval system performance and human–computer interaction [15,41]. Very few studies actually apply all the digital evaluation criteria to assess a digital library. Many of the studies focus on the evaluation of usability of digital libraries. After reviewing usability tests in selected academic digital libraries, Jeng [35] found that ease of use, satisfaction, efficiency, and effectiveness are the

main applied criteria. Some of the evaluation studies extend to assess performance, content and services of digital libraries while service evaluation mainly concentrates on digital reference [11]. Other evaluation studies also look into the impact of digital libraries [41].

As we can see, very few studies actually apply all the digital evaluation criteria to assess not the digital library, but the resources, that is, the documents. However, as much of the success of a digital library depends on the provided resources are interesting for the users, the user satisfaction is fundamental to evaluate a digital library [24,23]. Some developments using this idea can be found in [33].

3. A new recommender system that combines relevance and quality of items to recommend research resources in a UDL

In this paper we face the recommendations generation process about research resources as a task with two elements to consider: finding research resources that are relevant to the UDL users and, on the other hand, finding valid research resources from the standpoint of the item quality. We propose a new approach to combine the estimated relevance of an item along with the item quality.

We work from the method presented in [49], where we proposed an alternative way to obtain accurate and useful knowledge about the user preferences. The system allows users to provide their preferences by means of incomplete fuzzy linguistic preference relations [1,43], and in such a way, we facilitate users the expression of their preferences and, consequently, the determination of user profiles. The system completes the incomplete preference relations using the tools proposed in [1,3]. However, the recommendation strategy applied in [49] is a simple content-based approach. Now we improve that recommender system by implementing a hybrid recommendation strategy based on 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. With this dual perspective, we get minimize the cold-start problem because our system switch from one approach to another, depending on the circumstances. On the other hand, users do not need to provide their preferences about all resources, but the system presents them only a selection of the most representative resources, and the users only provide their preferences about that reduced number of resources.

Moreover, we incorporate a new module which performs a *re-ranking*, combining the estimated relevance of an item along with the item quality. The idea is that resources usually preferred to other are considered to have good quality in order to satisfy the users' needs. In such a way, the recommender system generates more useful and accurate recommendations. But these improvements do not affect the complexity of the recommendation process because we simply substitute the previous recommendation scheme with other with similar complexity. Furthermore, this proposal affects only the selected resources, which is a reduced number of items, and thus, it does not affect the overall complexity.

The proposed model is applied in a UDL to improve the recommendations generated by the system and help users to access relevant research resources. In Fig. 3 we show the basic operating scheme of this recommender system which is based on the following components:

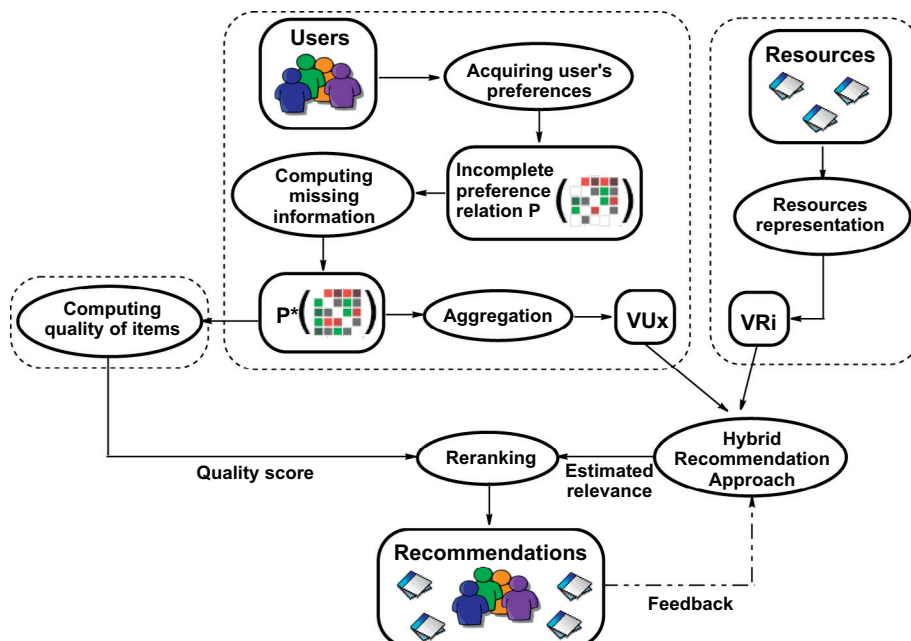


Fig. 3. Basic operating scheme.

1. *Resources representation.* The system obtains an internal representation of the resources based on their scopes.
2. *User profiles representation.* The system obtains an internal representation of the user based on their preferred research resources and topics of interest.
3. *Hybrid recommendation approach.* The system generates the recommendations according to the hybrid filtering approach.
4. *Computing quality of items.* From the preference matrix about research resources, the system obtains a quality score for each item.
5. *Re-ranking.* The system aggregates the estimated relevance of a research resource and its quality score in a single score.
6. *Feedback phase.* The users provides the system their opinions about the received recommendations.

