

A survey on automatic recommenders in the context of institutional repositories

Eugenio P. Concepción
Universitat Oberta de Catalunya
econcepcion@uoc.edu

ABSTRACT

The present article provides a general review of automatic recommendations tools in digital libraries. In this context, we aim to analyse three elements: how digital libraries are conceptualized (that is, how they are represented as software entities), the algorithms used by automatic recommenders to perform their logic, and the modelling of users' interactions in this kind of environments. In order to define the complete context of the research, we will start by defining what a digital library is and the way it is represented in common software management applications. Then, we will review the basic definitions related to recommenders and the usual algorithms. We focus on hybrid recommendation techniques because they provide more accurate results and help to get over the inherent weak points of basic techniques used individually. Finally we will analyse the users' interactions in digital libraries.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Information Filtering, Retrieval models

Keywords

Automatic Recommender, Institutional Repositories, Digital Libraries, Learning Enhanced Technologies, Hybrid Recommenders, Context-aware Recommenders.

1. CONTEXT

The development of e-learning environments in the last years has brought new usage scenarios for academic libraries. Even in traditional education, digital access to libraries resources has become essential (Saracevic, 2004). By means of digital libraries, students and researchers have new abilities when looking for educational resources, but also new challenges. As Smeaton (2005) states, a simple search function, a usual service of any digital library, increasingly leads to user frustration as user needs become more complex and the volume of managed information increases. So, search engines of digital libraries can quickly generate a number of result pages for a certain topic, and users must tackle the problem of selecting and filtering the most convenient resources from a large collection. Moreover, as Verbert (2011) indicates, finding relevant resources can be even more difficult when requirements are not always fully known by the learner, such as his level of competence or adequate technical format.

In order to help them in this task, digital libraries offer advanced search features, and even certain capabilities of collaborative search have been proposed (Smeaton, 2005). Up to date, most of these collaborative tools are mainly based in social criteria. These systems take advantage of users' social context to generate feasible alternatives for them. But, as we will see, by means of this technique we can only generate a subset of possible results.

In order to establish a formal definition, we will consider digital libraries in the same way as Smeaton (2005). Digital libraries are collections of information that have associated services made accessible to users communities by means of different technologies. Among

the resources included in these collections there are digital versions of books, research papers, articles, and also other multimedia contents like video or audio.

1.1 Automatic recommenders

Automatic recommenders are a well-known personalisation tools that can be applied in several contexts. Recommender systems are an extensively studied and well established field of research and application (Adomavicius, 2005). Major search engines like Google and electronic shops like Amazon have incorporated recommendation technology in their services in order to personalize their results. Even more, as Verbert (2011) reminds, these systems have been researched and deployed extensively over the last decade in various application areas, including e-commerce and e-health. According to Smeaton (2005), personalisation is defined as the ways in which information and services can be tailored to match the unique and specific needs of an individual or a community. This is achieved by adapting presentation, content, and/or services based on a persons task, background, history, device, information needs, location, or any other descriptive feature. Recommender systems are a particular type of personalisation that learn about those persons or communities needs and then proactively identify and recommend information that meets those needs.

According to Burke (2002), we can consider four different basic recommendation techniques:

- Collaborative: in this approach, recommendations are generated using exclusively information about rating profiles provided by users. The quid of this technique is to locate users with similar rating history with target user and generate recommendations according to their neighbourhood.
- Content-based: recommendations are generated according to products features and the ratings provided by other users. With this information, the system learns to classify products from users rating.
- Demographic: recommendations are based on users' demographic profiles. The recommender creates an initial classification based on users' background and then enriches it with its ratings.
- Knowledge-based: recommendations are generated according to system inferences about users' needs and preferences. These systems are usually provided with some extra logic in order to match product features with users' needs.

However, all these techniques, used individually, have shortcomings that need to be fixed by different means. In contrast, hybrid techniques, based on a combined approach are able to overcome the weaknesses of single techniques and enhance the quality of provided results. Burke (2002) lists all known hybridization techniques and provides a complete table showing redundant or useless combinations of them.

