A LINGUISTIC RECOMMENDER SYSTEM FOR UNIVERSITY DIGITAL LIBRARIES TO HELP USERS IN THEIR RESEARCH RESOURCES ACCESSES

Porcel, C., University of Jaén, Department of Computer Science, Jaén, Spain, cporcel@ujaen.es
Lizarte, M.J., University of Almería, Department of Evolutional and Education Psychology, Almería, Spain, mjlizarte@ual.es
Herrera-Viedma, E., University of Granada, Department of Computer Science and A.I., Granada, Spain, viedma@decsai.ugr.es

Abstract

The Web is one of the most important information media and it is influencing in the development of other media, as for example, newspapers, journals, books, libraries, etc. Moreover, in recent days people want to communicate and collaborate. So, libraries must develop services for connecting people together in information environments. Then, the library staff needs automatic techniques to facilitate that a great number of users can access to a great number of resources. Recommender systems are tools whose objective is to evaluate and filter the great amount of information available on the Web. We present a model of a fuzzy linguistic recommender system to help University Digital Library users in their research resources accesses. This system recommends researchers specialized and complementary resources in order to discover collaboration possibilities to form multi-disciplinaryy groups. In this way, this system increases social collaboration possibilities in a university framework and contributes to improve the services provided by a University Digital Library.

Keywords: *Recommender Systems, Fuzzy Linguistic Modeling, University Digital Libraries*

1 INTRODUCTION

In the last years the concept of digital library is growing. Digital libraries are information collections that have associated services delivered to user communities using a variety of technologies. 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 (Callan J., et al. 2003; GonÇalves M. A. et al. 2004; Renda M.E. and Straccia U. 2005).

Digital libraries are the logical extensions of physical libraries in the electronic information society. These extensions amplify existing resources and services. As such, digital libraries offer new levels of access to broader audiences of users and new opportunities for the library. In practice, a digital library makes its contents and services remotely accessible through networks such as the Web or limited-access intranets (Marchionini G.).

The library staff takes over handle and enables users the access to the documents which are more interesting for them, taking into account their needs or interest areas. The library staff searches, evaluates, selects, catalogues, classifies, preserves and schedules the digital documents access (GonÇalves M. A. et al. 2004).

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). As digital libraries become commonplace and as their contents and services become more varied, the users expect more sophisticated services from their digital libraries (Callan J., et al. 2003; GonÇalves M. A. et al. 2004; Renda M.E. and Straccia U. 2005).

A service that is particularly important 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 relevant items are sent to the users (Marchionini G.). One interesting extension of this concept is to use the connectivity inherent in digital libraries to support collaborative filtering, where users rate or add value to information objects and these ratings are shared with a large community, so that popular items can be easily located or people can search for objects found useful by others with similar profiles (Hanani U. et al. 2001; Marchionini G.; Reisnick P. and Varian H.R. 1997).

Digital libraries have been applied in a lot of contexts but in this paper we focus on an academic environment. University Digital Libraries (UDL) provide information resources and services to students, faculty and staff in an environment that supports learning, teaching and research (Chao H. 2002).

Recommender systems are becoming popular tools for reducing information overload and the use of this kind of systems allows the recommendation of resources interesting for the users, at the same time that these resources are inserted into the system. In the UDL framework, recommender systems can be used to help users (teachers, students and library staff) to find out and select their information and knowledge sources. In (Herrera-Viedma E., Porcel C. et al. 2007) we proposed a recommender system to advise research resources in a UDL, but we think that the recommendation approach could be improved in this application using a hybrid approach and incorporating the multi-disciplanary recommendations.

For this reason, in this paper we propose a fuzzy linguistic recommender system to achieve major advances in the activities of UDL in order to improve their performance. The system is oriented to researchers and it recommends two types of resources: in the first place, specialized resources of the user research area, and in the second place, complementary resources in order to include resources of related areas that could be interesting to discover collaboration possibilities with other researchers and to form multi-disciplinary groups. As in (Porcel C. et al. 2009) we combine a recommender system, to filter out the information, with a multi-granular Fuzzy Linguistic Modeling (FLM), to represent and handle flexible information by means of linguistic labels (Chang S.L. et al. 2007; Chen Z. and Ben-Arieh D. 2006; Herrera F. and Martínez L. 2001; Herrera-Viedma E. et al. 2003).

