

CBR Proposal for Personalizing Educational Content

A. Gil, S. Rodríguez, F. De la Prieta, J.F. De Paz, and B. Martín

Abstract. 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). AIREH (Architecture for Intelligent Recovery of Educational content in Heterogeneous Environments) is a multi-agent architecture that can search and integrate heterogeneous educational content within the CBR model proposes. The recommendation model and the technologies reported in this research applied to educational content are an example of the potential for personalizing labeled educational content recovered from heterogeneous environments.

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

1 Introduction

The significant evolution in the methods for managing and organizing large volumes of digital content in Internet has led to a phenomenon of globalization and decentralization. The education sector is a significant generator and consumer for digital content placed in large databases in Internet. The main problems to manage all the data is due to the various ways the information is characterized, contained in specialized data repositories each of which must be accessed by their own methods, resulting in dozens of communication protocols. Users have access to technologies that allow them to obtain educational content ubiquitously and in real-time. The information and technological overload become a problem in the search and selection for the educational contents. These aspects placed the current educational environment as a relevant sector in the development and integration of emerging

A. Gil · S. Rodríguez · F. De la Prieta · J.F. De Paz · B. Martín
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, fcofds, eureka}@usal.es

solutions to personalizing the management processes for the search, retrieval and integration of heterogeneous educational content.

In previous works has been presented an architecture recovery based educational content partner organizations called AIREH [1]. The main novelty was that proposed architecture has the capability of dynamic and adaptive planning to carry out an optimal distribution of the tasks of the organization's member agents enabling intelligent content retrieval and flexibility in the highly dynamic environment for which it have raised. The proposed architecture allows searching multiple repositories simultaneously based on federal search services in the environment of the called Learning Object Repositories (LOR). The LOR contains educational content, called Learning Object (LO) [2], and tagged in any of the educational metadata standards. The architecture AIREH supports mechanisms to implementing the recommendation or ranking for educational content recovered. This paper proposes a new approach to filter the educational content retrieved based on Case-Based Reasoning (CBR) as a basic feature in the active search of educational content within the architecture AIREH.

The remainder of this paper is organized as follows: Section 2 describes relevant works related to educational content recommendation; Section 3 introduces a new approach to applying CBR to the tagged educational content 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

Recommender systems select user information from tastes and preferences on a particular topic. The area of the store was the first field of application of these systems. Nowadays recommender systems are present in all areas of services offered via Internet such as music, news pages or virtual libraries. Its use is also taking a major impact on all types of e-learning systems [3]. This paper seeks to obtain a kind of repository customized for each user based on the effective management of search sequences in different repositories.

There are different methodologies to enable filtering information mechanisms in response to relevant aspects about user and context. To perform this task the system should learn from what the user finds as useful, interesting, or better, for which need to be in contact with the user. In addition to personalized information filtering the mechanism must be flexible to changes in user interests over time by establishing criteria for the treatment of the dynamic content.

The processes of learning and adaptation of recommender systems are based on the discovery of users' pattern through their online transactions or interactions with content. Learning methods are implemented in information filtering systems through automated reasoning mechanisms and learning mechanisms from the field of artificial intelligence mostly. Recommender systems are based on machine learning on the user's previous experiences in context. Among the various technologies for machine learning, this work makes use of Case-Based Reasoning (CBR) as a paradigm for learning and reasoning through experience. CBR models

allow the resolution of new situations by adapting solutions and customization through automatic reasoning about solutions to problems given above.

With so many educational content repositories, a major challenge is to efficiently find the most suitable contents for the users. This objective has attracted much research in the field of the selection and recommendation. Researchers and developers of e-learning have begun to apply information retrieval techniques with technologies for recommendation, especially collaborative filtering [4], or web mining [5], for recommending educational content. A recent review of these applications can be seen in [6]. 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).

Based on collaborative filtering recommendation algorithms with close neighbors are the first works developed by Altered Vista [7, 8]. These works explore how to collect user reviews of learning resources and propagate them in the form of word-of-mouth recommendations. Others as RACOFI (Rule-Appling Collaborative Filtering) propose a collaborative filtering by rules, 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 [9] have proposed an improvement to collaborative filtering that takes into account the gradual accumulation of information and focus on end users. Manouselis *et al.* [10] 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. [11-14] make use of algorithms based on reviews from other users according to interests which are extracted through nearest neighbor algorithms. These correlation-based algorithms are used to calculate an index score on the usefulness of learning objects through the analysis of comments from students with similar profiles. These algorithms improve preference-based selection algorithms by incorporating aspects of student preferences.

