



Periodicals  Literature

Keyword  Title  Author  Topic

The Free Library > Business and Industry > Business, International > International Journal on E-Learning > January 1, 2005  
The Free Library > Science and Technology > Electronics > International Journal on E-Learning > January 1, 2005  
The Free Library > Social Sciences > Education > International Journal on E-Learning > January 1, 2005  
The Free Library > Date > 2005 > January > 1 > International Journal on E-Learning

## Smart recommendation for an evolving e-learning system: architecture and experiment.

### [Enterprise E-learning](#)

Enterprise Rapid E-Learning - what to consider. In-depth whitepaper  
[www.mohive.com](http://www.mohive.com)

### [Platform Symphony](#)

High performance grid solutions for financial services.  
[www.platform.com](http://www.platform.com)

### [CORBA Middleware](#)

CORBA ORBs and Services from the smallest DSP to the largest network  
[www.prismtech.com/middleware](http://www.prismtech.com/middleware)

Ads by Google

[Link to this page](#)



In this article, we proposed an evolving e-learning system which can adapt itself both to the learners and to the open Web, and we pointed out the differences of making recommendations in e-learning and other domains. We propose two pedagogy features in recommendation: learner interest and background knowledge. A description of a paper's value, similarity, and ordering are presented using formal definitions. We also study two pedagogy-oriented recommendation techniques: content-based and hybrid recommendations. We argue that while it is feasible to apply both of these techniques in our domain, a hybrid collaborative filtering technique is more efficient to make "just-in-time" recommendations. In order to assess and compare these two techniques, we carried out an experiment using artificial learners. Experiment results are encouraging, showing that hybrid collaborative filtering, which can lower the computational costs, will not compromise the overall performance of the recommendation system. In addition, as more and more learners participate in the learning process, both learner and paper models can better be enhanced and updated, which is especially desirable for web-based learning systems. We have tested the recommendation mechanisms with real learners, and the results are very encouraging.

\*\*\*\*\*

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.

### [E learning](#)

User-friendly classroom management software - Download free trial  
[www.netop.com](http://www.netop.com)

### [3G Mobile Backhaul](#)

A Multiservice Ethernet solution [www.ciena.com](http://www.ciena.com)  
[www.ciena.com](http://www.ciena.com)

Ads by Google

In this article, we 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 either manually or automatically. Our proposed e-learning system adapts both to learners and the open Web. Figure 1 compares the traditional web-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/instructor. In an open evolving e-learning system, learning materials are automatically/manually found on the Web, but automatically 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 recommendations.

The rest of this article is arranged as follows. In the rest of this section, we will briefly discuss the overall system architecture. We will also point out the uniqueness of making recommendations for e-learning systems through a motivational example, followed by discussions on some background information on recommendation system and related work. In the second section, we will present the architecture and components of our system. In the third section, our proposed pedagogically-oriented paper recommendation techniques and concepts will be discussed in details, experiment results will also be shown in this section. We will include lessons learned from implementing our recommendation techniques and human subject studies in the fifth section.

### A Brief System Introduction

Our proposed system is designed to support an advanced course for senior undergraduate or graduate students, especially when they are required to read technical articles related to their course, such as journal articles, conference papers, book chapter, etc. In the rest of this article, the term "papers" is used to refer to those articles. However, this system can be generalized to a broader area, for example, for corporate learners or with richer learning materials, such as slide presentation, technical report, textbook, programming code, etc. As shown in Figure 1, there are two kinds of collaboration in the system: the collaboration between the system and the user, and the collaboration between the system and the open Web. The novelty with respect to our proposed system lies in its evolving paper repository, and its ability to make smart, adaptive recommendations based on the system's observations of learners' activities through out their learning and the accumulated ratings given by the learners.

[FIGURE 1 OMITTED]

To achieve the system goal, each paper must be tagged based on its content and technical aspects. Moreover, 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. Since new papers are added into the system and useless papers are deleted from the system, the system evolves according to the usage by the learners. Thus, the most important parts in the system are the recommendation module and the paper maintenance module, which become the main focus of our research.

Figure 2 illustrates the overall architecture of the system. 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 specified digital libraries to find more papers. In addition to automatic search, authorized instructors and learners are allowed to suggest papers to the system. The paper maintenance module will verify the suggested title by searching paper attributes from available digital libraries, and then it will refine the annotation of the paper according to the ratings by learners. It is through the recommendation module that personalized recommendations are made. The recommendation module consists of two sub-modules: the data clustering module and the focused collaborative filtering module. The data clustering module will cluster learners into a sub-class according to the purpose of the recommendation, while the focused collaborative filtering module will find the closest neighbor (s) of a target learner and recommend paper(s) to him/her according to the ratings by those closest neighbor(s). Tutors are responsible to set up the curriculum and provide basic learning material such as introduction part. Based on this information, the system can select a set of papers and find new papers if any. Learners are responsible to give ratings and other assessments at the beginning or during the middle of learning. For more detailed descriptions of each module, see (Tang, & McCalla, 2003a)

[FIGURE 2 OMITTED]

## What Makes Recommendations in E-learning Different 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, (Basu, Hirsh, & Cohen, 1998; Herlocker, Konstan, Borchers, & Riedl, 1999; Schein, Popescul, Ungar, & Pennock, 2002; Melville, Mooney, & Nagarajan, 2002). Particular issues for an e-learning recommender system include:

- \* Items liked by learners might not be pedagogically appropriate for them;
- \* Customization should not only be made about the choice of learning items, but also about their delivery (Kobsa, Koenemann, & Pohl, 2001);
- \* Learners are not expected to read too many papers.

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, although there are many high quality technical papers related to his/her interest. By contrast, in other domains, recommendations are made based purely on users' interests.

For the delivery of papers, 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 (Schafer, Konstan, Riedl, 2001). In our proposed system, we will organize papers not only based on their main research categories, but also their technical levels. In addition, making recommendations in the context of intelligent tutoring systems is more tractable than in other domains since learners' interests, goals, knowledge levels etc., may be better traced in a constrained learning environment.

Finally, the amount of papers recommended to a learner in each course is limited. It is commonly accepted that each learner read no more than twenty papers in a course. In some cases, the amount may less than five. However, the amount of ratings by learner can affect the accuracy of collaborative filtering. Besides, in most collaborative filtering, cold-start problem remains one of the most important problems. Cold-start problem is the situation when there is no rating available for a new paper, or there is a new user who has not rated any paper. The common solutions are by assigning a set of pre-selected items or randomly assigned a set of items. Since in e-learning we cannot ask every learner to read too many papers and new papers are published every year, the solutions above are not appropriate in our system. (Tang, & McCalla, 2004b) describe the adoption of artificial learners to successfully solve the cold-start problem in our system.