The paper is structured as follows. Section 2 revises some preliminaries, i.e., the basics of recommender systems and we analyze the approaches of FLM that we use to the system design, the 2-tuple FLM and the multi-granular FLM. In Section 3 we present a multi-disciplinary fuzzy linguistic recommender system to advise research resources in UDL. Section 4 reports the system evaluation. Finally, some concluding remarks are pointed out.

2 PRELIMINARIES

2.1 Recommender systems

Recommender systems could be defined as systems that produce individualized recommendations as output or have the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options (Burke R. 2002). They are becoming popular tools for reducing information overload and for improving the sales in e-commerce web sites (Burke R. 2007; Cao Y., and Li Y. 2007; Reisnick P. and Varian H.R. 1997).

This research area offers tools for discriminating between relevant and irrelevant information by providing personalized assistance for continuous information accesses, filtering the information and delivering it to people who need it (Reisnick P. and Varian H.R. 1997). Automatic filtering services differ from retrieval services in that in filtering the corpus changes continuously, the users have long time information needs (described by mean of user profiles instead of introducing a query into the system), and their objective is to remove irrelevant data from incoming streams of data items (Hanani U. et al. 2001; Marchionini G.; Reisnick P. and Varian H.R. 1997).

There are several approaches that have been proposed for the implementation of recommendation applications (Burke R. 2007; Hanani U. et al. 2001; Reisnick P. and Varian H.R. 1997). In this paper we propose the use of a hybrid approach to smooth out the disadvantages of each one of them and to exploit their benefits; using a hybrid strategy users are provided with recommendations more accurate than those offered by each strategy individually (Good N. et al. 1999). We focus on content-based and collaborative recommender systems, in which the users' information preferences can be used to define user profiles that are applied as filters to streams of documents. The construction of accurate profiles is a key task and the system's success will depend on a large extent on the ability of the learned profiles to represent the user's preferences (Quiroga L.M. and Mostafa J. 2002).

The recommendation activity is followed by a relevance feedback phase. Relevance feedback is a cyclic process whereby the users feed back into the system decisions on the relevance of retrieved documents and the system uses these evaluations to automatically update the user profiles (Hanani U. et al. 2001; Reisnick P. and Varian H.R. 1997).

2.2 Fuzzy Linguistic Modeling

The use of Fuzzy Sets Theory has given very good results for modeling qualitative information (Zadeh L.A. 1975) and it has proven to be useful in many problems, e.g., in decision making (Herrera F. et al. 1996; Xu Z.S. 2006), quality evaluation (Herrera-Viedma E. et al. 2006; Herrera-Viedma E. and Peis E. 2003), models of information retrieval (Herrera-Viedma E. 2001; Herrera-Viedma E. 2001; Herrera-Viedma E. et al. 2007; Herrera-Viedma E. et al. 2005), etc. It is a tool based on the concept of *linguistic variable* proposed by Zadeh (Zadeh L.A. 1975). Next we analyze the two approaches of FLM that we use in our system.

2.2.1 The 2-Tuple Fuzzy Linguistic Approach

The 2-tuple FLM (Herrera F. and Herrera-Viedma E. 1997; Herrera F. and Martínez L. 2000) is a continuous model of representation of information. To define it we have to establish the 2-tuple representation model and the 2-tuple computational model to represent and aggregate the linguistic information, respectively.

Let $S=\{s_0,...,s_g\}$ be a linguistic term set with odd cardinality, where the semantics of the labels is given by means of triangular membership functions and we consider that all terms are distributed on a scale on which a total order is defined. If a symbolic method aggregating linguistic information obtains a value $\beta \in [0,g]$, and $\beta \notin \{0,...,g\}$, then β is represented by means of 2-tuples (s_i,α_i), where s_i represents the linguistic label, and α_i is a numerical value expressing the value of the translation from the original result β to the closest index label, i, in the linguistic term set ($s_i \in S$). This model defines a set of transformation functions between numeric values and 2-tuples: $\Delta(\beta)= (s_i,\alpha_i)$ and $\Delta^{-1}(s_i,\alpha_i)= \beta \in [0,g]$ (Herrera F. and Martínez L. 2000).

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 (Herrera F. and Martínez L. 2000).