As aforementioned, we work with a multi-granular fuzzy linguistic approach [28,30] to model the user-system communication in order to allow a higher flexibility in the communication processes of the system. The system uses different label sets (S_1, S_2, \dots) 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., $S_i \in LH$. 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 take into account the following concepts that can be assessed in the system:

- **Importance degree** of a discipline with respect to a resource scope, which is assessed in S_1 . This degree is used to obtain an internal representation about the research resources.
- **Similarity degree** among resources or among users, which is assessed in S_2 .
- Predicted **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 .
- **Preference degree** of a resource regarding another one, which is assessed in S_5 .

The selected granularity must be sufficiently low as not to impose an excessive precision in the information you want to express and high enough to get a discrimination of the assessments in a limited number of degrees. Usually, the cardinality used in the linguistic models is an odd value, such as 7 or 9, not exceeding 11 labels. These classical values are based on Miller's observation line about human capacity [45], which indicated that can be handled reasonably and remember about 7 or 9 terms.

Following the linguistic hierarchy shown in Fig. 2, we use the level 2 (5 labels) to represent the importance degrees and preference degrees ($S_1 = S^5$ and $S_5 = S^5$), and the level 3 (9 labels) to represent the similarity degrees ($S_2 = S^9$), predicted relevance degrees ($S_3 = S^9$) and satisfaction degrees ($S_4 = S^9$). As the importance degrees are provided initially by UDL staff, we use a set of 5 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 the following subsections we explain all components in detail.

3.1. Resources representation

The considered resources are journal articles, conference contributions, 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 in the resource scope. We use the *vector model* [37] to represent the resource scope and a classification composed by 25 disciplines (see Fig. 4), i.e., a research resource i is represented as

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

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 UDL staff when they add new resources.

Example. We suppose that the UDL staff receive information about a paper about Computer Science. Then, one of the experts accesses to the application and inserts the new research resource. He/she fills all the fields of the form and he/she uses the interface shown in Fig. 4 to select the disciplines of the resource scope. We assume that the expert selects the discipline titled "Science and technology of computers" with an importance degree "Total" and the discipline titled "Telecommunications, electrical engineering, electronics and automatics" with an importance degree "High". These

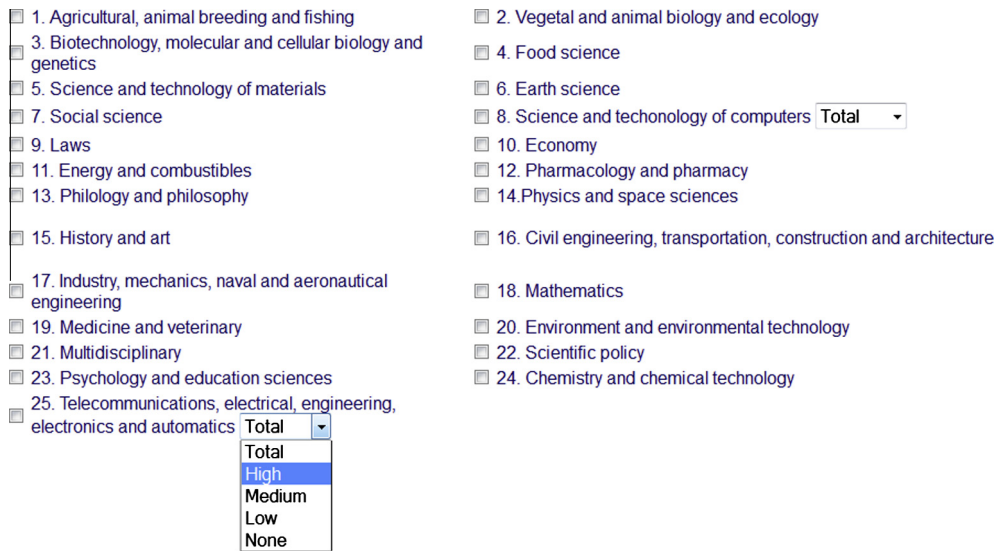


Fig. 4. Interface to define the disciplines of the resource scope.

disciplines are in the positions 8 and 25 respectively of the used classification. The rest of disciplines have an importance degree with a value of “None”. So, the scope of the new resource is represented in the following way:

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

where $VR_{i8} = (b_4, 0)$, $VR_{i25} = (b_3, 0)$ and the rest of positions have the value $(b_0, 0)$.

3.2. User profiles representation

Usually, users have to perform a great effort to provide their preferences, which the system uses to obtain the user profile. To reduce this effort and make easier the process for acquiring the user preferences, we use the method proposed in [49] where we presented an alternative method to obtain the user preferences.