The works by [15-18] suggest the need for selecting learning objects by taking into account the educational content described by their metadata. Based on semantic aspects by considering contextual information from the student's cognitive activities and the LO content structure works [19, 20] propose learning objects recommendation to suit the student's cognitive activities through an approach based on ontologies.

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 [5]. But it is interesting to highlight that most of the evaluated proposals point out 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. A closer look at the revised proposals underscores the lack of applications on real systems and educational content. The architecture used in this paper, AIREH, provides multiple perspectives to assess the recovery of educational content from a real, open and scalable environment. The support to actual educational context allows proposing, implementing and making a full study by actual tests on recommendation for recovered LOs based on a CBR.

3 AIREH Framework

It is necessary to provide the user with a framework that unifies the search and retrieval of educational content. AIREH (Architecture for Intelligent Recovery of Educational content in Heterogeneous Environments) is a multi-agent architecture that can search and integrate heterogeneous educational content. AIREH is based on a model with dynamic and adaptive planning capabilities to carry out the optimal distribution of the tasks for the organization's member agents, thus enabling intelligent content retrieval and flexibility in a highly dynamic environment (users, profiles, features, content, variability of learning object repositories', etc.).

This framework, see Figure 1, thus facilitates the learning process based on digital contents because it retrieves and filters learning objects properly while classifies according to rules. The generation of such rules comes from the organization of the content recovered based on educational metadata.

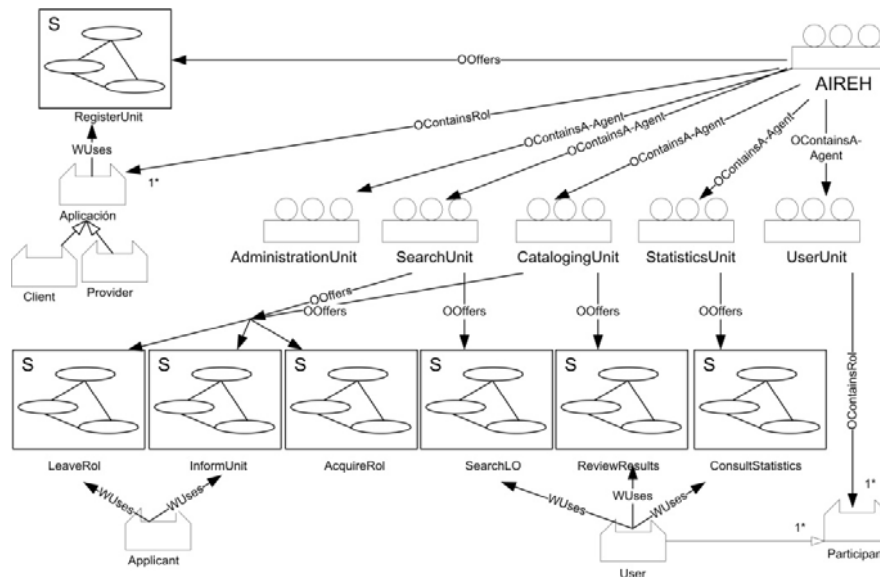


Fig. 1 Organizational model diagram congregational structure (external functionality)

The system offers educational resources as a final product encapsulated in the form of LO. These will be requested by users through a search process. It also provides statistical information such as product performance and use repositories of LO, identifying those that are used according to the search patterns. The mission of the organization will first maximize the system performance of queries by reducing time and increasing performance, and also, maximize the quality of results. The resolution of the issue of content retrieval has been addressed as the federated search mechanism by solving three phases: (1) the selection of repositories, (2) the recovery of content, and finally (3) the merger of results. The sum of several 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 Overview of the CBR Recommendation Strategy

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 [21, 22]. This involves the building of user models and the mechanism to anticipate and predict user preferences. In addressing these types of task, recommender systems draw on previous results from machine learning and other AI technology advances.

In order to bring the user with a single ordered list of Learning Objects that incorporates user's relevance criteria this work proposes a CBR reasoning model. In this regard, CBR-BDI agents use a system of case-based reasoning. To facilitate the construction of CBR-BDI agents, different jobs within the works of [22, 23] present a set of equivalence merge the two technologies through a generic library.

