

Intelligent Recovery Architecture for Personalized Educational Content

A. Gil, S. Rodríguez, F. De la Prieta, B. Martín, and M. Moreno

Abstract. Multi-agent systems are known for their ability to adapt quickly and effectively to changes in their environment. This work proposes a model for the development of digital content retrieval based on the paradigm of virtual organizations of agents. The model allows the development of an open and flexible architecture that supports the services necessary to conduct a search for distributed digital content dynamically. AIREH (Architecture for Intelligent Recovery of Educational content in Heterogeneous Environments) is based on the proposed model; it is a multi-agent architecture that can search and integrate heterogeneous educational content through a recovery model that uses a federated search. A major challenge in searching and retrieval digital content is to efficiently find the most suitable for the users. This paper proposes a new approach to filter the educational content retrieved based on Case-Based Reasoning (CBR). The model and the technologies presented in this research are an example of the potential for developing recovery systems for digital content based on the paradigm of virtual organizations of agents. The advantages of the proposed architecture are its flexibility, customization, integrative solution and efficiency.

Keywords: Multi-agent systems, e-learning, learning objects, repositories, Case Base Reasoning, recommender systems.

1 Introduction

The last decade has seen a significant evolution in the methods for managing and organizing large volumes of digital content. The information is characterized in

A. Gil · S. Rodríguez · F. De la Prieta · B. Martín · M. Moreno
Departamento de Informática y Automática – Facultad de Ciencias
University of Salamanca – Spain
Plaza de la Merced s/n
37008 Salamanca
e-mail: {abg, srg, fer, eureka, mmg}@usal.es

various ways, and contained in specialized data repositories each of which must be accessed by their own methods, resulting in dozens of communication protocols. Therefore, there is a need for research on the techniques, tools and methodologies that provide a technological solution that enables the adaptation of mechanisms in an environment that is the sum of open contexts highly dynamic, and heterogeneous.

The education sector is a significant generator, consumer and depository for educational content. These aspects position the current educational environment in a relevant sector in the development and integration of emerging solutions to management processes for the location and distribution of digital content. This openness in communication and the urgent need to achieve true interoperability between educational application environments have created the need for research in the search, retrieval and integration of heterogeneous educational content. The problems arising from the integration of educational content are usually caused by the multiple characterizations of the content, the vast amount of educational content distributed, and the access to them required by different users.

This paper proposes a model for the development of digital content retrieval based on the paradigm of virtual organizations of agents. The model allows the development of an open and flexible architecture called AIREH (Architecture for Intelligent Recovery of Educational content in Heterogeneous Environments) that supports the services necessary to conduct a personal search for distributed digital content dynamically.

The remainder of this paper is organized as follows: Section 2 describes relevant works related to Learning Object selection and recommendation, Section 3 introduces a new approach to applying CBR to the LO recommendation domain. Selected results of a comprehensive evaluation of the approach are presented in Section 4. The paper closes with relevant conclusions and an outlook to future work in section 5.

2 Related Work

With so many Learning Object Repositories, a major challenge is to efficiently find the most suitable LOs for the users. Researchers and developers of e-learning have begun to apply information retrieval techniques with technologies for recommendation, especially collaborative filtering [1], or web mining [2], for recommending educational content. A recent review of these applications can be seen in [3]. The features that handle these information filtering techniques in this context are the attribute information of education item (content-based approach) and the user context (collaborative approach).

One of the first works in this context was developed by Altered Vista: a system in which instructional techniques are evaluated based on collaborative filtering recommendation algorithms with close neighbors [4, 5]. These works explore how to collect user reviews of learning resources and propagate them in the form of word-of-mouth recommendations. RACOFI (Rule-Appling Collaborative Filtering) proposes a collaborative filtering by rules, with an architecture for the custom selection of educational content [5]. The author's recommendation combines both approaches to reduce recommendation by integrating a collaborative filtering

algorithm that works with user ratings of a set of rules of inference, which creates an association between the content and rate of recommendation. McCalla [6] proposes an improvement to collaborative filtering called the ecological approach to designing e-learning systems. Key aspects of this proposal take into account the gradual accumulation of information and focus on end users. Manouselis *et al.* [7] have conducted a case study with data collected from the CELEBRATE portal users to determine an appropriate collaborative filtering algorithm.

