Smart Recommendation for an Evolving E-Learning System

Tiffany Ya TANG and Gordon MCCALLA
Dept. of Computer Science, University of Saskatchewan
57 Campus Drive, Saskatoon, SK S7N 5A9, CANADA
yat751@mail.usask.ca, mccalla@cs.usask.ca

Abstract. The majority of current web-based learning systems are closed learning environments where courses and learning materials are fixed and the only dynamic aspect is the organization of the material that can be adapted to allow a relatively individualized learning environment. In this paper, we propose an evolving web-based learning system which can adapt itself not only to its users, but also to the open Web. More specifically, the novelty with respect to the system lies in its ability to find relevant content on the web, and its ability to personalize and adapt this content based on the system's observation of its learners and the accumulated ratings given by the learners. Hence, although learners do not have direct interaction with the open Web, the system can retrieve relevant information related to them and their situated learning characteristics.

1. Introduction

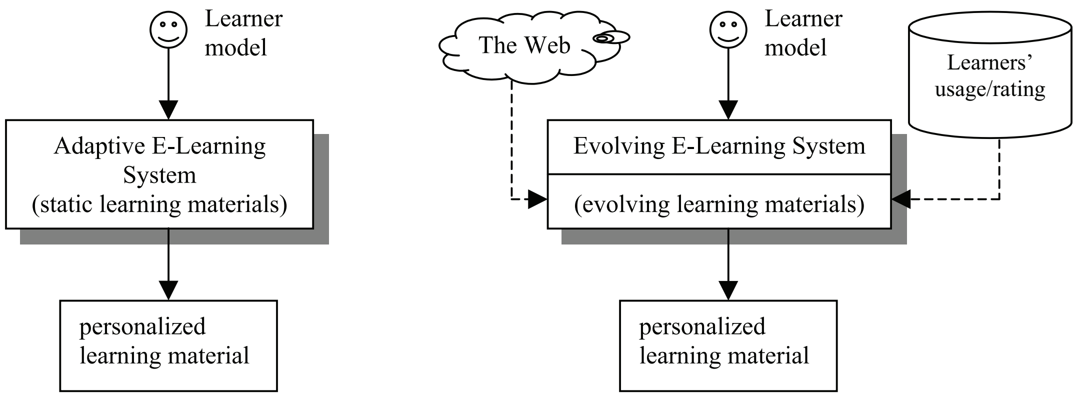
Research on e-learning has gained more and more attention thanks to the recent explosive use of the Internet. However, the majority of current web-based learning systems are closed learning environments, where courses and materials are fixed and the only dynamic aspect is the organization of the material that can be adapted to allow a relatively individualized learning environment.

In this paper, we will propose an evolving web-based learning system which can adapt itself not only to its users, but also to the open Web in response to the usage of its learning materials. Our system is open in the sense that learning items related to the course could be added, adapted, or deleted. Our proposed e-learning system adapts both to learners and the open Web. Figure 1 compares the traditional we-based adaptive learning system and our proposed open evolving learning system.

In a traditional adaptive e-learning system, the delivery of learning material is personalized according to the learner model. However, the materials inside the system are a priori determined by the system designer/tutor. In open evolving e-learning system, learning materials are automatically found on the web and integrated into the system based on users' interactions with the system. Therefore, although users do not have direct interaction with the open Web, new or different learning materials in the open Web can enrich their learning experiences through personalized paper recommendations.

1.1 A Brief System Introduction

Our tutorial system is designed to support an advanced course on data mining and web mining. It will contain 14 chapters, covering basic concepts and operations of data mining and web mining and their applications in e-business, intelligent tutoring, bioinformatics, recommender systems, etc. The system will initially consist of approximately 100 papers and two glossaries. We are also building a web crawler to crawl NEC's CiteSeer (the largest scientific digital library for computer science) to find new papers. The system is aimed at graduate students (computer science, bioinformatics, engineering, and business majors) and industry practitioners.



(a) Adaptive e-learning system (b) Evolving e-learning

Figure 1. A comparison of evolving e-learning system vs. adaptive e-learning system

There are two kinds of collaboration in the system. One is the collaboration between the system and the user, another is the collaboration between the system and the open Web. Users do not have direct interactions with the open Web; though the system can retrieve relevant information related to a learner and his/her learning characteristics.