It is commonly recognized that the sources of data on which recommendation algorithms can perform include users' demographic data. In our article, we will consider a special kind of user data different from the majority of recommender systems, (i.e., pedagogically-oriented data). The pedagogically-oriented data is different in the sense that it can directly affect as well as inform recommendation process, thus enhancing the quality of recommendations in the context of the web-based learning environments. The main pedagogical features used are the learner's goal and background knowledge, although other factors such as learning preferences are also important. To illustrate, consider the three learners A, B, and C in Table 1.

Suppose we have already made recommendations for the learners A, B and C. From Table 1, we can conclude that learner A and B have some over-lapping interests, but since their knowledge background differs, especially with respect to their technical background, the papers recommended to them would be different. But for learner B and C, although they have different application interests, their technical background is similar; therefore, they might receive similar technical papers. In the next section, we will describe related work and underlying technique.

## BACKGROUND KNOWLEDGE: RECOMMENDER SYSTEM

There are two basic approaches to providing personalized recommendations: content-based

and collaborative filtering (Jameson, Konstan, & Riedl, 2002). Regardless of the approach used, at the core of personalization is the task of building a model of the user.

#### Content-based filtering approach

Content-based approaches recommend items purely based on the contents of the items a user has experienced/consumed before. Representative content-based recommender systems include News Dude (Billsus, & Pazzani, 1999) and WebWatcher (Joachims, Freitag, & Mitchell, 1997). Since user profiles in the content-based approach are built through an association with the contents of the items, this approach tends to be quite narrowly focused and with a bias towards highly scored items. Moreover, the content-based approach only considers the preferences of a single user. The Collaborative Filtering (CF) approach is an approach capable of exploiting information about other similar users.

#### Collaborative filtering approach

CF makes recommendations by observing like-minded groups. It works by matching a target user against his/her neighbors who have historically had similar preferences to him/her. GroupLens is a pioneer rating-based automatic recommendation system which successfully adopts the CF approach (Resnick, Iacouvou, Suchak, Bergstrom, & Riedl, 1994). Firefly is another rating-based CF system for music albums and artists (Shardanand, & Maes, 1995). Compared to a content-based approach, the CF approach has received more popularity and worked very successfully both in research and practice (e.g., Melville, Mooney, & Nagarajan, 2002; Sarwar, Karypis, Konstan, & Riedl, 2000).

#### Hybrid approach

A purely content-based approach only considers the preferences of a single user, and concerns only the significant features describing the content of an item; whereas, a purely CF approach ignores the contents of the item, and only make recommendations based on comparing the user against clusters of other similar users. By combining these two techniques, perhaps we can have both individual as well as collective experiences with respect to the items being recommended. (Balabanovic, & Shoham, 1997) is one of the representative hybrid recommender systems.

## RELATED WORK

### Paper Recommendation

There are several related works concerning tracking and recommending technical papers. Basu et al., (2001) 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, Lawrence, and Giles (1999) refined CiteSeer, through an automatic personalized paper tracking module which retrieves each user's interests from well-maintained heterogeneous user profiles. Woodruff, Gossweiler, Pitkow, Chi, and Chard, (2000) discuss an enhanced digital book with a spreading-activation mechanism to make customized recommendations for readers with different types of background and knowledge. McNee, Albert, Cosley, Gopalkrishnan, Lam, Rashid, Konstan, and Riedl, (2002) investigate the adoption of collaborative filtering techniques to recommend papers for researchers. They only study 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. 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. Recker, Walker, and Lawless (2003) study the pedagogical characteristics of a web-based resource through Altered Vista, where teachers and learners can submit and review comments provided by learners. However, although they emphasize the importance of the pedagogical features of these educational resources, they did not consider the pedagogical features in making recommendations.

## Dynamic Curriculum Sequencing and Adaptive Hypermedia

Recently, adaptive hypermedia has been studied extensively. According to Brusilovsky (2001), there are two kinds of adaptation: adaptive navigation ("link level") and adaptive presentation ("content level"). Adaptive presentation is then sub-grouped into text and multimedia adaptation, while adaptive navigation is mainly sub-grouped into link ordering (Kaplan, Fenwick, & Chen, 1993), link annotation (Pazzani, Muramatsu, & Billsus, 1996), and link hiding (including removal, hiding and disabling (De Bra, & Calvi, 1998)). Early research in adaptive hypermedia has concentrated mostly on adaptive presentation technology (Boyle, & Encarnacion, 1994), capable of adaptively presenting the content of a given page or collections of pages which have been viewed by a user. More recently, more aspects of learners are utilized in order to tailor the delivered content, (e.g., Stern, & Woolf, 2000). It is obvious that the contents of the pages are used as clues to derive important learning features of students such as their interests, knowledge state, etc. From another perspective, 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 technologies are increasingly reported in the literature (De Bra, & Calvi, 1998; Weber, & Brusilovsky, 2001).

## Document Value

The majority of scientific literature as well as other document retrieval systems (and many other systems) has been focusing on finding documents relative to users' interests, to name a few (Bollacker et al., 1999; McNee et al., 2002). Recently, there have been approaches that augment the mostly commonly adopted similarity-based retrieving. Among them, Paepcke, Garcia-Molina, Rodrigues-Mula, and Cho (2000) propose a context-aware content-based filtering. In particular, context-aware content-based filtering attempts to determine the contextual information about a document, for example, the publisher of the documents, the time when the document was published, etc. For instance, they argued that "documents from the New York Times might be valued higher than other documents that appear in an unknown publication context." This contextual information provides additional rich information for users, thus constitutes a very important aspect of the value-ness of the item.

Our proposed approach takes into account of one type of contextual information: the pedagogical feature of learners. In particular, we argue that users' pedagogical goal and interest should be regarded as two of the most critical considerations when we are making recommendations in e-learning systems. We believe that the value added to the paper, in term of new knowledge that the learner learned from the paper, depends on the richness of information and the learner willingness to digest it. Moreover, this willingness depends on learner interest and motivation, which is reflected as learner goal.

## COMPONENTS AND TECHNIQUES

### Paper Repository

Electronic versions of all papers, including magazine articles, conference papers, workshop papers, etc., will be stored in the Paper Repository. If the electronic version is not available, then the hyperlink pointed to the paper or relevant publication information will be stored in the system. In addition, the repository also stores ratings and comments by users. Generally, the repository is a database with PaperID# as the primary key.

Four major tables are used in the database: MainTable, ActivePaperList, CandidatePaperList and UserRating. MainTable contains all completely verified paper information but comments by users. ActivePaperList contains only information of active papers which are ready to be recommended to users. CandidatePaperList contains a list of "new" papers which have not completely verified. A completely verified paper will be deleted from CandidatePaperList and added into MainTable. UserRating contains all user ratings and other comments in text format. Generally, ActivePaperList is a part of MainTable. The reason to have it separately is to facilitate the recommendation module, because a higher cost of processing a larger data and a frequent access of the ActivePaperList by the recommendation module. The UserRating is used to monitor user progress and to refine paper information.

