

Learning materials recommendation using good learners' ratings and content-based filtering

Khairil Imran Ghauth · Nor Aniza Abdullah

Published online: 9 March 2010

© Association for Educational Communications and Technology 2010

Abstract The enormity of the amount of learning materials in e-learning has led to the difficulty of locating suitable learning materials for a particular learning topic, creating the need for recommendation tools within a learning context. In this paper, we aim to address this need by proposing a novel e-learning recommender system framework that is based on two conceptual foundations—peer learning and social learning theories that encourage students to cooperate and learn among themselves. Our proposed framework works on the idea of recommending learning materials with a similar content and indicating the quality of learning materials based on good learners' ratings. A comprehensive set of experiments were conducted to measure the system accuracy and its impact on learner's performance. The obtained results show that the proposed e-learning recommender system has a significant improvement in the post-test of about 12.16% with the effect size of 0.6 and 13.11% with the effect size of 0.53 when compared to the e-learning with a content-based recommender system and the e-learning without a recommender system, respectively. Furthermore, the proposed recommender system performed better in terms of having a small rating deviation and a higher precision as compared to e-learning with a content-based recommender system.

Keywords Content-based filtering · E-learning · Recommender system · Peer learning

Introduction

The popularity of e-learning has created huge amounts of educational resources. Hence, locating the suitable learning references has become a big challenge. One way to address

K. I. Ghauth (✉)

Faculty of Information Technology, Multimedia University, 63100 Cyberjaya, Malaysia
e-mail: khairil-imran@mmu.edu.my

N. A. Abdullah

Faculty of Computer Science and Information Technology, University of Malaya,
50603 Kuala Lumpur, Malaysia

this challenge is the use of recommender systems. A recommender system is a tool that helps users to identify interesting items from a large pool of items. Furthermore, to recommend quality learning materials, it is needed to devise a new approach that is not solely based on the learning content itself, but one that takes into account the student's opinion as input for further classification of the material (Bobadilla et al. 2009). However, the focus of recent efforts in research has been more on the recommender systems' accuracy based on the learning content, neglecting the student's input. This work aims to fill that gap through incorporating learners' ratings in the content classification and recommendation process. In this regard, the success of the implementation of a variety of recommender systems in e-commerce, for recommending from a large number of items, has been inspiring for e-learning researchers and has been reviewed for this work (Soonthornphisaj et al. 2006; Liang et al. 2006; Tang and McCalla 2003).

Among popular approaches used in recommender systems are collaborative filtering, content-based filtering, and hybrid filtering. Collaborative filtering identifies the interesting items from similar users' opinions by calculating the nearest-neighbor from a rating matrix. New items that have not been rated by the user and are of interest to the nearest-neighbor will be recommended to the user. In contrast, content-based filtering uses features of items to infer recommendations. Whereby, items with similar content to the current viewing item will be recommended to the user (Felferning et al. 2007). On the other hand, hybrid filtering combines both content-based filtering and collaborative filtering techniques for producing a recommendation (Adomavicius and Tuzhilin 2005). As in other domains, recommender systems in e-learning can differ in many ways depending on what kind of object is to be recommended (e.g., courses to enroll, learning materials, and etc.) and whether context of learning is considered important (Soonthornphisaj et al. 2006; Liang et al. 2006; Kim and Baylor 2006).

While recommender systems have become a popular method of suggesting items, peer learning has emerged as an effective way of learning (Topping 2005). Topping (2005) defined peer learning as the acquisition of knowledge and skill through active helping and supporting among status equals or matched companions. It involves people from similar social groupings who are not professional teachers helping each other to learn, while learning themselves by so doing. Help and support among peers can be demonstrated in many ways such as teaching and/or sharing materials. Topping (2005) used the term "peer helper" for someone who is considered to be among the "best students" and who acts as a surrogate teacher, in a linear model of the transmission of knowledge, from a teacher to peer helpers to other learners. The idea of learning from the best students or good learners is also strongly supported by Social Learning Theory (Bandura 1977). Social Learning Theory (Bandura 1977) stated that people can learn by observing the behavior of others and the outcome of those behaviors. Furthermore, the theory also mentioned that other people will most likely exhibit the behavior if the outcome is positive. This theory strongly supports the idea of learning from good learners, whereby exhibiting good learners' behavior (i.e., focusing on highly rated items) can increase performance.

Hence, the incorporation of a peer-review mechanism (i.e., the good learners' rating) into a recommender system, seems promising of unprecedented effectiveness, and is what this work has investigated. In this paper, we propose an e-learning recommender system from the combination of two types of recommendation systems, namely: (i) content-based recommendation and (ii) recommendation based on good learners' rating. The objective of the first recommendation type is to recommend the additional learning resources that are similar to those of the viewing item. It ensures that the recommended items always remain within the learning context. The second recommendation type aims to guide learners in

selecting good learning resources in order to improve their understanding of the learning topic. The terms “good learners” and “items” used in this paper can be defined as follows. Good learners are the learners who have studied the learning materials, and completed the post-test evaluation and achieved marks above 80%. Items on the other hand are learning materials that can be divided into chapters or sub chapters and accompanied with the item attributes. Item attributes consist of author, title, and keywords. The term ‘item’ is used interchangeably with ‘document’ in this paper.