In our proposed system, learners are free to learn whatever they want, which is different from the majority of other web-based learning systems. The 14 chapters of basic learning materials are fixed, what keeps changing is the collection of papers to be recommended to the learners. The novelty with respect to our proposed system lies in its evolving paper repository (through the crawling of CiteSeer), and its ability to make smart, adaptive recommendations based on the system's observations of learners' activities throughout their learning, and the accumulated ratings given by the learners. Each paper is tagged based on its content and technical aspects. Learners are required to give feedback (ratings) towards the papers recommended to them. Therefore, according to both the usage and ratings of a paper, the system will adaptively change a paper's tags, and determine whether or not the paper should be kept, deleted or put into a backup list.

Two of the major techniques that would be adopted include collaborative filtering and data clustering which have seldom been reported in the artificial intelligence in education literature.

1.2 What Makes Recommendations Different in E-learning from that in Other Domains

Making recommendations in e-learning is different from that in other domains (the most studied domain of recommender system is movie recommendations, [1-4]). Particular issues for an e-learning recommender system include:

- Items liked by learners might not be pedagogically appropriate for them

For example, a learner without prior background on the techniques of web mining may only be interested in knowing the state-of-the-art of web mining techniques in e-commerce. Then, it should be recommended that he/she read some review papers, for example, an editorial article by two of the leading researchers in this area [5], although there are many high quality technical papers related to his/her interest. On the other hand, for the learner coming from industry with some prior knowledge who wants to know how web mining can be utilized to solve e-commerce problems, [6] should be recommended, because the paper is the KDD-Cup 2000[1] organizers' report on how web mining can support business decision making for a real-life e-commerce vendor, and points out challenges, as well as lessons learned from the competition, which can benefit both researchers and industry practitioners.

By contrast in other domains, recommendations are made based purely on users' interests.

- Customization should not only be made about the choice of learning items, but also about their delivery [7].

  For example, some instructors will recommend learners to read an interesting magazine article, such as a related article in *Communications of ACM*, before a technical paper, because they believe it will help learners understand the technical paper and make them less intimidated. However, this is not the case in e-commerce recommendations, where site managers prefer to leave the list of recommended items unordered to avoid leaving an impression that a specific recommendation is the best choice [8].

In our proposed system, we will organize papers not only based on their main research categories, but also their technical levels. For example, review papers, workshop papers, highly technical papers etc.

In addition, making recommendations in the context of intelligent tutoring system is more tractable than in other domains in that learners' interests, goals, knowledge levels etc, may be better traced in a constrained learning environment.

1.3 Related Work

Related work can roughly be categorized into two areas of research: (i) dynamic curriculum sequencing and adaptive hypermedia, (ii) recommender systems.

1.3.1 Dynamic Curriculum Sequencing and Adaptive Hypermedia

Adaptive hypermedia has been studied extensively in the literature. Generally, there are two kinds of adaptation: adaptive navigation (link level) and adaptive presentation (content level) [9]. Early research in adaptive hypermedia concentrated mostly on content-based adaptation [10, 11], that is adaptively presenting the content of a given page or collections of pages which have been viewed by a user. The contents of the pages are used as clues to derive important features of learners such as their interests, knowledge state etc. Part of this branch of study can be viewed alternatively as content-based recommendations when users' past reading items/pages are recorded and analyzed. Over the past few years, link-orientated adaptation

[1] Since 1997, the most prestigious data mining conference, ACM SIGKDD Conference has organized a KDD CUP, an annual competition that aims to promote application-oriented KDD (Knowledge Discovery in Database) technologies. Each year, tasks and real-life data are provided. The data covers a wide range of applications. KDD CUP 2000 focused on e-commerce.

technologies are increasingly reported in the literature [12, 13, 14]. Among them, ELM-ART [13] is one of the most extensive hyper-educational systems (indeed, it is an adaptive on-line textbook), which supports several key features such as adaptive navigation, curriculum sequencing, and personalized diagnosis of learner solutions for learners with different prior knowledge. Generally, learning items are all pre-stored and not changeable; what keeps changing is the order in which course items are delivered, as also described in [15]. From this perspective, this work is different from our proposed system where learning items are dynamically added, modified or even deleted. Therefore, a learner might find different learning items if he/she logs on to the system at different times.