Some solutions take a hybrid approach. [8, 9, 10] make use of algorithms based on reviews from other users according to interests which are extracted through nearest neighbor algorithms. However, all of the selected learning objects are treated equally without any distinction between them, which would allow more precise assessment criteria of the user, affecting the very pattern of preference of the user.

The works by [11, 12] suggest the need for selecting learning objects by taking into account the educational content described by their metadata, which falls in line with this work. They propose a mechanism called Contextualized Attention Metadata (CAM) to capture information about the actions along the life cycle of learning objects, including their creation, labeling, supply, selection, use and maintenance. Others are based on semantic aspects by considering contextual information automatically from the student's cognitive activities and the LO content structure [13, 14].

There are other approaches that require direct human intervention in their assessment such a LORI (Learning Object Review Instrument) tool, based on the advice of [15].

All the evaluated proposals concluded that the incorporation of mechanisms to assess attributes related to the educational content as well as aspects of user context and their interaction with the content, create effective recommendation mechanisms. There are a growing number of papers proposing systems to recommend learning resources, evidenced by the lack of operational solutions as confirmed by recent work [16].

The architecture proposed in this paper provides multiple perspectives to assess the recovery of educational content from a real, open and scalable environment, and will also will be a support mechanism to implement the recommendation or ranking for the recovered LOs.

3 The AIREH Proposal

The cornerstone of this work is the recovery of learning objects in a real environment using federated search in different Learning Objects repositories. It is necessary to provide the user with a framework that unifies the search and retrieval of objects, thus facilitating the learning process that filters and properly classifies learning objects retrieved according to some rules. The generation of such rules for the organization of the items recovered will be based on educational metadata and will provide useful content to end users.

The structural diagram, see Figure 1, adapted according to the pattern of congregation AIREH unit, contains many features (and products). In addition,

supplier and customer roles are refined into these new units to specialize in the functionality or the utilization of specific services associated with these types of products.