The system asks users to provide their preferences on some research resources, usually a limited number of resources; in our case we use five research resources. The choice of these research resources is made by the UDL staff taking into account the relevance supplied by the users and the research resource scope. We propose users to represent their preferences by means of incomplete fuzzy linguistic preference relations. Then, the system presents 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 results presented in [3], it is enough that the users provide only a row of the preference relation. Then, we use the method proposed in [3] to complete the 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:

- 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 assuring that these preferences are as consistent as possible. The system shows users the five most representative resources, $R = \{r_1, \dots, r_5\}$, and asks them to express their preferences by means of an incomplete fuzzy linguistic preference relation (see Fig. 5). The users only fill those preferences that they wish,

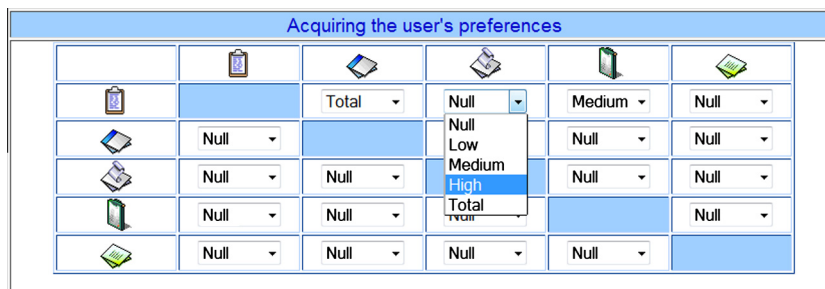


Fig. 5. Interface to define the user preferences about the five most relevant resources.

assigning labels of S_5 . 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 opinion. As aforementioned, the simplest case would be to provide a relation with only one row of preference values:

$$P = \begin{pmatrix} - & p_{12} & p_{13} & p_{14} & p_{15} \\ x & - & x & x & x \\ x & x & - & x & x \\ x & x & x & - & x \\ x & x & x & x & - \end{pmatrix} \quad (9)$$

Then, the system completes the preference relation P using the method proposed in [3], and obtains the relation P^* :

$$P^* = \begin{pmatrix} - & p_{12} & p_{13} & p_{14} & p_{15} \\ p_{21}^* & - & p_{23}^* & p_{24}^* & p_{25}^* \\ p_{31}^* & p_{32}^* & - & p_{34}^* & p_{35}^* \\ p_{41}^* & p_{42}^* & p_{43}^* & - & p_{45}^* \\ p_{51}^* & p_{52}^* & p_{53}^* & p_{54}^* & - \end{pmatrix} \quad (10)$$

where $p_{ij} \in S_5$ are the degrees inserted by the user about 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., user preference vector**, firstly we calculate the user preference degrees on each considered resource according to the preference relation P^* , and secondly, we use these 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 \bar{x}^e (Definition 3). Then, the preference degree of the resource i for the expert called DG_i , is computed as follows:

$$DG_i = \bar{x}^e[p_{i1}^*, \dots, p_{i5}^*] \quad (11)$$

Then, to obtain the user preference vector x , i.e. $VU_x = (VU_{x1}, VU_{x2}, \dots, VU_{x25})$, from the aggregation of the vectors that represents the characteristics of the chosen research resources, i.e., $\{VR_1, \dots, VR_5\}$, weighted by mean of the user preference degrees $\{DG_1, \dots, DG_5\}$. To do that, we use the linguistic weighted average operator defined in Definition 5, and then each position $k = \{1, \dots, 25\}$ of the vector VU_x , is computed as follows:

$$VU_{xk} = \bar{x}_1^w[(VR_{1k}, DG_1), \dots, (VR_{5k}, DG_5)] \quad (12)$$

3.3. Hybrid recommendation approach

In this phase the system filters the incoming information to deliver it to the fitting users. As aforementioned, we implement a *hybrid recommendation* strategy, which switches between a content-based recommendation approach and a collaborative one. Burke [6] proposes a classification composed by different strategies according to the method of combining any two (or more) pure techniques to build a hybrid recommender system. In this sense, our proposal is based on a *switching hybrid approach*, which uses one technique or another, depending on some criterion. A system may try one technique and if the confidence of the results is not satisfactory, it may switch to another technique. In our system, depending on the case, a content-based recommendation approach or a collaborative one is applied. The former is applied when a new item is inserted and the latter is applied when a new researcher is registered.

This process is based on a matching process developed by similarity measures, such as Euclidean Distance or Cosine Measure [37]. In particular, we use the standard cosine measure but defined in a linguistic framework:

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

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_3 , 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.

In the following, we explain both recommendation strategies.

3.3.1. 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:

1. Compute the linguistic similarity degree between VR_i and $VU_e : \sigma_l(VR_i, VU_e) \in S_2$.
2. 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.
3. If i is considered a related resource for e , then the system recommends this resource i to e with a predicted relevance degree $i(e) \in S_3 \times [-0.5, 0.5]$ which is obtained as follows:
 - (a) Look for all research resources stored in the system that were previously assessed by e , i.e., the set of resources $K = \{1, \dots, k\}$ such that there exists the linguistic satisfaction assessment $e(j) \in S_4$, $j \in K$ and $\sigma_l(VU_e, VR_j) \geq (s_4^9, 0)$.
 - (b) Then,

$$i(e) = \bar{x}_l^w((TF_{S_3}^{S_4}(e(1), 0), TF_{S_3}^{S_2}(\sigma_l(VR_i, VR_1))), \dots, (TF_{S_3}^{S_4}(e(k), 0), TF_{S_3}^{S_2}(\sigma_l(VR_i, VR_k)))))$$

where \bar{x}_l^w is the linguistic weighted average operator (Definition 5) and $TF_{t'}^t$ is the transformation function between a 2-tuple that belongs to level t and another 2-tuple in level $t' \neq t$ (Definition 6).