1.3.2 Recommender Systems

As described earlier, our proposed e-learning system makes individualized recommendations of materials for learners chosen from a dynamically evolving paper repository. There are several related works concerning tracking and recommending technical papers. Basu et al. [16] define the paper recommendation problem as: "*Given a representation of my interests, find me relevant papers*." They studied this issue in the context of assigning conference paper submissions to reviewing committee members. Reviewers do not need to key in their research interests as they usually do; instead, a novel autonomous procedure is incorporated in order to collect reviewer interest information from the web. Bollacker et al. [17] refine CiteSeer, through an automatic personalized paper tracking module which retrieves each user's interests from well-maintained heterogeneous user profiles. Woodruff et al. [18] discuss an enhanced digital book with a spreading-activation mechanism to make customized recommendations for readers with different types of background and knowledge. McNee et al. [19] investigate the adoption of collaborative filtering techniques to recommend papers for researchers. They do not address the issue of how to recommend a research paper; but rather, how to recommend *additional* references for a target research paper. In the context of an e-learning system, additional readings in an area cannot be recommended purely through an analysis of the citation matrix of the target paper, because the system should not only recommend papers according to learners' interests, but also pick up those not-so-interesting-yet-pedagogically-suitable papers for them [20]. In some cases, pedagogically valuable papers might not normally be of interest to learners and papers with significant influence on the research community might not be pedagogically suitable for learners. Therefore, we cannot simply present all highly relevant papers to learners; instead, a significantly modified recommending mechanism is needed.

In addition, common to all the above approaches is the utilization of citation technique [16 - 19], self-citation technique [17], fused citation technique [18] techniques, the first two of which have been widely adopted for document retrieval. In this paper, we might consider making use of the citation matrix to make recommendations. In order to keep up with the most up-to-date research on the subject, the system carries a paper-updating module powered by an imbedded web crawler, responsible for accommodating new papers and removing some old-fashioned papers. A description of this module will be presented later in this paper.

1.4 Organization of the Paper

The rest of this paper is organized as follows. In section 2, we will present the overall system architecture, followed in section 3 by a description of a running example of how the system makes recommendations to learners. In section 4, we will discuss details of each component of the system and techniques we are applying achieve their respective functions. We conclude this paper by discussing our current state of research, interesting applications of our research approach, and future work.

2. Overall System Architecture

Figure 2 illustrates the overall architecture of the system. Basic learning materials are in the form of chapters which cover both techniques and applications of data mining and web mining. There is a Paper Repository where papers related to the course are actively maintained through the Paper Maintenance Module, which includes a web crawler which can occasionally crawl CiteSeer to find more papers. Learners are responsible for assessing the papers recommended to them; they provide a user/ratings matrix to support collaborative filtering for paper recommendations. In addition to their active assessment of the papers, their browsing sequences, are used to group them into clusters of users with similar interests. It is through the Smart Recommender that personalized recommendations are made. Papers in the Repository are dynamically changing, with new papers added in by the web crawler, and obsolete papers being deleted based on learners' assessments. Hence, from the learner's perspective, the system is evolving to adapt to his/her learning activities and other learners.

Before we proceed to describe the components of the system, a simple running example is given in the next section to better illustrate a typical flow of the system.