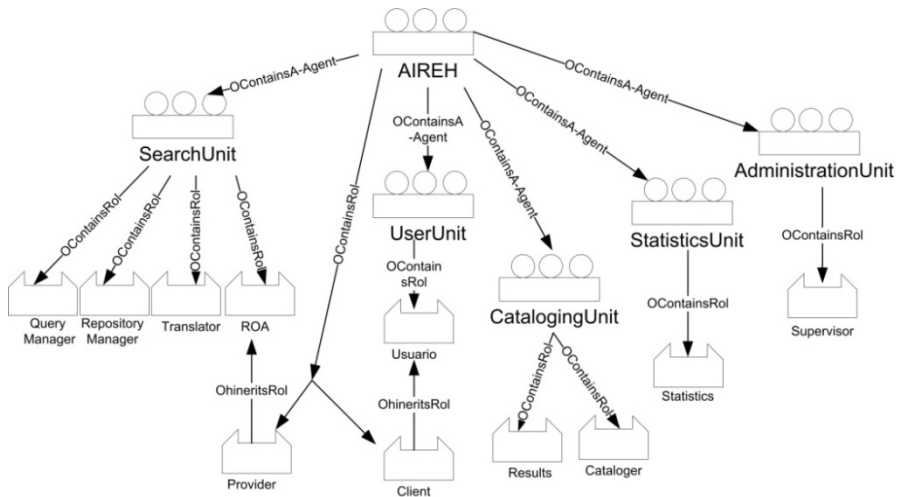


Fig. 1 Functional View

The federated search mechanism has been addressed as the resolution of the issue of content retrieval by solving three phases: (1) the selection of repositories, (2) the recovery of content, and finally (3) the merger of results. The first phase of the proposed federated search works by selecting the most appropriate ROA according to statistical parameters and performance during the search and retrieval session of content, and is based on user research through key terms. A second stage of filtering incorporates aspects of quality of the retrieved objects according to two criteria (1) the assessment of the quality of metadata retrieved and (2) the assessment of LO estimated by users through collaborative techniques. The sum of these criteria on the ranking of retrieved objects in the system provides a hybrid recommendation that begins with a refined content-based recommendation and collaborative features. This stage ends with the third phase of the federated search, and includes recommendations on merging the content, which improves the quality of the retrieved content for the user that generated the query.

3.1 Proposed Recommendation Strategy

The recommendations of real world sales are always based on knowledge about the items as well as about user tastes, preferences or interest determines the seller in the context. AI research has been focused the so-called recommender systems to model this tasks.

The Case Base Reasoning (CBR) is a particular search technique widely used in nearest neighbor recommender systems. Recovery techniques and their

adaptation to CBR techniques have become effective for the development of recommender systems [17].

The development of a single ordered list of Learning Objects that incorporates user's relevance criteria in this work is one of the tasks that the agent model AIREH implements with a CBR reasoning model. For this initial problem is defined based on the elements of the context, as shown in Figure 2. In this regard, CBR-BDI agents use a system of case-based reasoning.

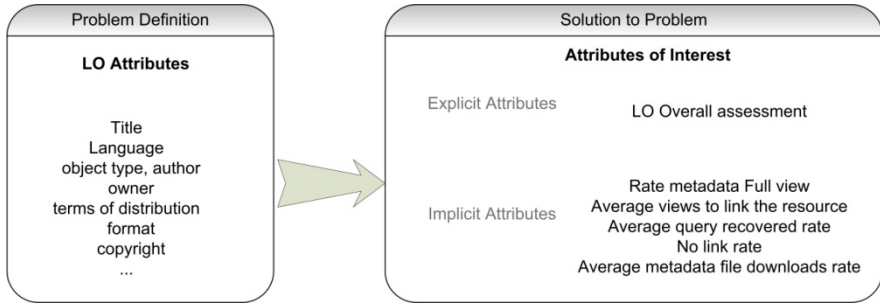


Fig. 2 Example of representation of cases in the LO domain

Given a set of educational content $p = \{p_1, p_2, \dots, p_i\}$, each of them, p_i , is characterized by a set of attributes from the set goals total T of possible attributes. So each unit of educational content is represented by a vector $p_i = \{p_{i1}, p_{i2}, \dots, p_{in}\}$. In the specific case of learning objects the total set of attributes defined, $T = \{title, language, keywords, format, \dots\}$, is extracted mainly from the information in the markup language in accordance with standard tagged used (LOM, DC, etc.) While attributes are introduced that reflect other aspects of the environment such as the repository to which it belongs.

Field Case	Element
USER	UserProfile
QUERY	Query
PREF	User Preference
STIME	Time Stamp

Fig. 3 Case Attributes

Figure 3 shows a description of the problem formalized by case attributes included to estimate other objective property but removing it from the set of objective parameters favoring a modular design ranking algorithm allows flexibility in implementation. Using these attributes, the system is able to get the description of the problem to solve.

Each user in the process has a different level of interest in each of the items. These interests are expressed by the user, as part of the so-called explicit or

attributes captured automatically by the system due to the interactivity of the user, generating the so-called implicit attributes. In this work it distinguishes the two types of parameters. The explicit parameter values user vote on any particular LO, while the implicit data is obtained through user interaction with the interface.

Each user has experience in several items. As each case represents the user experience on a particular article, the complete case base for a user is the user profile representation in the context modeling. Thus, the recommendation system maintains a database of cases representing each user profile.