So, there are seven possible combinations:

- Weighted: The final score is a linear combination of the values generated by the different recommendation techniques.
- Switching: The system chooses among recommendation components and applies the selected one.
- Mixed: Recommendations from different recommenders are presented together for users to make a choice.
- Feature Combination: Recommender uses features derived from different sources and combines them in a single recommendation algorithm.
- Feature Augmentation: One recommendation technique processes a feature or set of features, which is then part of the input to the next technique.

- Cascade: Recommenders have different priorities, and the lower priority ones are used to make the final decisions.
- Meta-level: The first recommendation technique is used to generate a model, which is then used by the next technique.

As we have seen, automatic recommenders usually rely on a number of single techniques, or hybrid recommendation algorithms. Every technique presents both advantages and drawbacks, and all of them have been profusely documented (Burke, 2003). The common point is that all these algorithms use information about users and / or resources to generate their recommendations.

1.2 Contents representation

Certain recommendation techniques, like content-based or knowledge-based, need to represent content metadata. These data can be modelled in a custom way or following a standard. The present point introduces a brief summary of the main related standards. All the following definitions are taken from Barker (2010).

1.2.1 IEEE LOM

The IEEE LOM (Learning Object Metadata) is an open standard for the description of learning objects, and is composed of two parts: a conceptual data schema (defined by IEEE in 2002), and a XML binding of that schema (also defined by IEEE in 2005). The definition of learning object used in this standard is “*any entity digital or non-digital that may be used for learning, education, or training*”. The LOM data schema specifies which characteristics of a learning object must be described and what vocabularies can be used for these descriptions; it also defines how this data model can be amended by additions or constraints.

The LOM data schema consists of a conceptual hierarchy of elements where the first level is composed of nine categories, each of which contains sub-elements; these sub-elements may be simple elements containing data, or they may be complex elements that aggregate sub-elements. The data model also specifies which elements may be repeated (individually or as a group).

The semantics of LOM elements are determined by their context. The semantic of each element can be affected by the parent element in the hierarchy and sometimes by other elements in the same container. The data schema also specifies the value space and data type for each of the simple data elements. The value space defines the restrictions, if any, on the data that can be entered for that element.

1.2.2 SCORM

The SCORM (*Sharable Content Object Reference Model*) is a collection and harmonization of specifications and standards that defines the interrelationship of content objects, data models and protocols such that objects are sharable across systems that conform to the same model (SCORM, 2013). This specification promotes reusability and interoperability of learning content across Learning Management Systems (LMS). The SCORM has releases dating back to 2000 with SCORM 1.0. SCORM 1.2, released in 2001 is the final version of SCORM before the integration of sequencing. Beginning in 2004, SCORM began to version with different editions of SCORM 2004. The most recent release (2009) is SCORM 2004 4th Edition.

1.2.3 Dublin Core Metadata

The *Dublin Core Metadata Initiative* (abbr. DCMI) develops several metadata standards, but the point of interest for this article is the *Dublin Core Element Set* (DCMI, 2008), which has been standardized as ISO Standard 15836-2003. The core Element Set is described as “*broad and generic, usable for describing a wide range of resources*” (DCMI, 2008). The

range of resource types to which Dublin Core metadata is applicable is emphasized in the formal definition of a resource, used elsewhere by DCMI, as "anything which might be identified" (Powell, 2007). These fifteen elements are:

“contributor, coverage, creator, date, description, format, identifier, language, publisher, relation, rights, source, subject, title and type.”

However, while the coverage of the Element Set is broad, it is not exhaustive. That is, there are many characteristics of resources that are not covered, some of which are important in specialized domains. For example there is no way to describe the intended audience of a resource. For these reasons, the original Dublin Core Element Set has been amended with several extensions, as follows:

- The DCMI Abstract Model (Powell, 2007) *"defines the nature of the components used [in Dublin Core metadata] and describes how those components are combined to create information structures"*.
- DCMI Metadata Terms (DCMI, 2008b) defines all the metadata terms maintained by the DCMI. The terms are divided into properties, vocabulary encoding schemes, syntax encoding schemes and classes. Classes are formal categories of resources that share important characteristics, e.g. "bibliographic resources" (books, journal articles) or "file formats".
- The Singapore Framework for Dublin Core Application Profiles (Nilsson, 2008) defined an approach for creating and documenting application profiles based on the Dublin Core abstract model and metadata vocabularies, such as the DCMI Metadata Terms, that are compatible with the model.
- Guidelines for encoding Dublin Core metadata in RDF, XML and HTML/XHTML meta and link elements.