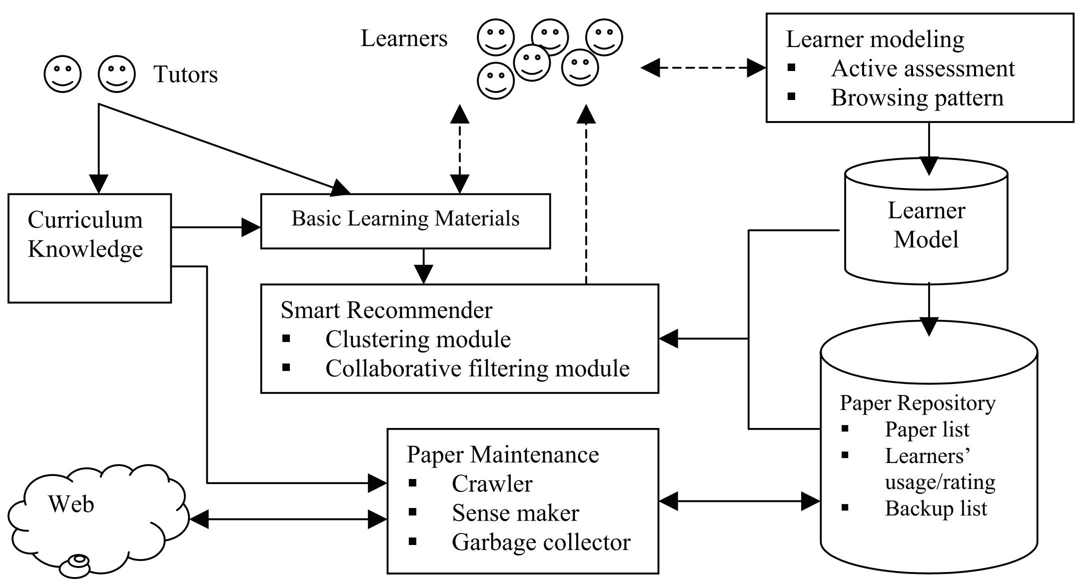


Figure 2. The architecture of the proposed system

3. A Simple Running Example

Suppose there are three learners Alex, John, and Jane, where Alex and John belong to the same cluster of learners, but they log on to the system at different times, and Jane belongs to another cluster of learners. Alex has been observed to browse pages related to web mining for e-commerce. He will be encouraged to read more on the topic, through a specific collection of research papers and magazine articles based on the technical quality he is seeking to explore. Meanwhile the system also learns that Alex tries to avoid papers with a high technical level, and is deemed to be a novice learner without much prior background knowledge. In all, Alex is grouped as *(novice learner, application, e-commerce)*, that is, Alex is a novice learner interested in web mining techniques for application problems, in particular e-commerce. In this case, a list of papers recommended to Alex is shown in Table 1.

John, although also grouped as *(novice learner, application, e-commerce)*, engages in his learning at a later time than Alex, and there have been some papers added in; therefore, the recommended papers are slightly different. In addition to the above three papers, two additional papers are added into the list as shown in Table 1 (Paper 6 and Paper 8 are highlighted in the table).

Jane is an advanced learner, with more interest in technical aspects of data mining for e-commerce. Therefore, Jane belongs to the group *(advanced learner, techniques, e-commerce)*, resulting in the five different papers on the recommendation list (Table 1).

Table 1. The list of papers recommended to Alex, John and Jane

Learner	Learner Cluster	Lists of Recommended Papers
Alex	Novice learner Application E-commerce	Paper 1, Paper 2, Paper 3
John	Novice learner Application E-commerce	Paper 1, Paper 2, Paper 3, **Paper 6, Paper 8**
Jane	Advanced learner Techniques E-commerce	Paper 10, Paper 11, Paper 13, Paper 15, Paper 20

It is clear that even for learners with the same interest (for example, in this case, e-commerce), there may be different papers to be recommended. Meanwhile, since the system is evolving, for learners logging on to learn through the system at different times, learning materials can be different and individualized for recommendations, compared to a traditional web-based learning system, where only the delivery order is individualized.

In the next section, we will describe in details the components of the system.

4. Components and Techniques

4.1 Paper Repository

E-versions of all papers, including magazine articles, conference papers, workshop papers, etc. will be stored in the Paper Repository.

4.1.1 Features of Papers Considered and Paper Tagging

Organization of papers is based on the following factors:

- Paper recency (publication date[2]), authors, etc. can be treated as special temporal and context features to use for matching purposes. For example, a recent conference paper,

[2] According to Paepcke et al [23], reputation of the publisher should also be counted to judge the value of a document. In our research since we will only include papers published in reputable conferences, workshops and journals, we do not consider publisher reputation to be a significant variable.

which previously appeared in a workshop, is more likely to be recommended when the system evolves.

- Publication type, e.g. journal paper, conference paper, workshop papers, technical reports etc.
- Paper length.
- Paper keywords. A paper's keyword(s) are mainly used to put it into a given category.

There are two kinds of tag associated with each paper: content and technical tag.
Content tags include:

- paper title
- category contents in terms of keywords
- publication year
- publication place
- author(s)
- paper length (used to determine whether the paper is a long, short or poster paper)
- publication type (journal/conference/workshop/technical reports)

Technical tag includes:

- technical level (for novice, medium or advanced learners)
- readability
- usefulness

Technical tags are usually added manually when the paper is newly added. It can be inferred and will be adjusted based on the feedback given by learners.

Some papers have already been acclaimed as seminal works in their research areas. Thus, these papers are stored and indexed as mandatory readings. For example, Agrawal et al's seminal work on association rule mining [21]. Indeed, in the context of our system, each paper can be regarded as a learning object, with appropriate tags for maintenance, updating, or deleting.