2.2.2 The Multi-Granular Fuzzy Linguistic Modeling

In any fuzzy linguistic approach, an improtant parameter to determine is the *granularity of uncertainty*, that is, the cardinality of the linguistc term set. When different experts have different uncertainty degrees on the phenomenon or when an expert has to assess different concepts, then several linguistic term sets with a different granularity of uncertainty are necessary (Herrera F. and Martínez L. 2001). In such situations, we need tools to manage multi-granular linguistic information. In (Herrera F. and Martínez L. 2001) a multi-granular 2-tuple FLM based on the concept of linguistic hierarchy is proposed. A *Linguistic Hierarchy*, *LH*, is a set of levels *l(t,n(t))*, where each level *t* is a linguistic term set with different granularity *n(t)* (Herrera F. and Martínez L. 2001). The levels are ordered according to their granularity. We can define a level from its predecessor level as:

$$l(t,n(t)) \rightarrow l(t+1,2\cdot n(t)-1).$$

In (Herrera F. and Martínez L. 2001) was defined a family of transformation functions between labels from different levels. To define the computational model, we select a level to make the information uniform and then we can use the operators defined in the 2-tuple FLM.

3 A MULTI-DISCIPLINARY RECOMMENDER SYSTEM TO ADVISE RESEARCH RESOURCES IN UDL

In this section we present a fuzzy linguistic recommender system designed using a hybrid approach and assuming a multi-granular FLM. It is used to advise researchers on the best research resources that could satisfy their information needs in a UDL. Moreover, the system recommends complementary resources which could be used by the researchers to meet other researchers of related areas with the aim to discover collaboration possibilities and so, to form multi-disciplinary groups. In this way, it improves the services that a UDL could provide researchers.

The UDL staff manages and spreads a lot of information resources, such as electronic books, electronic papers, electronic journals, official dailies and so on (Callan J., et al. 2003; Renda M.E. and Straccia U. 2005). Nowadays, this amount of information is growing up and they are in need of automated tools to filter and spread that information to the users in a simple and timely manner. So, the digital libraries should anticipate the users' needs and recommending about resources that could be interesting for them. Digital libraries must move from being passive, with little adaptation to their users, to being more proactive in offering and tailoring information for individuals and communities, and in supporting community efforts to capture, structure and share knowledge (Callan J., et al. 2003; GonÇalves M. A. et al. 2004; Renda M.E. and Straccia U. 2005).

We present a hybrid recommender system combining the content-based and collaborative approaches (Burke R. 2007; Hanani U. et al. 2001; Lekakos G. and Giaglis G.M. 2006). The system filters the incoming information stream and delivers it to the suitable researchers according to their research areas. It recommends users research resources of their own research areas and of complementary areas. We use typical similarity functions based on a threshold values to identify research resources of the own areas (Porcel C. et al. 2009). On the other hand, we use Gaussian similarity functions to identify research resources of the complementary areas (Bordogna G. and Pasi G., 1993; Yager R.R. 2007).

To represent the linguistic information we use different label sets, i.e. the communication among the users and the system is carried out by using multi-granular linguistic information, in order to allow a higher flexibility in the communication processes of the system. Therefore, the system uses 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*. 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. In our system, we distinguish three concepts that can be assessed:

- **Importance degree** (S_1) of a discipline with respect to a resource scope or user preferences.
- **Relevance degree** (S_2) of a resource for a user.
- **Complementary degree** (S_3) between the resource scope and the user topics of interest.

Specifically we follow a *LH* of 3 levels (3, 5 and 9 labels), but we only use two levels. We use the level 2 (5 labels) to assign importance degree ($S_1 = S^5$) and the level 3 (9 labels) to assign relevance degrees ($S_2 = S^9$) and complementary degrees ($S_3 = S^9$). Using this *LH* the linguistic terms in each level are:

- $S_5 = \{b_0 = Null = N; b_1 = Low = L; b_2 = Medium = M; b_3 = High = H; b_4 = Total = T\}$
- $S_9 = \{c_0 = Null = N; c_1 = Very\ Low = VL; c_2 = Low = L; c_3 = More\ Less\ Low = MLL; c_4 = Medium = M; c_5 = More\ Less\ High = MLH; c_6 = High = H; c_7 = Very\ High = VH; c_8 = Total = T\}$

The system has three main components: resources management, user profiles management and recommendation process (see figure 1).