1.3 Users modelling

In the same way that certain systems represent contents to be recommended, other systems need to create a model for the users (i.e. collaborative or demographic recommenders). There are two main standards that we should consider:

1.3.1 IEEE PAPI

The IEEE PAPI (*Public and Private Information for Learners*) was developed to provide a specification for modeling students syntactically and semantically. It includes structures to represent students' knowledge, learning styles, abilities, and any further personal information.

The specification provides several granularity levels and also a logical classification, and an extensions mechanism to allow personalization. It considers cultural and academic conventions and also provides security management.

1.3.2 IMS LIP

IMS LIP (*Learner Information Package*) specifies an XML-based syntax to allow the interchange of students' information between systems. Almost all of its elements are optional, so it is designed to be flexible and easily extensive. It can be used individually or organizationally. Atomic data are structured in eleven levels (accessibility, activity, affiliation, etc.). It incorporates the IEEE PAPI specification.

1.4 Technology Enhanced Learning

For the purpose of this article, automatic recommenders should be considered in the context of what is called Technology Enhanced Learning (TEL). Manouselis (2011) defines TEL as a field that aims to design, to develop and to test technologies centred on enhancing learning

practices of both individuals and organisations. It is therefore an application domain that studies technologies that support all forms of teaching and learning activities.

Again, as states Manouselis (2011), information retrieval, understood in terms of searching for relevant learning resources, have great importance as a central activity in TEL and therefore recommender systems have attracted increased interest.

1.5 Context-aware recommender systems

Unfortunately, the algorithms underlying regular recommender systems are not directly transferable to the area of TEL (Verbert, 2011). These algorithms use information about users and resources to generate recommendations. Purposely, most TEL recommender systems rely on users' profiles to gather additional information, as opposed to traditional recommenders, that focus on users' likes or interests. The knowledge level of the learner is often used to personalize recommendations, such as his/her knowledge of course concepts or past academic grades. Due to the fact that learning process usually takes place in a notably complex and heterogeneous environments, the use of contextual information relating to the user by recommenders has attracted major interest. Such contextualization is being researched as a paradigm for building intelligent systems that can better predict and anticipate the needs of users, and act more efficiently in response to their behavior (Verbert, 2011).

At this point, it is necessary to establish a definition for the concept of context. One of the most cited definitions was provided by A. Dey (2001), and cited by Verbert (2011). They define context as "*any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves*".

According to Verbert (2011), traditional recommenders (collaborative, content-based, knowledge-based and hybrid recommender systems) deal with two types of entities, users and items. As stated before, TEL applications have certain specificities, and this two-dimensional approach could be insufficient. Verbert (2011) proposes to incorporate additional data about users in order to enhance recommender's efficiency by means of data contained in the context. Such data can be used to adapt recommendations based on individual learner characteristics, such as learning goals and knowledge levels, and additional contextual information such as available time, location, people nearby, etc. This approach has been denominated context-aware recommender systems (CARS).

Adomavicius and Tuzhilin (2011) were pioneers in CARS. The authors developed a research where the user/item paradigm was extended to support additional dimensions capturing the context in which recommendations are made. Such contextual information can be obtained in a number of ways:

- Explicit context, that is, information is captured from users' manual input. For example, registration forms are often used to capture information of users, or rating input can be used to retrieve interests and preferences.
- Implicit methods, which capture contextual information automatically from the environment, for instance by obtaining the current location of the user by means of his browser.
- Contextual information can also be inferred by analyzing user interactions with tools and resources, for instance to estimate the current task of the user.

Verbert (2011) refers the different paradigms that have been proposed to incorporate contextual information in the recommendation process. Basically we must consider two approaches:

- A first recommendation using *context-driven querying and search approach* that takes advantage of contextual information to query or search a certain repository of resources (e.g., restaurants) and presents the best matching resources (e.g., nearby restaurants that are currently open) to the user.
- A second contextual *preference elicitation and estimation approach* that attempts to model and learn contextual user preferences. These recommender systems are built on knowledge of partial contextual user preferences and usually handle data records of the form <user, item, context, rating>. Each record therefore captures how much a user liked a particular item in a specific context.