4.2 Smart Recommender

4.2.1 Data Clustering Module

Clustering learners based on their learning interests is handled by the data-clustering module. In our approach, we will perform a data clustering technique as a first step to coarsely cluster learners based on their learning interests. Basically, for each prototypical user group, there is a generalized representative path associated with it, which is pre-defined [22]. It is obvious that these representative paths are the centers of their respective clusters. Therefore, learners' browsing sequences can be compared with these representative paths in order to calculate the similarity. In addition, since a learner might fall into more than one cluster, the clustering algorithm should allow overlapping clusters. For example, a learner might be interested in web mining applications in user modelling as well as intelligent tutoring systems.

Clustering is good at finding a densely populated group of users with close similarities, but it fails to provide personalized information for these users. In order to make up for this, individualization can be achieved by further performing a collaborative filtering technique. The advantages to first apply clustering is not only to scale down the candidate sets, but also to guide collaborative filtering into a more focused area where high quality, personalized recommendations can be made [24].

4.2.2 Focused Collaborative Filtering Module

After clustering is performed, learners are categorized in clusters based on their learning interest. However, recommendations cannot be made at this point, because even for learners with similar learning interests, their ability to consume papers can vary due to the dissimilarity of their knowledge level (as shown in our example in the previous section). Therefore, during this process, recommendations will be made not on the whole pool of users as most recommender systems do, e.g. [1- 4], but on the clustered areas as illustrated in Figure 3 below. Accordingly, the number of recommended papers will also shrink.

Collaborative filtering (CF) works by making recommendations or predictions based on a database of accumulated explicit ratings. One of the key steps involved for CF is to form neighbors for a target user. There are two common similarity measurements in CF literatures: Pearson-correlation based and Cosine-based similarity. Since Pearson-correlation neighborhood forming outperforms the Cosine-based approach [25], we apply the former in our study [24]. Details of the algorithm will not be described here.

		Paper 1	**Paper 2**	**Paper 3**	**Paper 4**	**…….**	**Paper 100**
Cluster1	**Learner1**						
	Learner2		Reduced Area for Focused CF				
	Learner 3						
Cluster2	**Learner 4**						
	….		Reduced Area for Focused CF				
	Learner 10						

Figure 3. An illustration for focused CF

4.3 Intelligent Paper Maintenance Module

The maintenance module is mainly responsible for updating (including adding, deleting, putting into backup list), collecting, and making sense of papers. Figure 4 shows the components of this module.

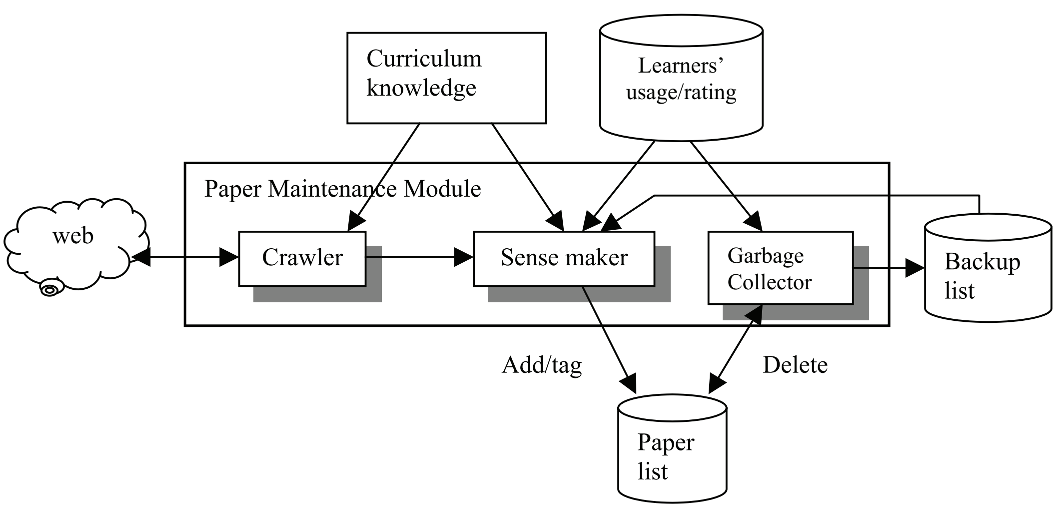


Figure 4 Intelligent Paper Maintenance Module

4.3.1 Topic-Driven Web Crawler

There is a web crawler embedded in this module, which is responsible for crawling NEC's CiteSeer to accommodate more up-to-date papers. Similar to the crawlers in, for example, [26, 27], our web crawler is a topic-driven web crawler[3], which exploits the content-similarity between course topics and candidate papers. In addition, we might also exploit the citation-matrix of each paper to retrieve papers closely related to the target paper(s). This technique has been successfully utilized in [17, 19, 28].