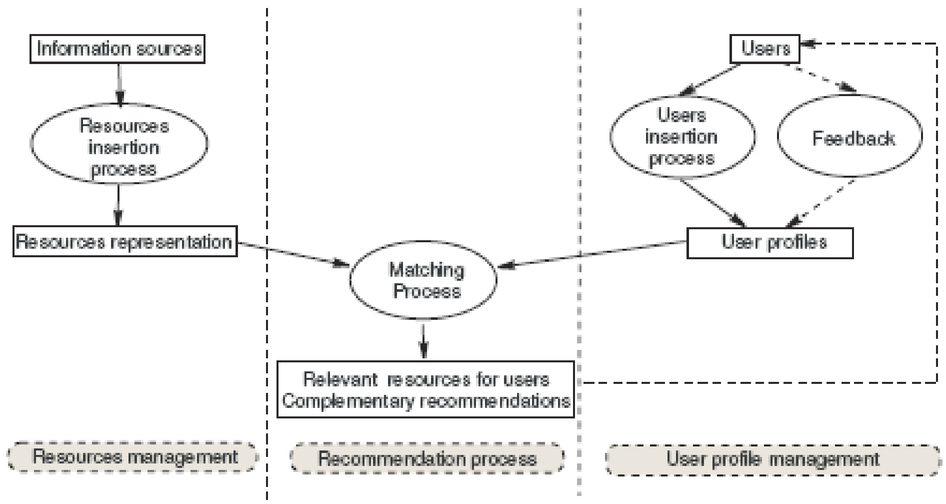

Figure 1. Structure of the system.

3.1 Resources management

This module is the responsible of the management and representation of the research resources; the system obtains an internal representation mainly based in the resource scope. To characterize a resource, the library staff must insert all the available information, as the title, author(s), kind of resource (if it is a book, or book chapter, or a paper, or a journal, or a conference, or an official daily and so on), journal (if it is part of a journal, the system stores the journal name), conference name and dates (if it is a conference), book (if it is a book chapter, the system stores the book title), official daily (if it is part of an official daily, the system stores the daily title), date, source, text, access link to the resource and its scope.

We use the vector model to represent the resource scope (Korfhage R.R. 1997). Thus, to represent a resource *i*, we use a classification composed by 25 disciplines (see table 1). In each position we store a linguistic 2-tuple value representing the importance degree of the resource scope with respect to the discipline represented by that position: $VR_i = (VR_{i1},\ldots, VR_{i25})$. Each component $VR_{ij} \in S_1$, with $j = 1\ldots25$, indicates the importance degree of the discipline *j* with regard to the resource *i*. These importance degrees are assigned by the library staff when they add a new resource.

3.2 User profile management

The users are the researchers registered in the system. The internal representation of their user profiles is based in the researchers topics of interest. To characterize an user, the system stores the following basic information: nickname, password (necessary to access the system), passport number, name and surname, department and center, address, phone number, mobile phone and fax, web, email (elemental information to send the resources and recommendations), research group (it is a string composed by 6 digits, 3 characters indicating the research area and 3 numbers identifying the group), preferences about resources (the users choose the kind of desired resources, i.e. if they want only books, or papers, etc.) and topics of interest.

Agriculture, animal breeding and fishing	Vegetal and animal biology and ecology
Biotechnology, molecular and cellular biology and genetics	Food science and techonology
Materials science and techonology	Earth science
Social science	Computers science and techonology
Law	Economy
Energy and combustibles	Pharmacology and pharmacy
Philology and philosophy	Physics and space sciences
History and art	Civil engineering, transportations, construction and architecture
Industrial, mechanics, naval and aeronautic engineering	Mathematics
Medicine and veterinary	Environment and environmental technology
Multi-disciplinary	Scientific policy
Psychology and education sciences	Chemistry and chemistry technology
Telecommunications, electric engineering, electronics and automatics	

Table 1. Disciplines classification.

We use also the vector model (Korfhage R.R. 1997) to represent the topics of interest. Then, for a user *x*, we have a vector: $VU_x = (VU_{x1},\ldots, VU_{x25})$, where each component $VU_{xy} \in S_1$, with $y = 1\ldots25$, stores a linguistic 2-tuple indicating the importance degree of the discipline *y* with regard to the user *x* topics of interest. These 2-tuples values are also assigned by the library staff.