3.3.2. Collaborative recommendations

When new users are inserted into the system, they receive recommendations about resources already inserted, 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 [25]. 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:

1. 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 , $y = 1 \dots n$ 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.
2. 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 $K = \{1, \dots, k\}$ such that there exists a linguistic satisfaction assessment $y(j) \in S_4$, $y \in \aleph_e$, $j \in K$, and $y(j) \geq (s_6^9, 0)$.
3. All the resources $j \in K$, are recommended to e with a predicted relevance degree $j(e) \in S_3 \times [-0.5, 0.5]$ which is calculated as follows:
 - (a) To look for all linguistic satisfaction assessments about resources that were well assessed by the nearest neighbors of e . That is, we recovery $y(j)$ with $j \in K$ and $y \in \aleph_e$.
 - (b) Then,

$$j(e) = \bar{x}_l^w((TF_{S_3}^{S_4}(y_1(j), 0), TF_{S_3}^{S_2}(\sigma_l(VU_e, VU_{y_1}))), \dots, (TF_{S_3}^{S_4}(y_n(j), 0), TF_{S_3}^{S_2}(\sigma_l(VU_e, VU_{y_n}))))$$

where $y_1, \dots, y_n \in \aleph_e$, \bar{x}_l^w is the linguistic weighted average operator (see Definition 5) and $TF_{t'}^t$ is the transformation function between a 2-tuple that belongs to level t and another 2-tuple in level $t' \neq t$ (Definition 6).

3.4. Computing the quality of research resources

In the literature we can find some approaches which allow us to evaluate the information quality in different scopes [8,12,32]. But, very few studies actually apply all the digital evaluation criteria to assess not the digital library, but the resources, that is, the documents. However, our goal is to assess that resources quality according to the user satisfaction but avoiding an intensive user's feedback. To do that, we could use several measures or concepts, but we are focused on a generalist academic environment, so we are dealing with nonspecific research resources. Then we evaluate the quality of research resources as the popularity of the resources, that is, from the users' perceptions on the research resources recommended by the system. If the resources of our system were more specific, such as scientific articles, we could use the author or journal reputation, such as the H-index of the author [2] or the impact factor of the journal. So, we propose to estimate the quality of a research resource based on its popularity. The main reason for adopting this approach is that we already have available the information about the preferred research resources such that we do not need more user interaction neither additional information about the resources to fix the quality of the resources.

As aforementioned, the system asks users to provide their preferences on five research resources, by means of an incomplete fuzzy preference relation. Then, the system uses the method proposed in [3] to complete this preference relation. We use this method to obtain the user profiles, but we also use this information about the preferred research resources to estimate the quality of these resources. Thus, we avoid further user's feedback that could be harmful regarding the usability of the system. We assume that research resources usually preferred over others have a higher quality. Then, the system asks users to provide their preferences on five research resources selected by the UDL staff taking into account the relevance supplied by the users and research resource scope. Then, the users provide their preferences about these resources by means of an incomplete fuzzy preference relation. Once the system completes the incomplete preference relation P and it obtains the relation P^* (see Section 3.2), we can count the times that each resource has been selected to be shown as well as the times

that each resource has been preferred over other. The displayed resources will vary over time, so the system must record each time a resource is selected and each time a resource is preferred to other. So, we estimate the quality of a item i as the probability that the item i be preferred over other having been selected, that is:

$$q(i) = p_i/s_i \quad (13)$$

where p_i is the total of times the resource i has been preferred to another one and s_i is the total of times the resource i has been selected.

3.5. Re-ranking

Once a research resource i is considered relevant for a user e , and both the estimated relevance degree of this resource i for e , $i(e) \in S_3$, and the resource quality score, $q(i) \in [0, 1]$, have been computed, the last step is to aggregate both in a single score. To do this, first we need to translate the research resource quality score to the values range in which the estimated relevance degree is defined, i.e., S_3 . We obtain this translated quality score, $tq(i)$, as follows:

$$tq(i) = q(i) \times g \quad (14)$$

where g is the granularity of S_3 , assuming $S_3 = S^g$ and $g = 8$.

Then, we use a multiplicative aggregation in which the estimated relevance is multiplied by the translated quality score, as follows:

$$\text{Final Relevance } (i) = \Delta \left(\frac{\Delta^{-1}(i(e) \times tq(i))}{g} \right) \quad (15)$$

where Δ and Δ^{-1} are the transformation functions between 2-tuples values and symbolic values defined in section 2.2.1. To obtain the final relevance degree in S_3 , we translate the final relevance value to the interval $[0, g]$.