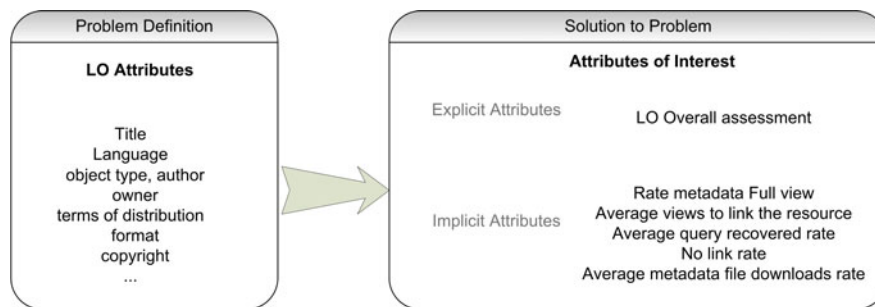


Fig. 2 Example of Case Representation in the educational content domain

A CBR depends largely on the structure and content representation and its collection of cases. The developed system is characterized by working with cases defined the characteristics of the educational context. Each case in the CBR case base is divided into two parts, see Figure 2. The first one describes a set of attributes of the named target (the definition of the problem in CBR terminology). The second element is the set of attributes that describe user interest in the issue because (the solution to the problem through a set of goals in CBR terminology).

With regard to the CBR cycle in the *recovery phase* the system retrieves similar items. Define the degree of similarity between the elements is the most important step in the recovery phase of CBR. The level of similarity between two items is computed using a global function of similarity, calculated by weighting based on weights of various measures of similarity. For numeric attributes, linear similarity functions were designed according to exponential. For attributes labeled, has made use of

WordNet semantic similarity algorithm [24]. With the efficient similarity measure defined, given a new case, we obtain an ordered list of similar cases. Using this feature of CBR, when a user is interested in a subject into the educational context, the recommendation system can generate a list ordered elements based on user profile.

The *reuse phase* consists of adapting the old solutions of retrieved cases for the new problem. Once the system has retrieved a set of above items (most similar), the system knows the user's interest defined by some interest attributes (solution in CBR terminology in Figure 2). Assuming the user interest in a new LO is similar to the interests of another user in similar LO in the reuse phase, the recommendation system calculates a value that encodes the degree of interest for the new element. This value is used to decide the ranking position of the set of LO recovered for the user depending on the grade.

The value of the interest, equation (1), is computed from the values of LO similar interests selected in the recovery phase. We calculate this interest by combining the numerical results assessing the degree of interest for each case in function of explicit and implicit interest.

$$I_c = \mathcal{M}(f^e(Int_j^e), f^i(Int_j^i)) \quad (1)$$

In the *revision phase*, the system evaluates the user's interest in the LO performed. The idea is to track user interaction through the system to find relevant information about the user's interest in the recommended LO through the list presented, as well as explicit and implicit information in order to retain the new case.

Finally, in the *retention phase*, the new element is inserted in the case base with the attributes of interest that were added in the revision phase.

The case definition provided in the CBR-BDI library provides a great power and flexibility to the developer, allowing them to customize the case to his particular problem. In this work the problem to be solved is to rank, according to criteria of customization, educational content retrieved from different repositories. For this initial problem is defined based on the elements of the context for which it is necessary to propose an ordered list of relevant results retrieved learning objects.

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.

4 Experimental Results

The recommendation is made by implementing the CBR proposed mechanism and according to the group of recovered cases. To validate the recommendation proposal, we evaluated the results obtained by the 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.

We evaluate the user perception about the quality of the recommendations made by the proposed mechanism along de CBR evolution. The evolution of the number

of cases in the case base allows for greater knowledge and appreciation of potential LO. 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.

The Figure 3 shows the user perception about the quality of the recommendations made by the proposed mechanism through the interface. Each user has evaluated over 12 months the quality of the learning objects in the first five places in the recommender. Success in the system is evaluated through user interaction with the recommended LO, as well as the assessment it makes of each. The user perceives an improvement in the time of the recommended LO.

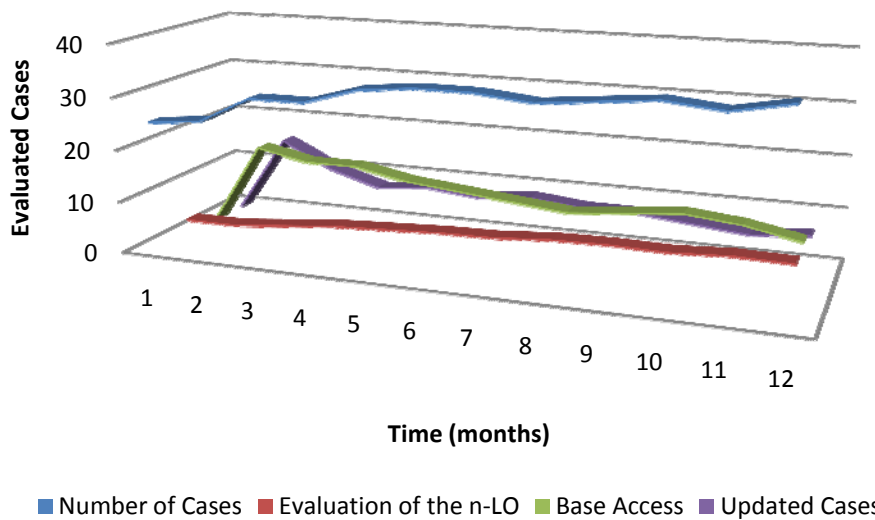


Fig. 3 Evaluation of the recommendations of the CBR

The Figure 3 shows also that the number of updated cases decreases as the system acquires experiences (the X axis represents the 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