In the last case, there are three approaches to deal with contextual preferences:

- Contextual pre-filtering: In this approach, contextual information is used to filter the dataset before applying a traditional recommendation algorithm.
- Contextual post-filtering: In which recommendations are generated on the entire dataset. The resulting set of recommendations is adjusted later using the contextual information.
- Contextual modeling approaches use contextual information directly in the recommendation function as an explicit predictor of a rating for an item.

Whereas contextual pre-filtering and post-filtering approaches can use traditional recommendation algorithms, the contextual modeling approach uses multidimensional recommendation algorithms.

Verbert (2011) highlights the current point of interest in CARS research: the influence of various parameters on the recommendation. This challenge has been identified by several authors and it is subject of debate in the community. She also outlines, citing Z. Yujie, the difficulty of describing clearly and uniformly what types of contexts are truly needed in CARS because of the variety of application scenarios and user needs.

1.6 Institutional Repositories

As relevant as the techniques applied by recommenders is the way in which digital contents are modelled. Conway (2008) offers a complete revision of how the digital contents perception has evolved in the last years. It focuses on the concept of Institutional Repository (IR) that introduces the use of metadata as part of the information managed. Minguillón (2010) emphasizes the differences between IR and generic digital collections. In the first case, contents are deposited in the repository together with their metadata, and are also accessible by means of several management operations.

In the present article we will use the term digital library and IR seamlessly. That is, an IR is equivalent, for our purposes, to the definition of digital library that we have stated before. Both can be abstracted as digital content repositories that contain information objects characterized by a specific set of metadata. These metadata are subsequently used to support the two main services of the IR: content preservation and content reusing.

2 SURVEY OF EXISTING RECOMMENDERS

In the TEL domain, a number of recommender systems have been introduced in order to propose learning resources to users (Manouselis, 2011). Such systems could potentially play an important educational role, considering the variety of learning resources that are published online. The following table summarizes some of the most relevant references gather by Manouselis (2011):

System / Authors	Description
<p>Altered Vista (Recker and Walker 2000; Recker and Wiley 2000; Recker and Walker 2003; Recker et al.2003; Walker et al. 2003)</p>	<p>It is a full system, focused on develop a collaborative filtering system for learning resources. The project was quite extensive, and it tried to explore how to collect user-provided evaluations of learning resources, and then to propagate them as verbal recommendations about the qualities of the resources.</p>
<p>RACOFI (Anderson et al. 2003; Lemire et al. 2005)</p>	<p>RACOFI is a prototype that combines two recommendation approaches. Firstly, it uses a collaborative filtering engine that works with ratings that users provide for learning resources, and then integrates it with an inference rule engine containing mining association rules between the learning resources and using them for recommendation.</p>
<p>QSIA (Rafaeli et al. 2004; Rafaeli et al. 2005)</p>	<p>QSIA (Questions Sharing and Interactive Assignments) is a full system developed for learning resources sharing, assessing and recommendation. This system is used in the context of online communities, in order to develop a social perspective in learning and to promote collaboration and online recommendation. It is not a totally automated recommender system because the user can decide whether to assume control on who advises (friends) or to use a collaborative filtering service.</p>
<p>CYCLADES (Avancini and Straccia 2005)</p>	<p>CYCLADES is also a full system focused on collaboration. It has proposed an environment where users search, access, and rate digital resources available in repositories found through the Open Archives Initiative (an opened interoperability initiative for digital archives). So that, the system can offer recommendations over resources that are stored in different archives and accessed through an open scheme.</p>
<p>CoFind (Dron et al. 2000 a,b)</p>	<p>CoFind is a prototype. It also uses digital resources that are freely available on the Web, but proposing a new approach by applying for the first time folksonomies (tags) for recommendations.</p>
<p>Learning object sequencing (Shen and Shen 2004)</p>	<p>This is a prototype proposed as a recommender system for learning objects that is based on sequencing rules that help users be guided through the concepts of an ontology of topics. The rules are fired when gaps in the competencies of the learners are identified, and then appropriate resources are proposed to the learners.</p>
<p>Evolving e-learning system (Tang and McCalla 2003; 2004a; 2004b;2004c; 2005)</p>	<p>It is a full system focused on providing an evolving e-learning system that includes a hybrid recommendation service. In this approach, resources are tagged according to their content and technical aspects, but learners also provide feedback about them in the form of ratings. Recommendation takes place both by using data clustering</p>