4.3.2 Sense-Maker

The Sense-Maker is mainly responsible for filtering out loosely related papers and grouping them into their appropriate topical categories. Paper tagging will be accomplished during this process, where the results of this process are candidate papers with appropriate tags. The sense making here is adaptively performed based on the collective learning behaviors and interests of users instead of an individual learner.

For example, the following paper can be categorized into at least two broad topics, which are web mining for e-commerce (application) and basics of association rule mining (techniques and operations):

> *Lin, W.Y., Alvarez, S.A., and Ruiz, C. Collaborative Recommendation via Adaptive Association Rule Mining. In Workshop of Web Mining for E-Commerce -- Challenges and Opportunities (WebKDD 2000), ACM-SIGKDD Conference on Knowledge Discovery in Databases (KDD'2000).*

The initial technical tag for the paper can possibly be set as of medium difficulty, which means that the paper is suitable for learners with some prior knowledge of association rule mining, since there are not many technical aspects described in the paper. But when there are accumulated ratings for the paper, the Sense-Maker can adaptively determine the appropriate technical tag for the paper. For instance, the majority of learners might find the paper to be highly technical which requires more extensive knowledge of both collaborative filtering and association rule mining, and their given ratings can be reflected in the paper's technical tag. Therefore, each paper's technical tag evolves according to the collective usage and ratings of its learners.

4.3.3 Garbage Collector

In order to keep the Paper Repository from growing too large, an intelligent Garbage Collector is used to decide whether or not to discard a paper completely or put it into a backup list for possible specialized needs. Although as pointed out in McCalla [33], patterns of user behavior might be needed to perform garbage collection, in our system, we will only focus on the usage of papers (as indicated below) as a reliable form of users' paper reading patterns, which indirectly determine the 'survival time' of a target paper. In addition, compared to the survival analysis proposed in [34], our module is simpler. In spite of it, we argue that in the context of our system, it is enough for us to capture both the overall usage and ratings of a target paper in order to determine whether or not to discard the paper.

---

[3] There has been huge amount of research concerning web crawlers, most notably [26, 27, 29, 30, 31, 32]. A detailed discussion of them is beyond the scope of this paper.

There are several criteria in determining whether a paper should be deleted or not. The first three factors, concerning the usage of a paper, mainly measure the frequency with which a target paper is browsed; the last two measure users' ratings of the target paper.

- Overall frequency (used in [35])
- Most recent frequency ([35])
- Overall cross-category frequency; Since as demonstrated in previous section, one paper might fall into more than two topical categories, its overall cross-category frequency measures its accumulative usage across these categories.
- Average rating
- Minimum acceptable rating

In our system, paper ratings are on a five-point scale, ranging from 1 to 5, the higher the better. If a paper consistently receives low ratings over a pre-defined period of time, it will be deleted.

5 Discussion and Future Work

Current web-based adaptive learning systems have been focusing on the interrelations between users and the system. Hence, the system, if deemed intelligent, must be capable of detecting users' needs, following their footsteps, and finally adapting to their needs. We argue that this is not enough. We have been ignoring the dynamics of the open Web. As such, we believe that two kinds of collaborations should be considered here: one is the collaboration between the system and its users; another is the collaboration between the system and the open Web in response to the changing needs of the users (Figure 1 (b)). A system, which can fulfill especially the second type of collaboration, would indeed help its users to keep up-to-date to the dynamics of information on the Web. It is evident that more rigorous collaborative research should be carried out between researchers from artificial intelligence in education, adaptive hypertext and hypermedia, web information retrieval, data mining, collaborative filtering, user modeling, intelligent user interfaces, computer supported collaborative work etc, in order to achieve these goals. It is our hope that our evolving e-learning system might be a first step towards a really open e-learning system.