3.6. Feedback phase

In this phase the recommender system recalculates and updates the ratings of the recommended 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 viceversa. The idea is to improve the generated recommendations taking into account the users' ratings, so it is very important the user provides this information which does not require too much interaction. This feedback activity is developed in the following steps:

1. The system recommends the user U a resource R , and then the system asks him/her his/her opinion or evaluation judgements about this recommended resource.
2. The user communicates his/her linguistic evaluation judgements to the system, $rc_y \in S_4$.
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 \bar{x}^e given in Definition 3.

4. Experiments and evaluation

In this section we present the evaluation of the proposed recommender system. To evaluate the performance of a recommender system, off-line and online experiments could be applied [56]. With an off-line setting, we could compare a method with other approaches without user interaction, using a standard data set. However, in this case we cannot use any standard data set, because we propose a new way to evaluate the quality of research resources and any standard data set has not this information about the preferred research resources. Consequently, in this study we only perform online experiments, i.e., practical studies where a group of users interact with the system and report us their experiences. When the users receive a recommendation, they provide a feedback to the system rating 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.

In this sense we perform two kind of experiments. We begin our study by considering whether the system properly recommends research resources interesting for the users; to do this, we compare the recommendations generated by the system with the recommendations generated by a group of experts (library staff). Afterward we focus on whether the system properly predicts the ratings of the research resources, by taking into account the ratings provided by the user in the previous experiment.

4.1. Data set

For the online evaluation, we have considered a data set with 200 research resources related with different areas, which were included into the system following the indications described in Section 3.3.1. We limited these experiments to 30 users who completed the registration process and they inserted their preferences about the five most relevant resources presented by the system (like in Fig. 5). 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 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 library staff. On the other hand, when the system recommended these 100 resource, it requested the user rates such resources. We registered the ratings provided for the users about the recommended resources to compare with the predictions generated by our system.

4.2. Assessing the capacity of recommendations

In the scope of recommender systems, precision, recall and F1 are widely used measures to evaluate whether a recommender system properly recommends items that the user will consider relevant [56]. To calculate these metrics we need to build a contingency table to categorize the items with respect to the information needs, i.e., the items are classified both as relevant or irrelevant and recommended or not recommended (see Table 2). These metrics are calculated as follows.

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

$$P = \frac{N_{rr}}{N_{rr} + N_{ir}} \tag{16}$$

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

$$R = \frac{N_{rr}}{N_{rr} + N_{rn}} \tag{17}$$

F1 is a combination metric that gives equal weight to both precision and recall, and it is calculated as follows [10,54]:

$$F1 = \frac{2 \times R \times P}{R + P} \tag{18}$$

4.2.1. Results of online experiments

In this section we show the results obtained with the online experiments. We used the online experiment data set outlined in previous subsection. By comparing the recommendations generated by the system with the recommendations provided by the library staff, we obtained the experimental confusion matrix shown in Table 3 and the contingency table shown in Table 4, in which the items are classified both as relevant or irrelevant and recommended to the user or not recommended. From the confusion matrix we can see that the system made 2807 correct predictions and 193 incorrect predictions, so that the error rate is 6.43% and the overall accuracy rate is 93.57%. For more detail, in Table 4 we can see the results obtained for each user. For example, for the user 1, the system selected 7 resources as relevant. However, from the information provided by the library staff, we could see that the system did not select 3 resources that library staff considered relevant for the user 1, and it selected 2 irrelevant resource for user 1. From this contingency table (Table 4), we obtain the corresponding precision, recall and F1 which are shown in Table 5. With this values the average of precision, recall and F1 metrics are

Table 2
Contingency table.

	Recommended	Not recommended
Relevant	Nrr	Nrn
Irrelevant	Nir	Nin

Table 3
Confusion matrix.

	Recommended	Not recommended
Relevant	200	102
Irrelevant	91	2607

69.13%, 66.63% and 0.6765, respectively, which reveal a good performance of the proposed system, and therefore, a good user satisfaction.

4.2.2. A comparative study

In order to obtain data to compare the new approach with some previous approaches, we performed a new study by analyzing the results obtained without considering the resources quality. We use the same data set and the same experimental design outlined in previous subsection but applied over the system proposed in [49]. In this case, we obtain a F1 measure of 0.5939, i.e., a smaller value than the obtained with our new approach. Therefore, by considering the resources quality the system performance is better than whether we consider only the relevance. Comparative results of both approaches are graphically displayed in Fig. 6.

4.3. Assessing the capacity of ratings predictions

To complete our experimental study, we wish to measure the accuracy of our system to predict the ratings a user would give to a resource. We use the **Mean Absolute Error (MAE)** [26,56], a commonly used accuracy metric which considers the average absolute deviation between a predicted rating and the user's true rating, that is:

$$MAE = \frac{\sum_{i=1}^n abs(p_i - r_i)}{n} \quad (19)$$

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.

Table 4
Experimental contingency table.