	<p>techniques to group learners with similar interests and by using collaborative filtering techniques to identify learners with similar interests in each cluster.</p>
<p>ISIS - Hybrid Personalised Recommender System (Drachslers et al. 2009c)</p>	<p>The ISIS system is a prototype that adopts a hybrid approach for recommending learning resources. The authors build upon a previous simulation study by Koper (2005) in order to propose a system that combines social-based (using data from other learners) with information-based (using metadata from learner profiles and learning activities) in a hybrid recommender system.</p>
<p>Multi-Attribute Recommendation Service (Manouselis et al. 2007)</p>	<p>This prototype uses a neighborhood-based set of collaborative filtering algorithms in order to support learning object recommendation. The main innovation of this study is that the engaged algorithms have been multi-attribute ones, allowing the recommendation service to consider multidimensional ratings that users provide on learning resources.</p>
<p>Learning Object Recommendation Model (Tsai et al. 2006)</p>	<p>Learning Object Recommendation Model (LORM) is a design that uses a hybrid recommendation algorithmic approach and describes resources upon multiple attributes, but has not yet reported to be implemented in an actual system.</p>
<p>Simulation environment (Nadolski et al. 2009)</p>	<p>This system is a simulation environment for combining different recommendation algorithms in a hybrid recommender system. The purpose of the system is to compare them against each other, regarding their impact on learners in informal learning networks. Authors compared various cost intensive ontology based recommendation strategies with light-weight collaborative filtering strategies.</p>
<p>ReMashed (Drachslers et al. 2009a,b)</p>	<p>ReMashed is a complete system developed to address learners in informal learning networks. A mash-up environment is created by combining users from different Web2.0 services like Flickr, Delicious.com or Sildeshare.com. Again, it is applied a hybrid recommender system that takes advantage of the tag and rating data of the combined Web2.0 sources to generate recommendations.</p>
<p>CourseRank (Koutrika et al. 2008;2009)</p>	<p>CourseRank is a full system used as an unofficial course guide for Stanford University students. In this system, the recommendation process approach focuses on querying a relational database with course and student information and the use of tuple operators to generate recommendations.</p>
<p>RPL recommender (Khribi et al. 2009)</p>	<p>In RPL, a hybrid approach has also been adopted. This prototype includes a recommendation engine that combines a collaborative filtering algorithm with a content-based filtering algorithm, using data that has been logged and</p>

	mined from user actions.
--	--------------------------

3 ANALYSIS OF APPLIED TECHNIQUES

The following point focuses on providing a comparative analysis of the previously mentioned recommenders. The purpose of this is to acquire a general view of the developed projects and to detect the main gaps or difficulties to be resolved.

Prior to review some of the recommenders mentioned above, it is necessary to expose the main advantages and disadvantages of used techniques. The following table contains a summary from Drachslar (2008):

Technique	Description	Advantages	Disadvantages
User-based CF	Users who rated the same item similarly probably have the same taste. Based on this assumption, this technique recommends the unseen items already rated by similar users.	No content analysis Domain-independent Quality improves Bottom-up approach Serendipity	New user problem New item problem Popular taste Scalability Sparsity Cold start problem
Item-based CF	Focus on items, assuming that the items rated similarly are probably similar. It recommends items with the highest correlation (based on ratings for the items).	No content analysis Domain-independent Quality improves Bottom-up approach Serendipity	New item problem Popular taste Sparsity Cold start problem
Demographic CF	Users with similar attributes are matched, and then it recommends items that are preferred by similar users (based on user data instead of ratings).	No cold start problem Domain-independent Serendipity	Obtaining information Insufficient information Only popular taste Obtaining metadata information Maintenance ontology
CB Case-based reasoning	Assumes that if a user likes a certain item, s/he will probably also like similar items. Recommends new but similar items.	No content analysis Domain-independent Quality improves	New user problem Overspecialisation Sparsity Cold start problem
CB Attribute-based techniques	Recommends items based on the matching of their attributes to the user profile. Attributes could be weighted for their importance to the user.	No cold start problem No new user/new item problem Sensitive to changes of preferences Can include non-item-related features Can map from user needs to items	Does not learn Only works with categories Ontology modelling and maintenance is required Overspecialisation

CF: Collaborative Filtering; CB: Content-Based

3.1 Problems concerning recommendation techniques

Among the disadvantages cited before, there are several classical problems in automatic recommenders that are necessary to explain.