As the system is based on a content-based approach, it could suffer the cold-start problem to handle new items or new users (Burke R. 2007). New items cannot be recommended to any user until they have been rated by someone. Recommendations for new resources are considerably weaker than more widely rated resources. To overcome this problem, in our system, as it was done in other systems (for example in Movielens (movielens)), when a new user is inserted, the first action to confirm his/her register is to access and assess more than 15 resources of all the resources in the system.

3.3 Recommendation strategy

In this phase the system generates the recommendations to deliver the research resources to the fitting users. We use the following strategies:

- Content-based approach is followed when a new resource is inserted into the system.
- Collaborative approach is followed when a new researcher is inserted into the system. He/she receives information about resources, previously inserted, interesting for him/her.

Both processes are based in a matching process developed between user profiles and resource representations (Hanani U. et al. 2001; Korfhage R.R. 1997). To do that, we can use different kinds of similarity measures, such as Euclidean distance or Cosine Measure. Particularly, we use the standard cosine measure (Korfhage R.R. 1997). However, as the components of the vectors used to represent user profiles and research resources are 2-tuples linguistic values, then we define the cosine measure in a 2-tuple linguistic context:

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

where g is the granularity of the used term set, n is the number of terms used to define the vectors (i.e. the number of disciplines) and (v_{ik}, α_{vik}) is the 2-tuple linguistic value of term k in the vector V_i. With this similarity measure we obtain a linguistic value in S_1. As we represent both the resources and user topics of interest following the vector model, we can use σ_l to calculate the similarity among two resources, two users, or a resource and a user.

3.3.1 Insertion of a new resource

When a new resource is inserted into the system, the linguistic similarity measure $\sigma_l(V_i, V_j)$ is computed among the new resource scope vector (V_i) against all the stored resources in the system (V_j, $j = 1..m$ where m is the number of resources). If $\sigma_l(V_i, V_j) \geq \alpha$ (linguistic threshold value to filter out the information), the resource j is chosen. Next, the system searches for the users satisfied with these chosen resources (previously they have rated the resource as good) and takes into account the user preferences (kind of resources) to consider the user or not. To obtain the relevance of the resource i for a selected user x, the system aggregates the $\sigma_l(V_i, V_j)$ with the assessments previously provided by x about the similar resources and with the assessments provided by others users. To aggregate the information we transform the value $\sigma_l(V_i, V_j)$ in a linguistic label in S_2, using the transformation function defined in (Herrera F. and Martínez L. 2001).

If the calculated relevance degree is greater than a linguistic threshold μ, then, the system sends the resource information and its calculated linguistic relevance degree (label of S_2) to the selected users. If not, the system estimates if the resource could interesting as a complementary recommendation.

To obtain the complementary recommendations, the system calculates the linguistic similarity measure $\sigma_l(V_i, V_x)$ among the resource i and the user x (for all users). Then, it applies a multi-disciplinary function to the value $\sigma_l(V_i, V_x)$. This function must give greatest weights to similarity middle values (near *0,5*), because values of total similarity contribute with efficient recommendations but are probably known for the users. Similarly, null values of similarity show a null relationship between areas. To establish this function we can use the centered OWA operators in which the OWA weights are generated from a Gaussian type function (Yager R.R. 2007). In the proposed system we use a triangular function (figure 2):

$$g(x) = \begin{cases} 2x & \text{for } 0 \leq x \leq 1/2 \\ 2 - 2x & \text{for } 1/2 < x \leq 1 \end{cases}$$

Next, if the obtained multi-disciplinary value is greatest than a previously defined linguistic threshold γ, the system recommends the complementary resource. To express multi-disciplinary values as a linguistic label in S_3, the transformation function is used. Finally, the system sends to the appropriate users the resource information and its estimated linguistic complementary degree (label of S_3).

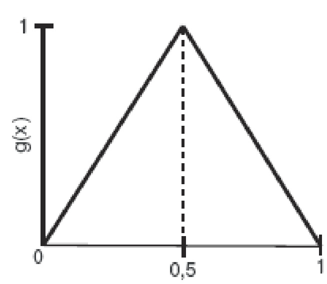

Figure 2. Triangular function.

3.3.2 *Insertion of a new user*

When a new user is inserted into the system, he/she is required to insert some information about the items that satisfied his/her topics of interest, so we use the collaborative approach to generate the recommendations. We follow a memory-based algorithm, which generates the recommendations according to the preferences of nearest neighbors, also known as nearest neighbor algorithms. These algorithms present good performance as related research reported (Symeonidis P. et al. 2008).