User	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Nrr	7	9	6	7	6	8	9	5	8	7	5	6	9	6	5
Nrn	3	5	3	4	3	4	4	3	5	3	2	3	5	3	3
Nir	2	3	2	5	2	3	4	2	4	4	3	2	4	4	4
User	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Nrr	8	9	6	4	7	6	7	7	6	5	5	6	8	6	7
Nrn	6	4	4	2	4	2	4	4	2	3	2	2	4	3	3
Nir	3	3	3	1	3	3	2	3	3	2	2	3	5	3	4

Table 5
Precision, recall and F1 values for all the users.

User	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Precision	0.78	0.75	0.75	0.58	0.75	0.73	0.69	0.71	0.67	0.64	0.63	0.75	0.69	0.60	0.56
Recall	0.70	0.64	0.67	0.64	0.67	0.67	0.69	0.63	0.62	0.70	0.71	0.67	0.64	0.67	0.63
F1	0.74	0.69	0.71	0.61	0.71	0.70	0.69	0.67	0.64	0.67	0.67	0.71	0.67	0.63	0.59
User	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Precision	0.73	0.75	0.67	0.80	0.70	0.67	0.78	0.70	0.67	0.71	0.71	0.67	0.62	0.67	0.64
Recall	0.57	0.69	0.60	0.67	0.64	0.75	0.64	0.64	0.75	0.63	0.71	0.75	0.67	0.67	0.70
F1	0.64	0.72	0.63	0.73	0.67	0.71	0.70	0.67	0.71	0.67	0.71	0.71	0.64	0.67	0.67

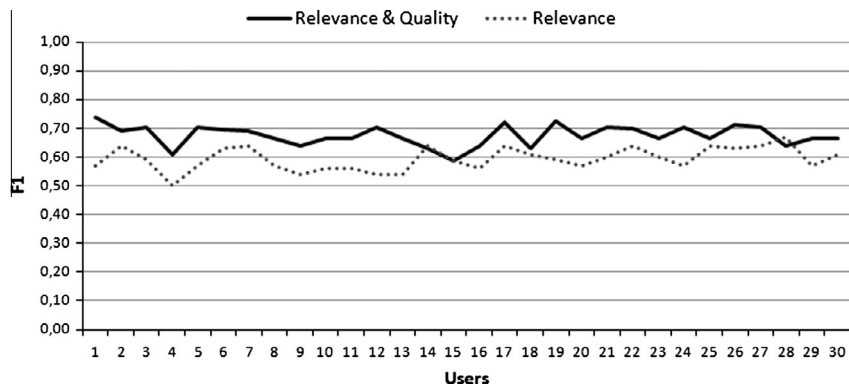


Fig. 6. F1 values for all the users.

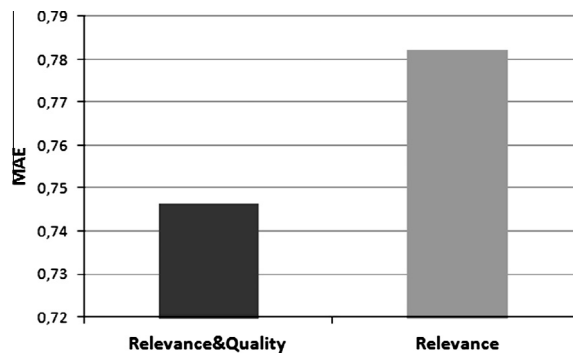


Fig. 7. Comparison of MAE.

As we have said, during the online study we registered the ratings provided for the users about the recommended resources to compare with the predictions generated by our system, and to calculate the MAE. Taking into account the average MAE for all the users, we obtain a final MAE of 0.7565.

We also have studied the approach proposed in [49], that is, without considering the resources quality. In order to compare the results predicting the ratings, we also calculate the MAE for this approach. In this case we obtain a MAE of 0.7823.

As we can see in the Fig. 7, the predictions obtained by using the quality of resources are better than the predictions obtained only with the relevance. Specifically we achieved an improvement of 4.80%. That is, the predictions generated with the new system are more close to the users' preferences.

5. Concluding remarks

In a UDL the selective dissemination of information about research resources is a service particularly important. The UDL 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. Recommender systems have been successfully applied in academic environments to help users in their access processes to relevant information. But these proposals do not take into account the quality of the resources. Now we face the recommendations generation process about research resources as a task with two distinct elements: On the one hand, finding research resources that are relevant to the users and on the other hand, finding valid research resources from the standpoint of the quality of items.

So, we have presented a hybrid fuzzy linguistic recommender system based on quality of the items and we apply it in a UDL to help the users to access relevant research resources. The system measures the item quality and it takes into account this measure like a new factor to be considered in the recommendation process. Thus, the new system incorporates a new module which performs a re-ranking process which takes into account the estimated relevance of an item along with the item quality. Besides, the system improves the feedback process using satisfaction degrees. We have performed online studies with the proposed system and the experimental results show us significant improvements over previous proposals.