3.1.1 Cold start problem

This is a classical problem affecting all of the learning-based techniques (i.e. collaborative, content-based and demographic) cited by Burke (2007). This problem consists of how handling new items or new users. In a collaborative system, for example, new items cannot be recommended to any user until they have been rated by someone. Recommendations for items that are new to the catalogue are therefore considerably weaker than more widely rated products, and there is a similar failing for users who are new to the system.

3.1.2 New user problem

This problem can be viewed as a case of the “cold start” problem (Burke, 2007). Because recommendations follow from a comparison between the target user and other users based solely on the accumulation of ratings, a user with few ratings becomes difficult to categorize.

3.1.3 New item problem

Similarly, a new item that has not had many ratings also cannot be easily recommended. This problem shows up in domains such as news articles where there is a constant stream of new items and each user only rates a few. It is also known as the “early rater” problem, since the first person to rate an item gets little benefit from doing so: such early ratings do not improve a user’s ability to match against others (Burke, 2003). This makes it necessary for recommender systems to provide other incentives to encourage users to provide ratings.

3.1.4 Overspecialisation

A typical problem with recommenders is over-specialization: users frequently see items that are very similar to what they liked in the past. While this approach produces relevant items, anecdotal evidence suggests that they may not be the most useful recommendations, due to their lack of novelty. Overspecialization leads to result sets with items that are too similar to one another, thus reducing the diversity of results and limiting user choices. Traditionally, the problem is addressed through attribute-based diversification—grouping items in the result set that share many common attributes (e.g., genre for movies) and selecting only a limited number of items from each group. It is, however, not always applicable, especially for social content recommendations.

3.2 Analysis of recommenders

The main criteria that will be analysed are the following:

- Type of recommender (single technique, hybrid, context-aware...)
- Techniques (knowledge-based, content-based, demographic...) and techniques combination
- Format of the metadata used for objects representation (SCORM, IEEE LOM, Dublin Core)
- User profiling (IEEE PAPI, IMS LIP)
- Potential disadvantages that the system could present according to the adopted strategy

The table below summarizes the indicated characteristics:

Recommender	Type	Techniques	Metadata format	User profiling	Potential disadvantages
Altered Vista	Single	Collaborative filtering	Not accessible		New user problem

					New item problem Popular taste Scalability Sparsity Cold start problem
RACOFI	Single	Collaborative filtering (STI Pearson and STIN2) and Item Average algorithm	Custom-made	Custom-made	New user problem New item problem Popular taste Scalability Sparsity Cold start problem
Isis	Hybrid	Ontology-based (content-based) and stereotype filtering (collaborative filtering)	Not mentioned	Not mentioned	Hybridization is supposed to overcome single techniques disadvantages (not enough information to claim it)
CYCLADES	Single	Content-based	Not mentioned	Not mentioned	New user problem New item problem Popular taste Scalability Sparsity Cold start problem
Multi-attribute	Single	Collaborative filtering	Not mentioned	Not mentioned	New user problem New item problem Popular taste Scalability Sparsity Cold start problem
CoFIND	Not available				
LORM	Hybrid	Knowledge-based, Collaborative filtering using Weighted result	SCORM, LOM	Custom-made	Weighed approach assumes implicitly that the relative value of the different techniques is more or less uniform across the space of possible items. But a collaborative recommender will be weaker for those items with a small number of raters
Simulation environment	Both	Collaborative filtering, ontology-based	Custom-made	Custom-made	N/A: Platform to test several techniques
RPL Recommender	Hybrid	Content-based, Collaborative filtering using Cascade + Feature augmentation	LOM	Custom-made	The application of such hybridization techniques can lead to one technique overshadowing the other