For each paper, there are three kinds of tags to describe it: content tag, technical tag, and usage tag. Content Tag is the paper's attributes available from most digital libraries. It includes PaperID#, Paper Title, Author(s), Publication Year, Publication Place (journal or conference

name), Publication Type (book chapter, journal, conference, etc.), Category Contents in terms of subject and keywords and paper length, etc. Technical Tag is additional information which describes the technical level of the paper content (for novice, medium or advanced learners) and can be used in content-based filtering (model-based recommendation). The granularity of technical level is determined by the instructor. For instance, in a Software Engineering course, we can use a coarse grain such as Statistics, Discrete Math, Internet Computing, OO-programming, etc. But for a course in Artificial Intelligence we may use finer grain such as Predicate Calculus, Probability Theory, Utility Theory, etc. Technical tags are usually added manually when the paper is newly added, or inferred and adjusted based on the feedback given by learners, which will be explained in the next section. In order to keep a complete record of paper usage, we also need a usage tag which includes: userID of users who rated the paper and ratings and the time the ratings are submitted.

## RECOMMENDATION MODULE

### 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 the recommendation goal, their interest, background knowledge, etc. Basically, for each prototypical user group, there are representative candidate papers associated with it. It is obvious that these representative papers are the centers of their respective clusters. In addition, since a learner might fall into more than one cluster, the clustering algorithm should allow overlapping clusters.

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 applying 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 (Tang, & McCalla, 2003b).

### Focused Collaborative Filtering Module

After clustering is performed, learners are categorized in clusters based on their learning goal, interests, etc. 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., Herlocker et al., 1999; Schein et al., 2002; Melville, Mooney, & Nagarajan, 2002), but on the clustered areas (Tang, & McCalla, 2003b).

[FIGURE 3 OMITTED]

## 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.

### Topic-Driven Web Crawler

There is a web crawler embedded in this module, which is responsible for crawling digital libraries. To date, there has been a huge amount of research concerning web crawlers. Similar to the crawlers in other literatures, our web crawler is a topic-driven web crawler, which exploits the content-similarity between course topics and candidate papers. In the simplest form, a database is used to store the links to available e-journals. The database also stores the template (rules) for retrieving the information inside the journal.

### 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.

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.

[FIGURE 4 OMITTED]

#### 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 (2000), patterns of user behavior might be needed to perform garbage collection, in our system we will only focus on the usage of papers 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 (Pitkow, & Pirolli, 1997), 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.

There are several criteria in determining whether a paper should be deleted or not. For example, overall frequency within the specified category most recent frequency (Debevc, Meyer, & Svecko, 1997), overall cross-category frequency (since one paper might fall into more than two topical categories, its overall cross-category frequency measures its accumulative usage across these categories), average rating, and minimum acceptable rating. The first three factors concerning the usage of a paper mainly measure the frequency with which a target paper is recommended and read; the last two measure users' ratings of the target paper. If a paper consistently receives low ratings over a pre-defined period of time, it will be deleted.

#### MODEL AND FEASIBILITY STUDY

In this section we will provide formal concepts used in our system, followed by a simulation study in order to check the effectiveness of collaborative filtering in the context of learning.

In order to show that learner interest is not the primary factor in recommendation, we conducted a small-scale survey by asking our colleagues and friends (most of them have completed or are currently in a graduate program). Survey results substantiate our previous claims that uninteresting, yet pedagogically valuable papers should be recommended. These pedagogically useful, yet uninteresting papers (items are not false positives) Sarwar et al., (2000), because they could be helpful in one way or another to fulfill learners' learning expectations. Details of the survey can be found at (Tang, & McCalla, 2004a).

#### Pedagogically-Oriented Paper Recommendation

Our goal can be stated as follows. Given a collection of papers and a learner's profile, recommend and deliver a set of pedagogically suitable materials in an appropriate sequence, so as to meet both the learner's pedagogical needs and interests. Ideally, the system will maximize a learner's utility such that the learner gains a maximum amount of knowledge and is well motivated in the end. However, the content-based recommendation, which is achieved through a careful assessment of learner characteristics and then matches these models against papers, is very costly due to the following reasons:

\* When a new paper is added into the system, a detailed identification is required (e.g., tagging it with detailed information of the background knowledge needed for understanding it), which cannot be done automatically;

\* When a learner gains some new knowledge after reading a paper, a new matching process is required in order to find the next suitable paper for him/her, resulting in the updating of his/her learner model;

\* The matching between learner model and paper model may not be a one-to-one mapping, which increases the complexity of the computation.

Alternatively, we can use collaborative filtering (CF) to reduce the complexity of the recommendation process. The idea of CF is to let peer learners filter out those not suitable materials, while the system does not need to know the detailed characteristics of them. Hence, the matching process is not performed from learner models to learning materials, but from one learner model to other learner models. Since the system also utilizes some characteristics of papers and considers both learner knowledge and interest, then it is not a pure CF but a hybrid-CF. The remaining question is whether or not the hybrid-CF is as effective as the content-based recommendation. To answer this question, we carried out an experiment using artificial learners for two types of pedagogical-oriented recommendation techniques: pure content-based, which makes recommendations based on the matching of learner models to papers, and hybrid CF, which is based on peer learner recommendation.

#### A Formal Notation of Paper Recommendations

In a formal notion, we can state our recommendation steps as follow:

1. For a learner model  $l$ , find a group of similar learners,  $N(l)$ . (cluster of learners)
2. Given content  $C$ , find a group of relevant papers  $R(C)$ . (cluster of papers)
3. Find a subset of learners  $N' \subseteq N(l)$ , who have read/rated any paper in  $R(C)$ ; denoted by  $f: N(l) \rightarrow R(C) \times N'$ . (refine the cluster of learners)
4. Based on the attributes of  $R(C)$ , use content-based filtering to find a set of recommended papers  $R' \subseteq R(C)$  such that they match learner model  $l$ .
5. Or, based on the ratings given by  $N'$ , use collaborative filtering to find a set of recommended papers  $R' \subseteq R(C)$ .

The first step can be achieved without complex computation--for example, grouping by learners who have taken the same courses together, or grouping according to their learning profile (e.g., average grade of previous courses). We can also ignore this step if not many learners are recorded in the database. The second step can be realized by checking the subject and keywords of the papers. The third step is easily checked from the database. Finally, the fourth and fifth steps need more attention which will be described using the following basic definitions.