Formally, the recommendation system is defined through a stage where there are m users $U=\{u_1, u_2 \dots u_m \}$ and a set of n LO, known as $O=\{o_1, o_2 \dots o_n \}$. In addition, each user u_i has a list of k evaluations carried out on a set of LO I_{ui} , where $I_{ui} \subseteq I$. In this context, a recommendation made by the active user $u_a \in U$, consists of a set of N learning objects $I_r \subset I$ which assesses interest to that user in context. The set of values of n users on m -learning objects, is a two dimensional array.

The recommendation mechanism used is based on a hybrid method that combines filtering techniques based on collaborative content aspects. This collaborative aspect comes from the feedback of users of the LO and is collected through the interface, and becomes part of the attributes explicit in the case base described in this section.

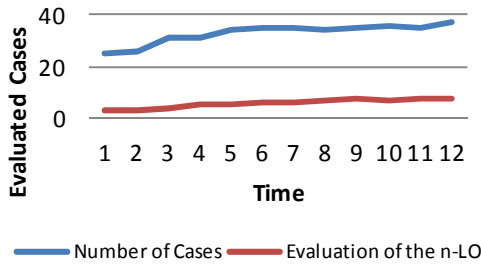


Fig. 4 Evaluation of the recommendations of the CBR

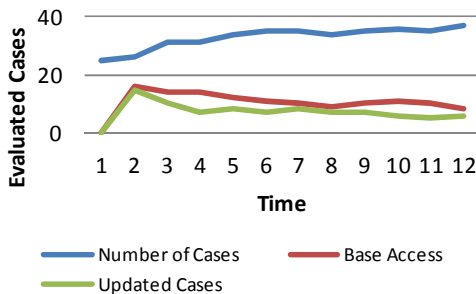


Fig. 5 Evolution of the CBR

4 Experimental Results

The recommendation is made according to the group of recovered cases. To validate the application, we compared the results obtained by the following evaluations of AIREH assessment by 6 months with 10 users. They perform a battery of queries from a selection of keywords from the computer science ground extracted from UNESCO codes.

The evolution of the number of cases in the case base allows for greater knowledge and appreciation of potential LO, as well as users. The Figure 4 shows how the user will increase the value of the first n recommended LO ($n = 5$) as the system solves new cases. Success in the system is evaluated through user interaction with the recommended LO, as well as the assessment it makes of each. This improvement is due to the system's ability to learn and adapt to lessons learned. Likewise, the experiences allow a better adaptation to the user profile.

Figure 5 shows that the number of updated cases decreases as the system acquires experiences (the x axis represents the number evaluations during the time period and the y axis quantifies the number of cases concerning the aspects evaluated (number of cases, access to the base and updates of cases). It is logical, since by increasing the variability of the cases captured and the number, the ability to find cases similar to the query that the user requires increases and may validate the recovered LO criteria defining user tastes and/or needs.

5 Conclusions and Further Work

A new approach for intelligent search of educational content was introduced by proposing an architecture called AIREH. This architecture is based on the application of virtual organizations of multi-agent systems and it has been used in a particular case study for federated search in repositories of learning objects and the subsequent recommendation of results.

This work demonstrate the possibility of modeling an efficient system for managing open systems from a model of an adaptive organization that provides personal recovery based on a CBR while is flexible and dynamic. The model and the technologies presented in this paper are an example of the potential for developing recovery systems for digital content based on the paradigm of virtual organizations of agents. The advantages of the proposed architecture are its flexibility, customization, integrative solution and efficiency.

Currently we have a prototype with some of the features implemented There are still many different perform evaluations to be covered for enhancing the performance of searching and retrieval educational content with personalization aspects. However, the research presented in this paper has provided suitable strategies for the future research towards enhanced content based retrieval systems.

Acknowledgements. This work has been partially supported by the project MICINN TIN 2009-13839-C03-03 and JCyL project SA225A11-2.