According to the table above, it can be said that most of the analysed recommenders are based on user-based collaborative filtering. This strategy has both advantages and disadvantages, but it seems that its ability to improve the quality of its recommendations over the time, and also its discovery capability (serendipity), are the differentiating factors. The rest of recommenders have chosen mainly content-based strategies that also seem to be quite popular. In both cases, it is clear that the use of single techniques has to deal with several issues, for example the “new user” problem. One way to resolve these issues is to apply hybridization, as RPL Recommender does; the other way consists of mitigate them by establishing custom-made solutions, as some others of the analyzed recommenders do. Amongst all the reviewed recommenders, those that have applied hybridization to enhance their results are more robust and richer than those that have simply applied isolated solutions to the different shortcomings attached to their recommendation strategy.

After reviewing the different models proposed, it seems that RPL Recommender is, not only the most evolved (because of being the most recent one and the advantage taken of previous works), but also one of the best documented. Authors have presented a well-documented introduction and a systematic description of the proposed system, introducing the last achievements in the matter. This work can be considered both as a reference and as a starting point for future works.

4 FUTURE WORK

In accordance with the present study, future work could consist in designing an automatic recommender for IR that overcomes detected difficulties. The core of the proposed solution would consist in an automatic recommender system based on a Knowledge-based algorithm combined with a Collaborative strategy using Feature Augmentation. This initial strategy could be enhanced by means of context-aware enrichments.

Knowledge-based technique generates its recommendations based on inferences about a user’s needs and preferences (Burke, 2002). Its approach is based on certain functional knowledge. This knowledge is about how a particular item meets a particular user need, and that leads it to reason about the relationship between a need and a possible recommendation. The same as the other studied techniques, this strategy applied separately has its own limitations, but it can be combined to beat them.

As stated before, CARS provide the necessary contextualization for generating more accurate results in the recommendations. By combining this context information with the output generated by a knowledge-based recommender, a better result will be obtained. Consequently, users should be modelled properly to acquire this contextual knowledge. It is convenient to base this characterization on any existing standard. Probably, LOM would be a good choice for this.

4.1 Proposed methodology

The methodological concerns can be divided in two parts. Firstly, a theoretical model of the proposed recommender will be designed. This can be considered the creative part of the study. The other part is the proposed methodology to evaluate the system in case it would be implemented and deployed in a real environment. So, in this stage it would be operating together with a real digital library, enhancing its main search and interaction features with recommendation capabilities.

The second part of the methodology is the evaluation of the designed recommender. According to the nature of the research objectives, the best way to evaluate the recommender is to perform a survey. This will allow us to gather statistical information about user experience when getting suggestions from the implemented recommender. After analyzing the data, conclusions will be summed up in a project report. However, the

proposed evaluation methodology is merely theoretical because the survey would need a reference implementation of the designed system.

ACKNOWLEDGMENTS

The following article has been supervised by Jordi Conesa Caralt as a research seminar of the Master of Multimedia Applications in the Universitat Oberta de Catalunya. I wish to thank him for making possible the continuing improvement of this work and for his technical support.

I also wish to especially thank my mother, for her material support, and my wife, for her emotional support.

REFERENCES

Adomavicius, G & Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering archive*, 17(6), 734-749.

Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. In *Recommender systems handbook* (pp. 217-253). Springer US.

Aksel, F. (2010). Enhancing Accuracy of Hybrid Recommender Systems Through Adapting the Domain Trends. M.S. Thesis. METU Library.

Avancini, H., & Straccia, U. (2005). User recommendation for collaborative and personalised digital archives. *International Journal of Web Based Communities*,1(2), 163-175.

Barker, P., & Campbell, L. M. (2010). Metadata for learning materials: an overview of existing standards and current developments. *Technology, Instruction, Cognition and Learning*, 7(3-4), 225-243.

Bellogín, A., Cantador, I., Díez, F., Castells, P., & Chavarriaga, E. (2013). An empirical comparison of social, collaborative filtering, and hybrid recommenders. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 4(1), 14.

Bellogín, A., Castells, P., & Cantador, I. (2011, July). Self-adjusting hybrid recommenders based on social network analysis. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval* (pp. 1147-1148). ACM.

Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *User modelling and user-adaptor interaction*,12(4), 331-370.