**Definition 1.** A paper in the domain being learned, denoted by  $r$ , is called commonly well selected if it is pedagogically suitable for all learners  $L$  under common learning constraints (time, prior knowledge, availability, etc.). The same definition applies for a set of all papers, denoted by  $[R.sup.c]$ , (i.e., it is commonly well selected if all papers  $r \in [R.sup.c]$  is commonly well selected).

**Definition 2.** A paper in the domain being learned is individually well selected if it is pedagogically suitable for a specific learner  $j \in L$ , under his/her individual learning constraints (common learning constraints plus individual learner characteristics, such as learning style, prior knowledge, preference, etc.). The same definition applies for all individually well selected papers, denoted by  $[R.sup.1]$ , (i.e., it is individually well selected if all papers  $r \in [R.sup.1]$  is individually well selected).

**Definition 3.** The set of all individually well selected papers is called the aggregate well selected paper, denoted by  $R$ .

Thus, we get  $[R.sup.c] = \bigcap_{j \in L} [R.sup.j.sup.1]$  and  $R = \bigcup_{j \in L} [R.sup.j.sup.1]$ . Additional paper beyond  $R$  is unnecessary. However, deciding  $[R.sup.1]$  is a non-trivial task, because in an ideal case, the tutor needs to decide proper pedagogical criteria in recommending the paper. In our proposed system, we left the learners and garbage collector to decide the set of  $R$ .

**Definition 4.** Similarity of two papers  $[r.sub.1]$  and  $[r.sub.2] \in R$ .

v-similarity (version-based):  $[r.sub.1]$  and  $[r.sub.2]$  share the same topic, might be written by



same authors, but one is a refined/updated version of another.

c-similarity (comparison-based): [r.sub.1] and [r.sub.2] discuss the same topic, with different approach.

t-similarity (technique-based): [r.sub.1] and [r.sub.2] use the same technique to solve two different problems.

s-similarity (simplicity-based): [r.sub.1] and [r.sub.2] concern the same topic and have the same level of simplicity in order to be understood.

Ideally, in content-based recommendation, we want to include papers with c-similarity but exclude v-similarity. According to learner interest and background knowledge, the system will recommend similar papers based of s-similarity. At current stage, we did not consider t-similarity. In CF, we do not consider those similarities which make our system less complicated.

Definition 5. Ordering of a set of papers [R.sup.S], where [R.sup.S] [??] R and |[R.sup.S]| > 1.

t-order: sequence of [R.sup.S] according to their technical difficulty.

l-order: sequence of [R.sup.S] according to their length.

p-order: sequence of [R.sup.S] according to the abstraction of their presentation.

r-order: sequence of [R.sup.S] according to the prestige of their publications.

c-order: sequence of [R.sup.S] according to the chronology of their publications.

In general, in the context of both content-based and collaborative filtering, we only consider c-order and l-order. The paper maintenance module will filter out papers from less prestigious publications by crawling only prestigious publishers and adding new papers only; thus it is not considered in the recommender module. The abstraction of paper presentation is not considered as well. The ordering in CF is realized after the system gets all the higher rated papers R' and is ready to deliver them to the learner. The straightforward way is by using t-order, from the easiest one, or combining it with l-order, with the shorter one first.

The "Conflict of Understanding and Interest" Problem

For paper p, and user U, we might have the learner sequence [U.sub.i.sup.t], where t is the time when the user accessed/read the paper. Therefore, when a user reads a paper at different time, he/she might have different ratings toward it, in that his/her understanding towards the paper might change (either for the better or for the worse) due to his/her own increasing background knowledge on the subject. This will also lead to a so-called "conflict of understanding and interest" problem where a user might provide largely different ratings towards a paper. But from both learners' and tutors' perspectives, this phenomenon is natural given the increasing pedagogical ability of learners as time goes by. Therefore, we will not make effort to "solve" this conflict; instead, these traces of living conflicts will be explored later to make a deep understanding of both the usage of a paper, and the learning curve of a learner.

It is obvious that as we can cluster users purely based on their browsing behaviors, we can also cluster the annotated user models with respect to a specific paper, or sequences of papers. Technically, the sequences of user models along with the collections of paper will provide rich information related to both users, user patterns, papers and paper usage patterns, which, in turn, can make more refined recommendations, provide both personalized and groupalized recommendations and form dynamic and collaborative groups based on clusters of learners with different interests, pedagogical backgrounds (Tang, & Chan, 2002).

## EVALUATING PEDAGOGY-ORIENTED HYBRID COLLABORATIVE FILTERING

As stated previously, we argue that two pedagogical features may be important in the recommendation system (RS): interest and knowledge. Moreover, content-based filtering may not as convenient as CF. In this section we will describe our experiment with both recommendation techniques using both pedagogical features.

## Simulation Setup

For the purpose of testing, we first generate 500 artificial learners and use 50 papers related to data mining as the main learning materials. The RS then delivers recommendations of 15 papers to each learner according to each individual learner model (pure content-based). Each artificial learner rates these papers according to their properties. After that, we generate 100 additional artificial learners, who become the target learners. Then, two recommendation techniques are applied for these target learners in order to evaluate their differences as well as performances. The first technique is the same as the technique used in the first 500 learners, (i.e., content-based recommendation). The second technique uses a hybrid-recommendation technique (model based with collaborative filtering).

## Learner Properties

In the simulation, we use minimal learner properties to generate artificial learners, as shown below:

\* Learner ID #.

\* Background knowledge as vector  $[(k_{sub.1}, k_{sub.2}) (k_{sub.3}, k_{sub.4}) (k_{sub.5}, k_{sub.6}) k_{sub.7}, k_{sub.8}, k_{sub.9}, k_{sub.10}]$ , where  $k_{sub.i}$  represents its strength on knowledge  $i$ -th, and  $k_{sub.i}$  [member of]  $[0, 1]$ . We assume that  $k_{sub.1}$  and  $k_{sub.2}$  are two basic mathematics topics,  $k_{sub.3}$  and  $k_{sub.4}$  are two discrete mathematics topics taught in computer science or mathematics,  $k_{sub.5}$  and  $k_{sub.6}$  are two statistics topics,  $k_{sub.7}$  is algorithm analysis, and  $k_{sub.8}$ ,  $k_{sub.9}$  and  $k_{sub.10}$  are topics in database, bioinformatics, and AI in education.  $k_{sub.1}$  is derived from truncated inverse standard lognormal distribution with  $[\sigma] = 1$  and reduced by factor  $1/5$ . And  $k_{sub.2}$  is lower than  $k_{sub.1}$  by the factor which also follows truncated standard lognormal distribution with  $s = 1$  and reduced by factor  $1/10$ .  $k_3, k_4$  and  $k_5$  are derived from truncated inverse lognormal distribution with  $[\sigma] = 1$  and reduced by factor  $1/5$ .  $k_{sub.6}$  is derived from truncated standard lognormal distribution with  $[\sigma] = 1$  and reduced by factor  $1/5$ .  $k_{sub.7}$  is derived from uniform distribution  $U[0, 1]$ .  $k_{sub.8}$ ,  $k_{sub.9}$  and  $k_{sub.10}$  are derived from truncated standard lognormal distribution with  $[\sigma] = 1$  and reduced by factor  $1/5$ .