References

1. Bobadilla, J., Serradilla, F., Hernando, A., MovieLens: Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems* (2009), doi:10.1016/j.knosys.2009.01.008
2. Khribi, M.K., Jemni, M., Nasraoui, O.: Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval. *Educational Technology & Society* 12(4), 30–42 (2009)
3. Manouselis, N., Vuorikari, R., Van Assche, F.: Collaborative recommendation of e-learning resources: an experimental investigation. *Journal of Computer Assisted Learning* 26, 227–242 (2010)
4. Recker, M., Walker, A., Lawless, K.: What do you recommend? Implementation and analyses of collaborative information filtering of web resources for education. *Instructional Science* 31(4-5), 299–316 (2003)
5. Lemire, D., Boley, H., McGrath, S., Ball, M.: Collaborative Filtering and Inference Rules for Context-Aware Learning Object Recommendation. *Technology and Smart Education* 2(3), 179–188 (2005)
6. McCalla, G.: The Ecological Approach to the Design of E-Learning Environments: Purpose-based Capture and Use of Information about Learners. *Journal of Interactive Media in Education, Special Issue on the Educational Semantic Web* 1(7), 18 (2004)
7. Manouselis, N., Vuorikari, R., Van Assche, F.: Simulated Analysis of Collaborative Filtering for Learning Object Recommendation. In: *SIRTEL Workshop, EC-TEL* (2007)
8. Aijuan, D., Baoying, W.: Domain-based recommendation and retrieval of relevant materials in e-learning. In: *IEEE International Workshop on Semantic Computing and Applications 2008 (IWSCA 2008)*, pp. 103–108 (2008)
9. Ghauth, K., Abdullah, N.: Learning materials recommendation using good learners' ratings and content-based filtering. In: *Educational Technology Research and Development*. Springer, Boston (2010), <http://dx.doi.org/10.1007/s11423-010-9155-4>, ISSN 1042-1629
10. Wang, T.I., Tsai, K.H., Lee, M.C., Chiu, T.K.: Personalized Learning Objects Recommendation based on the Semantic Aware Discovery and the Learner Preference Pattern. *Educational Technology and Society* 10(3), 84–105 (2007)
11. Ochoa, X., Duval, E.: Use of Contextualized Attention Metadata for Ranking and Recommending Learning Objects. In: *Proceedings of 1st International Workshop on Contextualized Attention Metadata: Collecting, Managing and Exploiting of Rich Usage Information*, pp. 9–16 (2006)
12. Wolpers, M., Najjar, J., Duval, E.: Tracking Actual Usage: the Attention Metadata Approach. *Educational Technology & Society* 10(3), 106–121 (2007)
13. Han, Q., Gao, F., Wang, H.: Ontology-based learning object recommendation for cognitive considerations. In: *8th World Congress on Intelligent Control and Automation (WCICA)*, July 7-9, pp. 2746–2750 (2010)
14. Ruiz-Iniesta, A., Jiménez-Díaz, G., Gómez-Albarrán, M.: Personalización en Recomendadores Basados en Contenido y su Aplicación a Repositorios de Objetos de Aprendizaje. *IEEE-RITA* 5(1), 31–38 (2010)

15. Vargo, J., Nesbit, J.C., Belfer, K., Archambault, A.: Learning object evaluation: Computer mediated collaboration and inter-rater reliability. *International Journal of Computers and Applications* 25(3), 198–205 (2003)
16. Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., Koper, R.: Recommender Systems in Technology Enhanced Learning. In: *Recommender Systems Handbook*, pp. 387–415. Springer (2011)
17. Montaner, M., López, B., de la Rosa, J.L.: Opinion-Based Filtering Through Trust. In: Klusch, M., Ossowski, S., Shehory, O. (eds.) CIA 2002. LNCS (LNAI), vol. 2446, pp. 164–178. Springer, Heidelberg (2002)