Burke, R. (2007). Hybrid Web Recommender Systems. *The adaptive Web. Lecture Notes in Computer Science*, 4321/2007, 377-408.

Conway, P. (2008). Modeling the digital content landscape in universities. *Library Hi Tech*, 26(3), 342-354. doi:10.1108/07378830810903283

DCMI (2008). Dublin Core Metadata Element Set, version 1.1. DCMI Recommendation. Retrieved 28 February 2009, from <http://www.dublincore.org/documents/dces/>

DCMI (2008b). DCMI Metadata Terms. DCMI Recommendation. Retrieved 28 February 2009, from <http://dublincore.org/documents/dcmi-terms/>

- Drachsler, H., Hummel, H. G., & Koper, R. (2008). Personal recommender systems for learners in lifelong learning networks: the requirements, techniques and model. *International Journal of Learning Technology*, 3(4), 404-423.
- Fernández-Tobías, I., Cantador, I., Kaminskas, M., & Ricci, F. (2012). Cross-domain recommender systems: A survey of the state of the art. In *Proceedings of the 2nd Spanish Conference on Information Retrieval. CERI*.
- Ferran, N., Casadesús, J., Krakowska, M., & Minguillón, J. (2007). Enriching e-learning metadata through digital library usage analysis. *The Electronic Library*, 25(2), 148–165. doi:10.1108/026404707110741296
- He, J & Chu, W.W. (2011). *Social Network-Based Recommender System (SNRS)*. Cambridge: Proquest, Umi Dissertation Publishing.
- Jain, P. (2011). New trends and future applications/directions of institutional repositories in academic institutions. *Library Review*, 60(2), 125–141. doi:10.1108/00242531111113078
- Khribi, M. K., Jemni, M., & Nasraoui, O. (2009). Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval. *Educational Technology & Society*, 12 (4), 30–42.
- Lynch, C. (2010). Where do we go from here? The next decade for digital libraries. *Notes*, 2010, 08-31.
- Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2011). Recommender systems in technology enhanced learning. In *Recommender systems handbook* (pp. 387-415). Springer US.
- Meyer, F., Fessant, F., Clérot, F., & Gaussier, E. (2012). Toward a new protocol to evaluate recommender systems. *arXiv preprint arXiv:1209.1983*.
- Minguillón, J., Rodríguez, M., & Conesa, J. (2010). Extending learning objects by means of social networking. *Advances in Web-Based Learning – ICWL 2010. Lecture Notes in Computer Science Volume 6483*, 2010, 220–229.
- Nilsson, M., Baker, T., & Johnston, P. (2008). The Singapore framework for Dublin Core application profiles. DCMI Recommended Resource. Retrieved 28 February 2009, from <http://dublincore.org/documents/singapore-framework/>
- Powell, A., Nilsson, M., Naeve, A., Johnston, P., & Baker, T. (2007). DCMI abstract model. DCMI Recommendation. Retrieved 28 February 2009, from <http://dublincore.org/documents/abstract-model/>
- Saracevic, T. (2004). Evaluation of digital libraries: an overview. *Notes of the DELOS WP7 workshop on the evaluation of Digital Libraries*, Padua, Italy.
- SCORM (2013). Advanced Distributed Learning SCORM Official Website (Consulted on Jan. 2013). <http://www.adlnet.org/scorm/>
- Smeaton, A. F., & Callan, J. (2005). Personalisation and recommender systems in digital libraries. *International Journal on Digital Libraries*, 5(4), 299–308. doi:10.1007/s00799-004-0100-1

Tsai, K. H., Chiu, T. K., Lee, M. C., & Wang, T. I. (2006, July). A learning objects recommendation model based on the preference and ontological approaches. In *Advanced Learning Technologies, 2006. Sixth International Conference on* (pp. 36-40). IEEE.

Verbert, K., Manouselis, N., & Ochoa, X. (2011). Context-aware recommender systems for learning: a survey and future challenges. *IEEE Transactions on Learning Technologies*, 6(1), 1–20.

Zuhadar, L., & Nasraoui, O. (2010). A Hybrid Recommender System Guided by Semantic User Profiles for Search in the E-learning Domain. *Journal of Emerging Technologies in Web Intelligence*, 2(4), 272–281. doi:10.4304/jetwi.2.4.272-281