\* Interest toward specific topics as vector  $[l_{sub.1}, l_{sub.2}, l_{sub.3}, \dots, l_{sub.12}]$ , where  $l_{sub.i}$  represents its interest on topic  $i$ -th, and  $l_{sub.i}$  [member of]  $[0, 1]$ . We assume that all interests except  $l_{sub.1}$  (general topical knowledge) are generated randomly following uniform distribution. And  $l_{sub.1}$  is generated using truncated inverse standard lognormal distribution with  $[\sigma] = 1$  and reduced by factor  $1/5$ .

\* Motivation as value  $M$  [member of]  $[0, 1]$ . Where 1 represents that the learner's willingness to spend more time to learn something not covered/understood before and 0 represents the learner's unwillingness to do so.  $M$  is generated using truncated standard lognormal distribution with  $[\sigma] = 1$  and reduced by factor  $1/5$ .

## Paper Properties

We use the following properties for the papers:

\* Paper ID #.

\* Technical knowledge as vector  $[(k_{sub.1}, k_{sub.2}) (k_{sub.3}, k_{sub.4}) (k_{sub.5}, k_{sub.6}) k_{sub.7}, k_{sub.8}, k_{sub.9}, k_{sub.10}]$ , where  $k_{sub.i}$  denotes the extensiveness of the knowledge  $i$ -th used inside the paper. The extensiveness of a knowledge means that a learner needs a good background of the corresponding knowledge in order to be able to understand the paper thoroughly. If the learner lacks that knowledge, then he/she can gain the corresponding knowledge by reading the paper carefully and spending more time to look at the references. We assume  $k_{sub.i}$  [member of]  $[0, 1]$ , and each of them represents the same topic as that described in learner properties. This feature indirectly affects the technical level of the paper.

\* Paper topics as vector  $[l_{sub.1}, l_{sub.2}, l_{sub.3}, \dots, l_{sub.10}]$ , where  $l_{sub.i}$  denotes the corresponding topic in the learner's interest.

\* Authority level, which is used to determine whether the paper is an important paper or not, for

example, a classical paper or highly cited paper, etc.

### Learning Constraints

Five core papers, (i.e., papers that are pedagogically required for all learners; for example, the seminal paper by Agrawal in 1994 introduced, for the first time, the notion of association rule mining and its technique), will be recommended by the system regardless of learners' interests or knowledge. Those papers are core papers in the simulated course. They are specifically chosen as follows: either paper ID #1 or #2, either paper #5 or #6, paper #8, one of paper #26 or #27 or #35, and either paper #33 or #48. Moreover, at least two papers with high technical level should be recommended to the learner. In total, 15 papers must be read and rated by each learner.

The above requirements define the constraints of recommendation, which differentiates the recommendation in an e-learning system from that in other application areas.

## MODEL-BASED RECOMMENDATIONS

### Rule Generations

The content-based recommendation is based on the following rules (learner-centric):

\* System starts with recommending acceptable papers in terms of learners' knowledge level (understandable) and the similarity of learners' interest towards the topic in the paper (the understandable level and the interest similarity will be described later). Up to eight authoritative papers will be selected first; if no more authoritative papers can be selected, then the system will recommend non-authoritative papers.

\* Two interested papers, but with very high technical level will be recommended in order to improve the learner's knowledge.

\* Some not-interested-yet-pedagogical-useful (authoritative) papers will be provided as the part of the learning requirement in the end.

After learners finish a paper, some additional knowledge may be acquired, which depends on the learner motivation. In our simulation, we assume that the increment is based on:

IF paper.[k.sub.j] > learner.[k.sub.j] AND paper.authority = TRUE THEN

learner.[k.sub.j] = (paper.[k.sub.j] - learner.[k.sub.j]) X learner.M X Interest X [w.sub.1] + learner.[k.sub.j]

IF paper.[k.sub.j] > learner.[k.sub.j] AND paper.authority = FALSE THEN

learner.[k.sub.j] = (paper.[k.sub.j] - learner.[k.sub.j]) X learner.M X Interest X [w.sub.2] + learner.[k.sub.j]

Where [w.sub.1] and [w.sub.2] represent factors that might affect learning speed after reading an authoritative/non-authoritative paper. They are two of the control variables in our experiment. Interest represents the similarity of a learner's interest to the paper's topic which will be described later. Moreover, the total gain made by the learner is defined as the value added from reading the paper, or

Value added = [[SIGMA].sub.i] ((new) learner.[k.sub.i] - (old) learner.[k.sub.i])

The rule to measure a learner's understanding (Understand) will be based on the knowledge gap between the learner's knowledge and those required to fully understand the paper. In addition, the similarity of learners' interest to the paper's topic (Interest) is generated according to the following rules:

y = 1 if [there exists]j such that learner.[l.sub.j] and paper.[l.sub.j] [greater than or equal to] 0.9

y = 0.9 if [there exists]j such that learner.[l.sub.j] and paper.[l.sub.j] [greater than or equal to] 0.8

...

$y = 0.1$  if [there exists]  $j$  such that learner.[l.sub.j] and paper.[l.sub.j] [greater than or equal to] 0.0

Interest = Max(y)

#### Rating Generation Rules

After a learner reads a paper, we need rating-generation rules to generate learner rating toward the paper. We use the following rules in our simulation:

1. If learners are interested in the topic AND understand the paper, a higher rating is generated. Or, matching based on both interests and background knowledge generates higher ratings 4 or 5 under the following formula.

Rate = Round (Interest X Understand X 2) + 3 if Interest [greater than or equal to] 0.7 and Understand [greater than or equal to] 0.7

2. Learners give ratings to a paper based on the amount of knowledge that could be acquired (value added) AND the understanding of the paper (easy to follow, or pedagogical-readiness), OR the importance of the paper to their interest.

#### Rating Implications

If the rating is high (e.g., 4 or 5), learner motivation will increase randomly following uniform distribution to upper bound value 1, with increasing rate  $x$ . If the rating is low (e.g., 1 or 2), learner motivation will decrease randomly also following uniform distribution to lower bound 0, also with decreasing rate  $x$ . If the rating falls into the medium of the scale (e.g. 3), learner motivation unchanged.  $x$  is another control variable which represents how much motivation a learner gain/loss after reading a paper.