Analyzing our system, we could conclude that its main limitation is the need for interaction with UDL staff to establish the internal representations for the research resources. With regard to future research, we believe that a promising direction is to study automatic techniques to establish the representation resources. 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. In this sense, we will study to include new measures aiming at evaluating other aspects than accuracy of recommendation approaches, such as novelty, coverage, trust, serendipity, diversity, utility or other aspects.

Acknowledgments

This paper has been developed with the financing of Projects TIN2010-17876, TIN2010-22145-C02-01, TIC5299 and TIC-5991.

References

- [1] S. Alonso, F. Cabrerizo, F. Chiclana, F. Herrera, E. Herrera-Viedma, Group decision-making with incomplete fuzzy linguistic preference relations, *International Journal of Intelligent Systems* 24 (2009) 201–222.
- [2] S. Alonso, F. Cabrerizo, E. Herrera-Viedma, F. Herrera, H-index: a review focused in its variants, computation and standardization for different scientific fields, *Journal of Informetrics* 3 (2009) 273–289.
- [3] S. Alonso, F. Chiclana, F. Herrera, E. Herrera-Viedma, J. Alcalá-Fdez, C. Porcel, A consistency-based procedure to estimating missing pairwise preference values, *International Journal of Intelligent Systems* 23 (2008) 155–175.
- [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 (2010) 4290–4311.

- [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 (2002) 331–370.
- [7] R. Burke, Hybrid web recommender systems, in: P. Brusilovsky, A. Kobsa, W. Nejdl (Eds.), *The Adaptive Web*, LNCS 4321, 2007, pp. 377–408.
- [8] F. Cabrerizo, J. López-Gijón, A. Ruiz-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) (2010) 455–472.
- [9] J. Callan, A. Smeaton, Personalization and recommender systems in digital libraries, in: Joint NSF-EU DELOS Working Group Report, 2003.
- [10] Y. Cao, Y. Li, An intelligent fuzzy-based recommendation system for consumer electronic products, *Expert Systems with Applications* 33 (2007) 230–240.
- [11] D.S. Carter, J. Janes, Unobtrusive data analysis of digital reference questions and service at the internet public library: an exploratory study, *Library Trends* 49 (2000) 251–265.
- [12] H. Chao, Assessing the quality of academic libraries on the web: the development and testing of criteria, *Library & Information Science Research* 24 (2002) 169–194.
- [13] S. Charlotte-Ahrens, *Recommender Systems: Relevance in the Consumer Purchasing Process*, 2011 (epubli).
- [14] P.-Y. Chen, S.-Y. Wu, Does collaborative filtering technology impact sales? empirical evidence from amazon.com, 2007. <<http://ssrn.com/abstract=1002698>>.
- [15] G.G. Chowdhury, S. Chowdhury, *Introduction to Digital Libraries*, Facet Publisher, London, 2003.
- [16] M. Claypool, A. Gokhale, T. Miranda, Combining content-based and collaborative filters in an online newspaper, in: Proceedings of the ACM SIGIR-99 Workshop on Recommender Systems-Implementation and Evaluation, 1999.
- [17] M. Deshpande, G. Karypis, Item-based top-n recommendation algorithms, *ACM Transactions on Information Systems* 22 (2004) 143–177.
- [18] L. Duen-Ren, L. Chin-Hui, L. Wang-Jung, A hybrid of sequential rules and collaborative filtering for product recommendation, *Information Sciences* 179 (2009) 3505–3519.
- [19] A. Edmunds, A. Morris, The problem of information overload in business organizations: a review of the literature, *International Journal of Information Management* 20 (2000) 17–28.
- [20] M.A. Gonçalves, E. Fox, L. Watson, N. Kipp, Streams, structures, spaces, scenarios, societies (5s): a formal model for digital libraries, *ACM Transactions on Information Systems* 22 (2) (2004) 270–312.
- [21] 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.
- [22] U. Hanani, B. Shapira, P. Shoval, Information filtering: overview of issues, research and systems, *User Modeling and User-Adapted Interaction* 11 (2001) 203–259.
- [23] R. Heradio, F. Cabrerizo, D. Fernández-Amorós, M. Herrera, E. Herrera-Viedma, A fuzzy linguistic model to evaluate the quality of library 2.0 functionalities, *International Journal of Information Management* 33 (4) (2013) 642–654.
- [24] R. Heradio, D. Fernández-Amorós, F. Cabrerizo, E. Herrera-Viedma, A review of quality evaluation of digital libraries based on users' perceptions, *Journal of Information Science* 38 (3) (2012) 269–283.
- [25] J. Herlocker, J. Konstan, J. Riedl, An empirical analysis of design choices in neighborhood-based collaborative filtering algorithms, *Information Retrieval* 5 (2002) 287–310.
- [26] J. Herlocker, J. Konstan, L. Terveen, J. Riedl, Evaluating collaborative filtering recommender systems, *ACM Transactions on Information Systems* 22 (1) (2004) 5–53.
- [27] F. Herrera, E. Herrera-Viedma, Aggregation operators for linguistic weighted information, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems* 27 (1997) 646–656.
- [28] 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 (2000) 43–58.
- [29] F. Herrera, L. Martínez, A 2-tuple fuzzy linguistic representation model for computing with words, *IEEE Transactions on Fuzzy Systems* 8 (6) (2000) 746–752.
- [30] 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) (2001) 227–234.
- [31] E. Herrera-Viedma, J. López-Gijón, Libraries' social role in the information age, *Science* 339 (6126) (2013) 1382.
- [32] 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) (2006) 538–549.
- [33] E. Herrera-Viedma, E. Peis, Evaluating the informative quality of documents in SGML format from judgements by means of fuzzy linguistic techniques based on computing with words, *Information Processing & Management* 39 (2) (2003) 233–249.
- [34] M. Hsu, A personalized english learning recommender system for esl students, *Expert Systems with Applications* 34 (2008) 377–408.
- [35] J. Jeng, What is usability in the context of the digital library and how can it be measured?, *Information Technology and Libraries* 24 (2) (2005) 47–56.
- [36] A. Jun, A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem, *Information Sciences* 178 (1) (2008) 37–51.
- [37] R. Korfhage, *Information Storage and Retrieval*, Wiley Computer Publishing, New York, 1997.
- [38] S. Lawrence, C. Giles, Searching the web: general and scientific information access, *Communications Magazine* 37 (1) (1999) 116–122.
- [39] 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) (2008) 1032–1048.
- [40] P. Lops, M. Gemmis, G. Semeraro, *Recommender Systems Handbook*, chap. Content-based Recommender Systems: State of the Art and Trends, in: F. Ricci, L. Rokach, B. Shapira, P.B. Kantor (Eds.), Springer, 2011, pp. 73–106.
- [41] G. Marchionini, Evaluation digital libraries: a longitudinal and multifaceted view, *Library Trends* 49 (2) (2000) 304–333.
- [42] G. Marchionini, Research and development in digital libraries, 2000. <http://ils.unc.edu/march/digital_library_R_and_D.html>.
- [43] L. Martínez, L. Pérez, M. Barranco, M. Espinilla, Improving the effectiveness of knowledge based recommender systems using incomplete linguistic preference relations, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 16 (2) (2008) 33–56.
- [44] G. Meghabghab, A. Kandel, *Search Engines, Link Analysis, and User's Web Behavior*, Springer-Verlag, Berlin Heidelberg, 2008.
- [45] G. Miller, The magical number seven or minus two: some limits on our capacity of processing information, *Psychological Review* 63 (1956) 81–97.
- [46] M. Pazzani, A framework for collaborative, content-based and demographic filtering, *AI Review* 13 (56) (1999) 393–408.
- [47] W. Pedrycz, *Granular Computing: Analysis and Design of Intelligent Systems*, CRC Press/Francis Taylor, Boca Raton, 2013.
- [48] A. Popescul, L. Ungar, D. Pennock, S. Lawrence, Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments, in: Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence (UAI), San Francisco, 2001.
- [49] C. Porcel, E. Herrera-Viedma, Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries, *Knowledge-Based Systems* 23 (2010) 32–39.
- [50] C. Porcel, J. Moreno, E. Herrera-Viedma, A multi-disciplinary recommender system to advice research resources in university digital libraries, *Expert Systems with Applications* 36 (10) (2009) 12520–12528.
- [51] P. Reisman, H. Varian, Recommender systems, *Special Issue of Communications of the ACM* 40 (3) (1997) 56–59.
- [52] M. Renda, U. Straccia, A personalized collaborative digital library environment: a model and an application, *Information Processing and Management* 41 (2005) 5–21.
- [53] L. Ross, P. Sennyey, The library is dead, long live the library! The practice of academic librarianship and the digital revolution, *The Journal of Academic Librarianship* 34 (2) (2008) 145–152.