Currently, the system is still being designed, the collaborative filtering module built, with our current focus on collecting the initial papers, and adding in other components. The techniques for the generalized clustering and focused collaborative filtering have been designed.

The system is expected to be 'alive' on the web when it is finished designing. As an initial step to test the system's functionality, we would create artificial learners to get a flavor of how the system could evolve as a result of its observations of these learners. We will conduct a human-subject study in order to gain insights to the issues involved in the system, once the system is built.

Acknowledgement

We would like to thank the National Sciences and Engineering Council of Canada (NSERC) for their financial support for this research.

References

[1] Basu, C., Hirsh, H. and Cohen. W.W. (1998) Recommendation as classification: using social and content-based information in recommendation. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI/IAAI 1998)*, 714-720.

[2] Herlocker, J.L., Konstan, J.A., Borchers, A. and Riedl, J. (1999) An algorithmic framework for performing collaborative filtering. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'99)*, Berkley, USA. 1999. 230-237.

[3] Schein, A., Popescul, A., Ungar, L. and Pennock, D. Methods and metrics for cold-start recommendations. (2002) In Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'02), Tampere, Finland, August 2002.

[4] Melville, P., Mooney, R. and Nagarajan, R. Content-boosted collaborative filtering for improved recommendations. (2002) In *Proceedings of the 18th National Conference on Artificial Intelligence (AAAI-2002)*, Edmonton, Canada. 2002. 187-192.

[5] Kohavi, R. and Provost, F. (2001) Applications of data mining to electronic commerce. *Data Mining and Knowledge Discovery*, an editorial of the special issue of Data Mining on Electronic Commerce, 5(1/2): 1-7.

[6] Kohavi, R., Brodley, C., Frasca, B., Mason, L. and Zheng, Z. J. (2000) KDD-Cup 2000 organizers' report: peeling the onion. *SIGKDD explorations*, 2(2): 86-98.

[7] Kobsa, A., Koenemann, J., Pohl, W. (2001) Personalized hypermedia presentation techniques for improving online customer relationships. *The Knowledge Engineering Review* 16(2), 111-155.

[8] Schafer, J.B., Konstan, J.A., and Riedl, J. (2001) Electronic commerce recommender applications. *Data Mining and Knowledge Discovery*, 5, (1/2), 115-152.

[9] Brusilovsky, P. Adaptive hypermedia. (2001) *User Modeling and User Adapted Interaction*, Ten Year Anniversary Issue (Alfred Kobsa, ed.) **11** (1/2): 87-110. 2001.

[10] De Rosis, F., De Carolis, B and Pizzutilo, S. (1993) User tailored hypermedia explanations. *INTERCHI'93 Conference Proceedings: Conference on Human Factors in Computing Systems, INTERACT'93 and CHI'93*, Amsterdam, The Netherlands. 169-170. 1993.

[11] Boyle, C., and Encarnacion, A.O. (1994) MetaDoc: an adaptive hypertext reading system. *User Models and User Adapted Interaction*. 4, 1-19. 1994.

[12] De Bra, P. and Calvi, L. (1998) AHA! An open adaptive hypermedia architecture. *The New Review of Hypermedia and Multimedia.* 4, 115-139. 1998.

[13] Weber,G. and Peter Brusilovsky, P. (2001) ELM-ART: an adaptive versatile system for web-based instruction. *International Journal of Artificial Intelligence in Education.* 12: 1-35. 2001.

[14] Brusilovsky, P. and Rizzo, (2002) R. Map-Based Horizontal Navigation in Educational Hypertext. *Journal of Digital Information.* 3 (1). 2002.

[15] Stern, M.K. and Woolf, B.P. (1998) Curriculum sequencing in a web-based tutor. In *Proc. 4th Intl' Conf. on Intelligent Tutoring Systems (ITS'98)*. 574-583.

[16] Basu, C, Hirsh, H., Cohen, W. and Nevill-Manning,C. (2001) Technical paper recommendations: a study in combining multiple information sources. *Journal of Artificial Intelligence Research*, 1, 231-252.

[17] Bollacker, K.D., Lawrence, S. and Giles, C.L. (1999). A system for automatic personalized tracking of scientific literature on the web. In *Proc. ACM Conference on Digital Libraries (DL 1999)*, 105-113.

[18] Woodruff, A., Gossweiler, R., Pitkow, J., Chi, E. and Card, S.K. (2000) Enhancing a digital book with a reading recommender. In *Proc. ACM CHI 2000.*153-160. 2000.