#### Hybrid Recommendations

The following rules are used for the hybrid recommendation.

\* Neighborhood finding. For each target learner (tlearner) find five neighbors (nlearner) based on the similarity of their interest and background knowledge. The similarity measurement is calculated based on the following:

Positive Similarity = [SIGMA] (nlearner.[l.sub.i]) IF tlearner.[l.sub.i] [greater than or equal to] nlearner.[l.sub.i]

Negative Similarity = [SIGMA] (nlearner.[l.sub.i] - tlearner.[l.sub.i]) IF tlearner.[l.sub.i] < nlearner.[l.sub.i]

Similarity = Positive Similarity - Negative Similarity

The similarity formula is used to find similarity in learner background knowledge. The rationale to adopt this measurement is when a learner has a lower interest than a target learner. The magnitude of the learner's interest is credible for recommending a paper to the target learner, therefore the positive similarity measures the total of learners' interest in such condition. However, if the learner's interest is higher than the target learner's interest, then an error may appear regarding the learner's recommendation, and the gap between those two interests may be the cause of the error. Therefore, the negative similarity denotes the sum of the gaps. The same is used for the similarity measure of two learners' background knowledge.

\* From these five nearest neighbors, we can get a set of candidate papers based on their ratings. In our simulation, each learner has rated 15 papers; therefore, at least 15 papers will be in the candidate set. Then, we order those papers in candidate set from the highest ratings from the closest neighbor to the lowest rating from the furthest neighbor.

\* The system will recommend up to eight authoritative papers starting from those receiving the

highest rating followed by recommending nonauthoritative papers. Then, the system will choose and recommend two very interesting and highly technical papers, and recommend five pedagogically required papers, if the learner has not read them. Finally, the system recommends the rest of the papers according to the rating order, up to 15 papers in total.

#### Evaluation Metrics and Control Variables

Those commonly adopted metrics (Herlocker et al., 1999), (e.g., ROC) in the research community cannot be applied here due to the inherent features of recommendation for e-learning. These metrics, for example, ROC, are mainly adopted to test the users' satisfaction in terms of item interest. However, we argue that since the most critical feature of recommending learning items is to facilitate learning (not just to provide interesting items), it is not applicable in our domain. Therefore, we propose two new metrics as follows:

\* Average learner motivation after recommendation

\* Average learner knowledge after recommendation

And for the purpose of comparison, we compare the percentage differences between content-based recommendation and hybrid recommendation.

In our simulation, control variables [w.sub.1], [w.sub.2] and  $x$  are adjusted to differentiate artificial learners as follows:  $x = 1$  for fast motivation change (FMC),  $x = 0.3$  for moderate (MMC),  $x = 0.1$  for slow (SMC), and  $x = 0$  (NMC). Moreover, we use eight pairs of ([w.sub.1], [w.sub.2]), which are (1, 0), (U[0, 1], 0), (U[0, 0.3], 0), (1, U[0, 0.3]), (U[0, 1], U[0, 0.3]), (U[0, 0.3], U[0, 0.3]), (1, U[0, 1]), (U[0, 1], U[0, 1]), (1, 1), where U[0, y] means a random value generated from a uniform distribution function. The pair value represents the effect of authoritative and non-authoritative papers in the increment of the learner's knowledge. For example, (1, 0) indicates that only authoritative papers can fully increase a learner's knowledge. And (1, 1) indicates that both authoritative and non-authoritative papers are equally weighted and can fully increase a learner's knowledge. Each group of experiments is repeated thirty times for statistical analysis.

## EXPERIMENT RESULTS AND DISCUSSION

### Experiment Results

Table 3 shows the results of the experimentation. The value shown in each cell is the pair value of the percentage difference between content-based recommendation and the hybrid-CF technique in terms of average learner knowledge and motivation. A negative value indicates that the content-based recommendation technique is better than hybrid-CF. And a positive value represents the reverse situation. For example, the pair value (0.65; 2.93) represents that using hybrid-CF is 0.65% and 2.93% better than using content-based in terms of the average learner knowledge and motivation respectively. All results are checked by t-test for equal mean hypothesis (assuming different variance). The value in italics inside the table shows that the null hypothesis is not rejected (for  $[\alpha] = 0.05$ ), or the difference between content-based and hybrid-CF is not statistically significant. If we exclude zero and italic values in Table 3, then there are 14 and 6 negative values for the difference of learner knowledge and motivation respectively, with the lowest values equal to -1.05% and -5.68% respectively. And there are 8 and 12 positive values for the difference of learner knowledge and motivation, with the highest values equal to 1.20% and 19.38%, respectively. Thus, we conclude that using hybrid-CF results in a lower performance in terms of learner average knowledge. However, since hybrid-CF usually needs lower computational cost than content-based recommendation (which is not measured here) and the performance loss is not big, hence hybrid-CF is very promising in e-learning system.

So far, it is unclear why the individual result of our simulations, especially some values which show high differences, especially when motivation changes quickly (FMC). However, we can conclude that using hybrid-CF may not always result in a lower performance. And if it happens, the difference may not higher than 5%. This conclusion is useful, since hybrid-CF needs lower cost than content-based recommendation. Thus, if the performance lost is not big, then hybrid-CF should be used instead of the traditional content-based recommendation.

### Discussion

Computer simulation has long served as a tool of applying artificial intelligence on intelligent tutoring systems (Chan, & Baskin, 1990; Tang, & Chan, 2002). Although a simulation program can only model part of the real environment where real learners involve, it can afford a powerful tool for gaining insights for paper recommendations in complex settings. Therefore, the simulation discussed here can serve as a guide in our future study. In fact, we have designed a follow-up human subject study, which extends the resulting recommendations made by the artificial learners on the real human learners to solve the cold-start problem in our domain successfully (Tang, & McCalla, 2004b).

## LESSONS LEARNED

One interesting approach associated with the recommendation module is that papers actually can be annotated by users themselves (so their user models), which is fundamentally different from current paper tagging techniques. Indeed, it is obvious that in such evolving e-learning systems, when more and more users interact with the system, each paper annotations will be greatly refined and automated, thus instead inform as well as improve the quality of future recommendations.

Another interesting issue concerns our experiments on both artificial learners and human subjects. More specifically, we conducted a human-subject study where each human user received papers based on the ratings given by the artificial learners (Tang, & McCalla, 2004c) and the rating generation mechanism follows what we describe in this paper. Experiment results showed that the majority of learners have struggled to reach a harmony between their interest and educational goal--they are willing to read noninteresting-yet-pedagogically-useful papers in order to acquire new knowledge for either their group project or their long-term goal. Hence, from this perspective, learners seem to be more tolerant than users in commercial recommender systems. Nevertheless, as educators, we should still maintain a balance of recommending interesting papers and pedagogically helpful ones in order to retain learners and continuously engage them throughout the learning process. This is especially true in our case, since most of the students in Hong Kong are more application-oriented.