- [54] B. Sarwar, G. Karypis, J. Konstan, J. Riedl, Analysis of recommendation algorithms for e-commerce, in: Proceedings of ACM E-Commerce 2000 Conference, 2000, pp. 158–167.
- [55] R. Schenkel, T. Crecelius, M. Kacimi, T. Neumann, J. Parreira, M. Spaniol, G. Gerhard Weikum, Social wisdom for search and recommendation, *Data Engineering Bulletin* 31 (2) (2008) 40–49.
- [56] G. Shani, A. Gunawardana, *Recommender Systems Handbook*, chap. Evaluating Recommendation Systems, in: F. Ricci, L. Rokach, B. Shapira, P.B. Kantor (Eds.), Springer, 2011, pp. 257–298.
- [57] L. Zadeh, The concept of a linguistic variable and its applications to approximate reasoning, Part I, *Information Sciences* 8 (1975) 199–249; L. Zadeh, The concept of a linguistic variable and its applications to approximate reasoning, Part II, *Information Sciences* 8 (1975) 301–357; L. Zadeh, The concept of a linguistic variable and its applications to approximate reasoning, Part III, *Information Sciences* 9 (1975) 43–80.
- [58] A. Zenebe, A. Nocio, Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems, *Fuzzy Sets and Systems* 160 (1) (2009) 76–94.