The first step is to identify the users most similar to the new user, using the linguistic similarity measure $\sigma_l(V_x,V_y)$ among the topics of interest vectors of the new user (V_x) against all users in the system (V_y, $y = 1..n$ where n is the number of users). If $\sigma_l(V_x,V_y) \geq \delta$ (linguistic threshold value), the user y is chosen as nearest neighbor of x. Next, the system searches for the resources which were interesting for the neighbors of x and takes into account the user preferences (kind of resources) to consider the resource or not. To obtain the relevance of a resource i for the user x, the system aggregates the $\sigma_l(V_x,V_y)$ with the assessments previously provided about i by the nearest neighbors of x. To aggregate the information, we transform the value $\sigma_l(V_x,V_y)$ in a linguistic label in S_2, using the transformation function.

Finally, if the calculated relevance degree is greater than the linguistic threshold μ, then, the system recommends to the new user the resource information and its calculated linguistic relevance degree (label of S_2). If not, the system proceeds to estimate if the resource could be interesting as a complementary recommendation for the new user.

Next, the system calculates the linguistic similarity measure $\sigma_l(V_x,V_y)$ among the user x and the resource i (for all resources). Then, it applies the multi-disciplinary function *g(x)* previously shown in figure 2, to the value $\sigma_l(V_x,V_y)$. If the obtained multi-disciplinary value is greatest than the linguistic threshold γ, the system recommends the resource as complementary. To express multi-disciplinary value as a linguistic label in S_3, the transformation function is used.

Finally, the system sends to the new users the information of all identified resources and its estimated linguistic complementary degree (label of S_3).

3.4 Feedback phase

In this phase the recommender system recalculates and updates the recommendations of the accessed resources. This feedback activity is developed in three steps:

- The system recommends the user *U* a resource *R*, and then it asks the user his/her opinion or evaluation judgements about it.
- The user gives his/her linguistic evaluation judgements, $rc_y \in S2$.
- This evaluation is registered in the system for future recommendations. The system recalculates the linguistic recommendations of *R* by aggregating the opinions provided by other users together with rc_y provided by *U*.

4 SYSTEM EVALUATION

To prove the system functionality, we have implemented a trial version, in which the system works only with few researchers. The main focus in evaluating this trial version is to determinate if it fulfills the proposed innovations, that is, the recommended information (specialized and complementary resources) is useful and interesting for the users. In a later version we will include the system in a UDL.

We calculate the precision, recall and F1, which are measures widely used to evaluate the quality of the recommendations (Cao Y., and Li Y. 2007; Cleverdon C.W. and Keen E.M. 1966; Sarwar B. et al. 2000):

- *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 be relevant.

- *Recall* is calculated as the ratio of the selected relevant items to the relevant items, that is, it represents the probability of a relevant items be selected.
- *F1* is a combination metric that gives equal weight to both precision and recall.

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

We have designed experiments in which the system is used to recommend research resources that best satisfy the preferences of 10 users. We considered a data set with 30 research resources of different areas, collected by the library staff from different information sources. These resources were included into the system following the indications above described. The users completed the registration process and evaluated 15 resources. The resources and the provided evaluations constituted our training data set. After this, we took into account other 20 resources that constituted the test data set. The system filtered these 20 resources and recommends them to the suitable users. Then, we compared the recommendations provided by the systems with the recommendations provided by the library staff. The average of precision, recall and F1 metrics are 63.52%, 67.94% and 65.05% respectively.

We can compare our system with the system proposed in (Cao Y., and Li Y. 2007) because they also calculate the precision, recall and F1, but the obtained values in that evaluation are very high. The values obtained in our system are worse than the obtained in (Cao Y., and Li Y. 2007) but they reveal a good performance of this trial version and therefore a great satisfaction of the users. In future works, we will implement the final version in a UDL and then we will evaluate the system using some datasets like bibsonomy (bibsonomy) or movilens (movielens), and comparing with other recommender systems.

5 CONCLUSIONS

Internet access has resulted in digital libraries that are increasingly used by diverse communities for diverse purposes, and in which sharing and collaboration have become important social elements. Users of UDL need tools to assist them in their processes of information gathering because of the large amount of information available on these systems. We have presented a multi-disciplinary fuzzy linguistic recommender system to spread research resources in UDL.