## CONCLUSION

In this article, we proposed an evolving e-learning system which can adapt itself both to the learners and the open the Web and pointed out the differences of making recommendations in e-learning and other domains. We propose two pedagogy features in recommendation: learner interest and background knowledge. A description of paper value, similarity, and ordering are presented using formal definitions. We also study two pedagogy-oriented recommendation techniques: content-based and hybrid recommendations. We argue that while it is feasible to apply both of these techniques in our domain, a hybrid collaborative filtering technique is more efficient to make "just-in-time" recommendations. In order to assess and compare these two techniques, we carried out an experiment using artificial learners. Experiment results are encouraging, showing that hybrid collaborative filtering, which can lower the computational costs, will not compromise the overall performance of the RS. In addition, as more and more learners participate in the learning process, both learner and paper models can better be enhanced and updated, which is especially desirable for web-based learning systems. We have tested the recommendation mechanisms with real learners, and the results are very encouraging. Details can be found at (Tang, & McCalla, 2004b, c).

Table 1 A comparison of learner model A, B and C

	Learner A	Learner B	Learner C
Knowledge in Statistics	Strong	Weak	Weak
Knowledge in Marketing and Management science	Strong	Weak	Weak
Knowledge in network security (e.g. SSL)	Strong	Weak	Weak
Interest	Network security,	Network security,	Data mining & web-mining

	social network	social network	application in e-commerce
Paper Preferences	Technical/theoretical	Application and magazine survey, technical/theoretical	Application and magazine survey

Table 3 The differences between content-based and hybrid recommendation (in percentage %). The first value in each cell represents the difference of final knowledge and the second value represents the difference of final motivation.

([w.sub.1], [w.sub.2])	FMC	MMC	SMC
(1, 0)	0.59; 2.77	-0.70; -0.06	-0.77; -0.42
(U[0, 1], 0)	0.98; 7.97	-0.28; 3.85	0.21; -0.32
(U[0, .3], 0)	-0.47; 15.15	-0.52; 0.75	0.33; -5.42
(1, U[0, .3])	-0.57; 1.61	-1.05; -1.05	-0.76; -0.90
(U[0, 1], U[0, .3])	0.30; 8.09	-0.44; 3.41	0.22; -0.01
(U[0, .3], U[0, .3])	-0.85; 19.38	-0.69; -0.19	0.06; -5.68
(1, U[0, 1])	-0.52; 1.13	-0.96; -0.8	-0.82; -0.84
(U[0, 1], U[0, 1])	0.96; 7.36	-0.15; 4.68	0.16; -0.06
(1, 1)	-0.34; 1.47	-0.69; -1.31	-0.47; -0.81
([w.sub.1], [w.sub.2])	NMC		
(1, 0)	-0.43; 0.00		
(U[0, 1], 0)	0.54; 0.00		
(U[0, .3], 0)	1.09; 0.00		
(1, U[0, .3])	-0.29; 0.00		
(U[0, 1], U[0, .3])	0.69; 0.00		
(U[0, .3], U[0, .3])	1.20; 0.00		
(1, U[0, 1])	-0.27; 0.00		
(U[0, 1], U[0, 1])	0.88; 0.00		
(1, 1)	-0.43; 0.00		

## References

- Balabanovic, M., & Shoham, Y. (1997) Fab: Content-based collaborative recommendation. *Communications of the ACM*, 40(3): 66-72.
- Basu, C., Hirsh, H., Cohen, W., & Nevill-Manning, C. (2001). Technical paper recommendations: A study in combining multiple information sources. *JAIR*, 1, 231-252.
- Billsus, D., & Pazzani, M. (1999). A hybrid user model for news story classification. *Proceedings of UM'99*, 99-108.
- Bollacker, K., Lawrence, S., & C. Lee Giles, C. L. (1999). A system for automatic personalized tracking of scientific literature on the web. *ACM DL*, 105-113.
- Boyle, C., & Encarnacion, A. O. (1994). MetaDoc: An adaptive hypertext reading system. *UMUAI*, 4, 1-19.
- Brusilovsky, P. (2001). Adaptive hypermedia. *UMUAI*, 11(1/2), 87-110.
- Chan, T., & Baskin, A.B. (1990). Learning companion systems. *ITS 1990*, 6-33.
- Debevc, M., Meyer, B., & Svecko, R. (1997). An adaptive short list for documents on the world wide web. *IUI 1997*, 209-211. U.S.A.
- De Bra, P., & Calvi, L. (1998). AHA! An open adaptive hypermedia architecture. *The New*

Review of Hypermedia and Multimedia, 4, 115-139.

Herlocker, J., Konstan, J., Borchers, A., & Riedl, J. (1999). An algorithmic framework for performing collaborative filtering. SIGIR'99, 230-237.

Jameson, A., Konstan, J., & Riedl, J. (2002). AI techniques for personalized recommendation. Tutorial notes, AAAI-02, Edmonton, Canada.

Joachims, T., Freitag, D., & Mitchell, T. (1997). WebWatcher: A tour guide for the World Wide Web. Proceedings of IJCAI'97, 770-775.

Kaplan, C., Fenwick, J., & Chen, J. (1993). Adaptive hypertext navigation based on user goals and context. UMUAI, 3(3), 193-220.

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

McCalla, G. (2000). The fragmentation of culture, learning, teaching and technology: Implications for the artificial Intelligence in education research agenda in 2010. IJAIED. 11(2): 177-196. 2000.

McNee, S., Albert, I., Cosley, D., Gopalkrishnan, P., Lam, S., Rashid, A., Konstan, J., & Riedl, J. (2002). On the recommending of citations for research papers. ACM CSCW'02, 116-125.

Melville, P., Mooney, R., & Nagarajan, R. (2002). Content-boosted collaborative filtering for improved recommendations. AAAI/IAAI 2002, 187-192. Edmonton, Canada.

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

Pazzani, M., Muramatsu, J., & Billsus, D. (1996). Syskill and Webert: Identifying interesting Web sites. AAAI'96, 54-61.

Pitkow, J., & Pirolli, P. (1997). Life, death, and lawfulness on the electronic frontier. ACM CHI 1997, 383-390.

Recker, M., Walker, A., & Lawless K. (2003). What do you recommend? Implementation and analyses of collaborative information filtering of web resources for education. Instructional Science, 31:299-316.

Resnick, P., Iacouvou, N., Suchak, N., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of Netnews. ACM CSCW'94, 175-186.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Analysis of recommendation algorithms for e-commerce. ACM EC'2000, 158-167. Minneapolis.