The new approach for intelligent search of educational content is introduced by the framework prototype called AIREH. This architecture is based on the application of virtual organizations of multi-agent systems and is applied to a federated search in repositories of learning objects and the subsequent recommendation of results. AIREH provides multiple perspectives to assess recovery of educational

content from a real, open and scalable environment while the support to implement recommendations or ranking mechanisms for LO recovered.

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 type of recommendations used in this work lies in a system combining the hybrid type, since the active user is recommended with educational content that other users have taken in the same field of interest, but taking into account the history of the content selected by the user in advance. 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 an operative prototype with some of the features implemented. There are still many others featured are increasing the functionality while different perform evaluations are covered for enhancing the performance of searching and retrieval educational content with personalization aspects. However, the research presented in this paper working on a CBR as learning strategy applied to recommendation 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. Gil, A.B., De la Prieta, F., Rodríguez, S.: Automatic Learning Object Extraction and Classification in Heterogeneous Environments. In: Pérez, J.B., Corchado, J.M., Moreno, M.N., Julián, V., Mathieu, P., Canada-Bago, J., Ortega, A., Caballero, A.F. (eds.) Highlights in Practical Applications of Agents and Multiagent Systems. AISC, vol. 89, pp. 109–116. Springer, Heidelberg (2011)
2. Chiappe, A., Segovia, Y., Rincon, H.Y.: Toward an instructional design model based on learning objects. *Educational Technology Research and Development* 55, 671–681 (2007)
3. 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)
4. Bobadilla, J., Serradilla, F., Hernando, A., Lens, M.: Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems* (2009), doi:10.1016/j.knosys.2009.01.008
5. 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)
6. 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)
7. 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)

8. 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)
9. 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* (7) (2004); Special Issue on the Educational Semantic Web 1, 18 (2004)
10. Manouselis, N., Vuorikari, R., Van Assche, F.: Simulated Analysis of Collaborative Filtering for Learning Object Recommendation. In: *SIRTEL Workshop, EC-TEL* (2007)
11. 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)
12. Ghauth, K., Abdullah, N.: Learning materials recommendation using good learners' ratings and content-based filtering. In: *Educational Technology Research and Development*, SN 1042–1629. Springer, Boston (2010), <http://dx.doi.org/10.1007/s11423-010-9155-4>
13. Tsai, K.H., Chiu, T.K., Lee, M.C., Wang, T.I.: A learning Object Recommendation Model based on the Preference and Ontological Approaches. In: *Proceeding of the Sixth International Conference on Advanced Learning Technologies, ICALT 2006* (2006)
14. 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)
15. Yang, Y.: An evaluation of statistical approaches to text categorization. *Journal of Information Retrieval* 1, 67–88 (1999)
16. Kerkiri, T., Manitsaris, A., Mavridou, A.: Reputation metadata for recommending personalized e-learning resources. In: *Proceedings of the Second International Workshop on Semantic Media Adaptation and Personalization, Uxbridge*, pp. 110–115 (2007)
17. 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)
18. Wolpers, M., Najjar, J., Duval, E.: Tracking Actual Usage: the Attention Metadata Approach. *Educational Technology & Society* 10(3), 106–121 (2007)
19. 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)
20. Ruiz-Iniesta, A., Jiménez-Díaz, G., Gómez-Albarrán, M.: Personalización en Recomendadores Basados en Contenido y su Aplicación a Repositorios de Objetos de Aprendizaje. In: *IEEE-RITA*, vol. 5(1), pp. 31–38 (2010)
21. 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)
22. Corchado, J.M., Laza, R.: Constructing Deliberative Agents with Case-based Reasoning Technology. *International Journal of Intelligent Systems* 18(12), 1227–1241 (2003)
23. Glez-Bedia, M., Corchado, J.M., Corchado, E.S., Fyfe, C.: Analytical Model for Constructing Deliberative Agents. *Engineering Intelligent Systems* 3, 173–185 (2002)
24. Jiang, J.J., Conrath, D.W.: Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy CoRRcmp-1g/9709008 (1997)