Expert Systems with Applications 36 (2009) 12520–12528

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A multi-disciplinar recommender system to advice research resources in University Digital Libraries

C. Porcel [a,*], J.M. Moreno [b], E. Herrera-Viedma [c]

[a] University of Jaen, Department of Computer Science, Jaen, Spain
[b] University of Murcia, Department of Computer Science, Murcia, Spain
[c] University of Granada, Department of Computer Science and A.I., Granada, Spain

ARTICLE INFO

Keywords:
Recommender systems
Fuzzy linguistic modeling
University Digital Libraries

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, and libraries. In this paper, we analyze the logical extensions of traditional libraries in the Information Society. In Information Society people want to communicate and collaborate. So, libraries must develop services for connecting people together in information environments. Then, the library staff need automatic techniques to facilitate so 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 to assist the users in their information access processes. We present a model of a fuzzy linguistic recommender system to help the *University Digital Libraries* users to access for their research resources. This system recommends researchers specialized and complementary resources in order to discover collaboration possibilities to form multi-disciplinar 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.

1. Introduction

In the last years the new concept of digital library is growing. *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 displayed on the 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 et al., 2003; Gonçalves, Fox, Watson, & Kipp, 2004; Renda & Straccia, 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, 2009).

The digital libraries are composed of human resources (staff) that take over handle and enable the users to access the documents that are more interesting for them, taking into account their needs or areas of interest. The library staff searches, evaluates, selects, catalogues, classifies, preserves and schedules the digital documents access (Gonçalves et al., 2004). Some of the main digital libraries functions are the following:

- To evaluate and select digital materials to add in its repository.
- To preserve the security and conservation of the materials.
- To describe and index the new digital materials (catalogue and classify).
- To deliver users the material stored in the library.
- Other managerial tasks.

Libraries offer different types of references and referral services (e.g., ready reference, exhaustive search, and selective dissemination of information), instructional services (e.g., bibliographic instruction and database searching), added value services (e.g., bibliography preparation, and language translation) and promotional services (e.g., literacy and 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 et al., 2003; Gonçalves et al., 2004; Renda & Straccia, 2005).

A service that is particularly important is the selective dissemination of information or filtering (Morales del Castillo, Pedraza-Jiménez, Ruíz, Peis, & Herrera-Viedma, 2009; Morales del Castillo, Peis, Moreno, & Herrera-Viedma, in press). Users develop profiles

* Corresponding author.
E-mail addresses: cporcel@ujaen.es (C. Porcel), jmmoreno@um.es (J.M. Moreno), viedma@decsai.ugr.es (E. Herrera-Viedma).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.04.038

that reveals their areas of interest and as new materials are added to the collection, they are compared to the profiles and relevant items are sent to the users (Marchionini, 2009).

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, Shapira, & Shoval, 2001; Marchionini, 2009; Reisnick & Varian, 1997).

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

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-disciplinar groups. As in (Porcel, López-Herrera, & Herrera-Viedma, 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, Wang, & Wang, 2007; Chen & Ben-Arieh, 2006; Herrera & Martínez, 2001; Herrera-Viedma, Cordón, Luque, López, & Muñoz, 2003; Herrera-Viedma, Martínez, Mata, & Chiclana, 2005; Herrera, Herrera-Viedma, & Martínez, 2008).

The paper is structured as follows. Section 2 revises some preliminaries, i.e., the concept and main aspects about recommender systems and 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-disciplinar fuzzy linguistic recommender systems to advice research resources in UDL. Section 4 reports the system evaluation and some experimental results. 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 has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options (Burke, 2002). They are becoming popular tools for reducing information overload and for improving the sales in e-commerce web sites (Burke, 2007; Cao & Li, 2007; Hsu, 2008; Reisnick & Varian, 1997).

It is a research area that 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 & Varian, 1997). Automatic filtering services differ from retrieval services in that in filtering the corpus changes continuously, the users have long time information needs (described by mean of user profiles instead of to introduce a query into the system) and their objective is to remove irrelevant data from incoming streams of data items (Hanani et al., 2001; Marchionini, 2009; Reisnick & Varian, 1997). A result from a recommender system is understood as a recommendation, an option worthy of consideration; a result from an information retrieval system is interpreted as a match to the user's query (Burke, 2007).

A variety of techniques have been proposed as the basis for recommender systems. We can distinguish four different classes of recommendation techniques based on the source of knowledge (Burke, 2007; Hanani et al., 2001; Reisnick & Varian, 1997):

- *Content-based systems*: They generate the recommendations taking into account the terms used in the items representation and the ratings that a user has given to them (Basu, Hirsh, & Cohen, 1998; Claypool, Gokhale, & Miranda, 1999). These recommender systems tend to fail when little is known about the user information needs.
- *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 (Good et al., 1999; Renda & Straccia, 2005).
- *Demographic systems*: A demographic recommender system provides recommendations based on a demographic profile of the user. Recommended items can be generated for different demographic niches, by combining the ratings of users in those niches (Pazzani, 1999).
- *Knowledge-based systems*: These systems generate the recommendations based on the inferences about items that satisfy the users from the information provided by each user regarding his/her knowledge about items that can be recommended (Burke, 2002).

All these techniques have benefits and disadvantages. However, we can use a hybrid approach to smooth out the disadvantages of each one of them and to exploit their benefits (Basu et al., 1998; Claypool et al., 1999; Good et al., 1999). In these kind of systems, the users' information preferences can be used to define user profiles that are applied as filters to streams of documents. Therefore, 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 & Mostafa, 2002).

The recommendation activity is followed by a relevance feedback phase. *Relevance feedback* is a cyclic process whereby the user feeds back into the system decisions on the relevance of retrieved documents and the system then uses these evaluations to automatically update the user profile (Hanani et al., 2001; Reisnick & Varian, 1997).

2.2. Fuzzy linguistic modeling

The use of fuzzy sets theory has given very good results for modeling qualitative information (Zadeh, 1975) and it has proven to be useful in many problems, e.g., in decision making (Cabrerizo, Alonso, & Herrera-Viedma, 2009; Herrera, Herrera-Viedma, & Verdegay, 1996; Mata, Martínez, & Herrera-Viedma, 2009), quality evaluation (Herrera-Viedma, Pasi, López-Herrera, & Porcel, 2006; Herrera-Viedma & Peis, 2003), models of information retrieval (Herrera-Viedma, 2001a, 2001b; Herrera-Viedma & López-Herrera, 2007; Herrera-Viedma, López-Herrera, Luque, & Porcel, 2007; Herrera-Viedma, López-Herrera, & Porcel, 2005), and political analysis (Arfi, 2005). It is a tool based on the concept of *linguistic variable* proposed by Zadeh (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 & Martínez, 2000) is a continuous model of representation of information which allows to reduce the loss of information typical of other fuzzy linguistic approaches