Schafer, J., Konstan, J., & Riedl, J. (2001). Electronic commerce recommender applications. Data Mining and Knowledge Discovery, 5, (1/2, 2001), 115-152.

Shardanand, U., & Maes, P. (1995). Social information filtering: Algorithms for automating 'word of mouth'. ACM CHI'1995, 210-217 Denver.

Schein, A., Popescul, A., Ungar, L.H., & Pennock, D. (2002). Proceedings of SIGIR'02. 253-260.

Stern, M. K., & Woolf, B.P. (2000). Adaptive content in an online lecture system. AH, 227-238

Tang, T.Y., & Chan, K.C.C. (2002). Feature construction for student group forming based on their browsing behaviors in an e-learning system. PRICAI 2002, LNCS 2417, 512-521.

Tang, T. Y., & McCalla, G. (2003a). Smart recommendations for an evolving e-learning system. Workshop on Technologies for Electronic Documents for Supporting Learning, AIED'2003.



Tang, T. Y. and McCalla, G. (2003b) Mining implicit ratings for focused collaborative filtering for paper recommendations. UM 2003, Workshop on User and Group Models for Web-based Adaptive Collaborative Environments, Johnstown, U.S.A.

Tang, T. Y. & McCalla, G. (2004a). On the pedagogically guided paper recommendation for an evolving web-based learning system. FLAIRS Conference, 2004. AAAI Press. 86-91.

Tang, T. Y. & McCalla, G. (2004b). Utilizing artificial learners to help overcome the cold-start problem in a pedagogically-oriented paper recommendation system. In Proceedings of AH 2004: International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems. August 23-26, Eindhoven, Netherlands, 245-254.

Tang, T. Y. & McCalla, G (2004c) Laws of Attraction: In Search of Document Value-ness for Recommendation. In Proceedings of ECDL 2004: the 8th European Conference on Digital Library, Sept. 12-17 2004, Bath, UK. 269-280.

Weber, G., & Brusilovsky, P. (2001). ELM-ART: An adaptive versatile system for web-based instruction. International Journal of AI in Education, 12: 1-35.

Woodruff, A., Gossweiler, R., Pitkow, J., Chi, E., & Card, S. (2000). Enhancing a digital book with a reading recommender. ACM CHI 2000, 153-160.

TIFFANY YA TANG

Department of Computing, Hong Kong Polytechnic University, Hong Kong

Department of Computer Science, University of Saskatchewan, Canada

cstiffany@comp.polyu.edu.hk

GORDON McCALLA

Department of Computer Science, University of Saskatchewan, Canada

yat751@mail.usask.ca; mccalla@cs.usask.ca

-----  
COPYRIGHT 2005 Association for the Advancement of Computing in Education (AACE)  
No portion of this article can be reproduced without the express written permission from the copyright holder.  
Copyright 2005, Gale Group. All rights reserved. Gale Group is a Thomson Corporation Company.  
-----

**Please bookmark with social media, your votes are noticed and appreciated:**



[Free e-learning guide](#)

See e-learning techniques, industry standards, terms and concepts.

[www.knowledgепresenter.com](http://www.knowledgепresenter.com)

[Data Integration Study](#)

Pernod Ricard Pacific selects Sonic ESB for SOA integration platform.

[www.progress.com/innovators](http://www.progress.com/innovators)

Ads by Google

Reader ratings: ★★★★★ 0 [0 vote(s)] You can rate this article by selecting 1 to 5 stars on the left.

**Reader Opinion**

Title: \_\_\_\_\_

Comment: \_\_\_\_\_



[RacerPro for Semantic Web](#)

Fast and scalable reasoner on Complex ontology & OWL data  
[www.franz.com/products/racer](http://www.franz.com/products/racer)

[SCORM E-Learning](#)

LCMS / LMS - Complete solution for Aerospace and Defence  
[www.lbs-ltd.com](http://www.lbs-ltd.com)

[Abstract/paper management](#)

World Class Online system Submit Review Pub Plan Session  
[www.ShockLogic.com](http://www.ShockLogic.com)

[Altio](#)

Real-time data visualisation, interaction and integration.  
[www.altio.com](http://www.altio.com)

Ads by Google

[The Free Library > Business and Industry > Business International > International Journal on E-Learning > January 1, 2005](#)  
[The Free Library > Science and Technology > Electronics > International Journal on E-Learning > January 1, 2005](#)  
[The Free Library > Social Sciences > Education > International Journal on E-Learning > January 1, 2005](#)  
[The Free Library > Date > 2005 > January > 1 > International Journal on E-Learning](#)

Publications by Name	Publications by Date	Authors	Literature
<a href="#">A-D</a> <a href="#">E-Q</a> <a href="#">P-T</a> <a href="#">U-Z</a>	<a href="#">84</a> <a href="#">85</a> <a href="#">86</a> <a href="#">87</a> <a href="#">88</a> <a href="#">89</a> <a href="#">90</a> <a href="#">91</a> <a href="#">92</a> <a href="#">93</a> <a href="#">94</a> <a href="#">95</a>	<a href="#">A</a> <a href="#">B</a> <a href="#">C</a> <a href="#">D</a> <a href="#">E</a> <a href="#">F</a> <a href="#">G</a> <a href="#">H</a> <a href="#">I</a> <a href="#">J</a> <a href="#">K</a> <a href="#">L</a> <a href="#">M</a>	<a href="#">A</a> <a href="#">B</a> <a href="#">C</a> <a href="#">D</a> <a href="#">E</a> <a href="#">F</a> <a href="#">G</a> <a href="#">H</a> <a href="#">I</a> <a href="#">J</a> <a href="#">K</a> <a href="#">L</a> <a href="#">M</a>
	<a href="#">96</a> <a href="#">97</a> <a href="#">98</a> <a href="#">99</a> <a href="#">00</a> <a href="#">01</a> <a href="#">02</a> <a href="#">03</a> <a href="#">04</a> <a href="#">05</a> <a href="#">06</a> <a href="#">07</a>	<a href="#">N</a> <a href="#">O</a> <a href="#">P</a> <a href="#">Q</a> <a href="#">R</a> <a href="#">S</a> <a href="#">T</a> <a href="#">U</a> <a href="#">V</a> <a href="#">W</a> <a href="#">X</a> <a href="#">Y</a> <a href="#">Z</a>	<a href="#">N</a> <a href="#">O</a> <a href="#">P</a> <a href="#">Q</a> <a href="#">R</a> <a href="#">S</a> <a href="#">T</a> <a href="#">U</a> <a href="#">V</a> <a href="#">W</a> <a href="#">X</a> <a href="#">Y</a> <a href="#">Z</a>

[Terms of use](#) | Copyright © 2007 [Einfex, Inc.](#) | [Feedback](#) | [For webmasters](#) | [Submit articles](#)