The proposed system is oriented to researchers and it advises them about resources that could be interesting for them. The system filters the incoming information stream to spread the information to the fitting users, and when new users are inserted into the system, they receive interesting information for them. To improve the services that a UDL provides, it additionally recommends complementary resources that allow researchers to discover collaboration possibilities with other colleagues and to form multi-disciplinary work groups. The multi-granular fuzzy linguistic modeling has been applied in order to improve the users-system interaction and the interpretability of the system activities. The experimental results show us the user satisfaction with the received recommendations.

Acknowledgements

This paper has been developed with the financing of FEDER funds with the PETRI project (PET2007-0460) and FUZZYLING project (TIN2007-61079).

References

Bibsonomy. http://www.bibsonomy.org/.

Bordogna G. and Pasi G. (1993). A fuzzy linguistic approach generalizing boolean information retrieval: A model and its evaluation. *Journal of the American Society for Information Science*, 44, 70-82.

Burke R. (2002). Hybrid recommender systems. *User Modeling and User-Adapted Interaction*, 12(4), 331-370.

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

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

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

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

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

Chen Z. and Ben-Arieh D. (2006). On the fusion of multi-granularity linguistic label sets in group decision making. *Computers and Industrial Engineering*, 51(3), 526-541.

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

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

Good N., Schafer J.B., Konstan J.A., Borchers A., Sarwar B.M., Herlocker J.L. and J. Riedl. (1999). Combining collaborative filtering with personal agents for better recommendations. *In Proc. of the Sixteenth National Conference on Artificial Intelligence*, 439-446.

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

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

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

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

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

Herrera-Viedma E. (2001). 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.

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

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

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

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

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

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

Herrera-Viedma E., Porcel C., López-Herrera A.G. and Alonso S. (2007). A Fuzzy Linguistic Recommender System to Advice Research Resources in University Digital Libraries. *Fuzzy Sets and Their Extensions: Representation, Aggregation and Models. Studies in Fuzziness and Soft computing, Springer,* Vol. 220, 570-589.

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

Lekakos G. and Giaglis G.M. (2006). Improving the prediction accuracy of recommendation algorithms: Approaches anchores on human factors. *Interacting with Computers*, 18, 410-431.

Marchionini G. Research and Development in Digital Libraries. http://ils.unc.edu/~march/digital_library_R_and_D.html. Last access: April, 2009.

Movielens - movie recommendations. http://movielens.umn.edu/login.

Porcel C., López-Herrera A.G. and Herrera-Viedma E. (2009). A recommender system for research resources based on fuzzy linguistic modeling, *Expert Systems with Applications*, 36, 5173-5183, doi: 10.1016/j.eswa. 2008.06.038.

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

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

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

Sarwar B., Karypis G., Konstan J. and Riedl J. (2000). Analysis of recommendation algorithms for e-comerce. *In Proceedings of ACM E-Commerce 2000 conference*, 158-167.

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

Xu Z.S. (2006). An approach based on the uncertain lowg and induced uncertain lowg operators to group decision making with uncertain multiplicative linguistic preference relations. *Decision Support Systems*, 41(6), 488-499.

Yager R.R. (2007). Centered OWA operators. *Soft Computing*, 11, 631-639.

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

ISBN 978-960-98566-7-6

[i] Crowned at last, Economist, April 2nd 2005

[ii] Person of the Year, TIME, December 26/ 2006 -January 1/2007

[iii] Zogby Poll, 27 February 2008, http://www.zogby.com/news/ReadNews.dbm?ID=1454

[iv] Newsweek, 2007 The New Wisdom of the Web, http://www.newsweek.com/id/45976/page/2

[v] Nielsen Online, 2008, Over 875 Million Consumers Have Shopped Online -- The Number of Internet Shoppers Up 40% in Two Years, http://www.nielsen.com/media/2008/pr_080128b.html

[vi] BizReport.com, June 5, 2007

[vii] BizReport.com, June 5, 2007

[viii] Possible roles can be: Informational, promotional, relational, educational or transactional

[ix] http://www.marketingpower.com/Community/ARC/Pages/Additional/Definition/default.aspx

[x] http://www.haveyoursay.com/

[xi] http://www.nespresso-whatelse.com/club/?xtor=