[19] McNee, S.M., Albert, I., Cosley, D., Gopalkrishnan, P., Lam, S.K., Rashid, A.M., Konstan, J.A and Riedl, J. (2002) On the recommending of citations for research papers. In *Proceedings of ACM International Conference on Computer Supported Collaborative Work (CSCW'02)*, 116-125. 2002.

[20] Tang, Y. T. and McCalla, G. (2003a) Towards pedagogy-oriented paper recommendation and adaptive annotations for a web-based learning system**.** To appear, in Workshop on Knowledge Representation and Automated Reasoning for E-Learning Systems, *18th International Joint Conference on Artificial Intelligence (IJCAI 2003)*, Acapulco, Mexico. August 9-15, 2003.

[21] Agrawal, R., Imielinski, T. and Swami, A. (1993) Mining association rules between sets of items in large databases. In *Proceedings of the ACM SIGMOD Int'l Conf. On Management of Data (ACM SIGMOD'93)*, Washington. 207-216.

[22] Tang, Y. T. and McCalla, G. (2002) Student modeling for a web-based learning environment: a data mining approach. In Proceedings of the 18th National Conference on Artificial Intelligence (AAAI 2002), Edmonton, Canada, July 28-August 1, 2002. 967-968. AAAI Press.

[23] Paepcke, A., Garcia-Molina, H., Rodriguez-Mula, G. and Cho, J. (2000) Beyond document similarity: understanding value-based search and browsing technologies. *SIGMOD Records*, March 2000. 29(1): 80-92.

[24] Tang, Y. T. and McCalla, G. (2003b) Mining the implicit ratings for focused collaborative filtering for paper recommendations**.** To appear, in Workshop on User and Group Models for Web-based Adaptive Collaborative Environments, *9th International Conference on User Modeling (UM 2003)*, June 2003, Johnstown, U.S.A. 2003.

[25] Breese, J.S., Heckerman, D. and Kadie, C. (1998) Empirical analysis of predictive algorithms for collaborative filtering. In *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI 1998).*

[26] Amento, B., Terveen, L. and Hill, W. (2000) Does "authority" mean quality? Predicting expert quality ratings of web documents. In *Proceedings of the 23rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2000),* 296-303. 2000.

[27] Diligenti, M., Coetzee, F., Lawrence, S., Giles, C.L. and Gori, M. (2000) Focused crawling using context graphs. In *Proceedings of the 26th International Conference on Very Large Databases (VLDB 2000),* Cairo. Morgan Kauffmann Publishers. 527-534.

[28] Popescul, A., Flake, G.W., Lawrence, S., Ungar, L.H. and Giles, C.L. (2000) Clustering and identifying temporal trends in document databases. In *Proceedings of IEEE International Conference on Advances in Digital Libraries (ADL 2000),* Washington, 173-182.

[29] Menczer, F. and Belew R. (2000) Adpative retrieval agents: internalizing local context and scaling up to the web. *Machine Learning*, 39(2/3): 203-242.

[30] Menczer, F., Pant, G., Srinivasan, P. and Ruiz, M.E. (2001) Evaluating topic-driven web crawlers. In *Proceedings of the 24th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '01).* September 2001, New Orleans, U.S.A.

[31] Chakrabarti, S., van den Berg, M. and Dom, B. (1999). Focused crawling: a new approach to topic-specific web resources discovery. In *Proceedings of the 8th ACM International World Wide Web Conference (WWW8),* 1999.

[32] Cho, J., Garcia-Molina, H. and Page, L. (1998) Efficient crawling through url ordering. In *Proceedings of the 7th ACM International World Wide Web Conference (WWW7).*

[33] McCalla, G. (2000) The fragmentation of culture, learning, teaching and technology: implications for the artificial Intelligence in education research agenda in 2010. *International Journal of Artificial intelligence in Education*. 11(2): 177-196. 2000.

[34] Pitkow, J. and Pirolli, P. (1997) Life, death, and lawfulness on the electronic frontier. In *Proceedings of ACM CHI 1997*, 383-390.

[35] Debevc, M., Meyer, Be. And Svecko, R. (1997) An adaptive short list for documents on the world wide web. In *Proceedings of ACM International Conference on Intelligent User Interface (IUI 1997),* 209-211. Orlando. U.S.A.