The remaining part of this paper is organized as follows. In ‘[Literature survey](#)’ section, the existing work on e-learning recommender systems is presented. ‘[System description](#)’ section introduces the overall system architecture and describes the proposed method which includes the recommendation framework and the mathematical model used for recommending the items. ‘[Methodology](#)’ section presents the analysis of the data and the results. This section explains the detailed information of the experiment setup for evaluating the recommendation mechanism. ‘[Implications](#)’ section discusses the limitations of the proposed system and the improvements needed to system components. Finally, ‘[Conclusion and future works](#)’ section provides the concluding remarks along with suggestions for future work.

Literature survey

As described earlier, our proposed e-learning recommender system makes recommendations based on content-based filtering and good learners’ ratings. The good learners’ ratings recommendation can be seen as a special type of collaboration among the learners even though the technique that we employed does not use the collaborative filtering algorithm. Content-based filtering recommends similar items based on the similarity of content while collaborative filtering recommends according to similar users’ preferences (Felfernig et al. 2007). Both techniques are suitable to recommend relevant learning materials to the user. There are several related works concerning content-based filtering and/or collaborative filtering techniques. Bobadilla et al. (2009) proposed a new equation for collaborative filtering that incorporated the learners’ score (from a test) into the calculations for item prediction. Their experiment revealed that the method achieved high item-prediction accuracy. Liang et al. (2006) applied the combination of content-based filtering and collaborative filtering to make personalized recommendations for a courseware selection module. The algorithm starts with user u entering some keywords on the portal of courseware management system. Next, the courseware recommendation module finds within the same user interest group of user u the k courseware with the same or similar keywords that others choose. A recommendation degree will be calculated for each k courseware by multiplying the degree of trust (similarity) between user u and other users and the evaluation of courseware by user u . Finally the top 5 recommended courseware is outputted according to the recommendation degree sequence. Their experiment showed that the algorithm used, was able to reflect users’ interests with high efficiency. Soonthornphisaj et al. (2006) used the collaborative filtering technique to predict the most suitable documents for the learner. The algorithm starts with the calculation of weight between all users and the active learner using Pearson correlation. Next, it selects the n users that have the highest similarity to the active learner in order to form the neighborhood. Finally, the rating prediction is calculated using the weight combination obtained from the neighborhood. They were also proposing a new e-learning framework using web

services that had the ability to aggregate recommended materials from other e-learning web sites and predicted more suitable materials for learners.

Some researchers have combined the collaborative filtering and/or content based filtering with data mining techniques. The data mining techniques gathered information about the user behavior, such as navigation history obtained from the log file, to produce recommendations. This technique is suitable to recommend the sequence of learning materials (i.e., learning path) rather than the learning material itself. Zaiane (2002) used web mining techniques to build agents that could recommend online learning activities or shortcuts in a course website which improved course navigation as well as assisted with the online learning process. Khribi et al. (2009) computed online automatic recommendations based on learners' recent navigation histories as well as exploiting similarities and dissimilarities among user preferences and among the contents of the learning resources. They used web usage mining techniques together with content-based and collaborative filtering to compute relevant links to recommend to active users. Liu and Shih (2007) designed a material recommendation system based on association rule mining and collaborative filtering. Since the user's preferences were predetermined (from the result of the web usage mining), the system was able to reduce the work load that was required to develop the system's search engine, as well as the complexity of the content parsing needed for improving the recommendation.

There are several other methods that have been used in e-learning recommender systems including clustering techniques, metadata, Item Repository Theory, and neural networks. Each of the techniques produced their own recommendation results but they required intensive computation. Tang and McCalla (2003) proposed an evolving web-based learning system that was able to find relevant content on the web, personalize and adapt the content based on the system observation of its learners and the accumulated ratings given by the learners without the learners having to directly interact with the open Web. They used a clustering technique to cluster the learners before using collaborative filtering, to calculate learners' similarities for content recommendation. Kerkiri et al. (2007) designed a framework that exploited both description and reputation metadata to recommend personalized learning resources. Their experiment proved that the use of reputation metadata augmented the learner's satisfaction by retrieving those learning materials that were evaluated positively. Chen et al. (2005) used Item Repository Theory which estimated the abilities of online learners and recommended appropriate course materials to learners. The obtained feedback from the learners showed that they felt their understanding of the content of the recommended learning materials was high (score of 0.825 from the scale of 0–1 where 0 indicates not understood and 1 indicates understood). Their proposed system has also successfully adjusted the level of difficulty of the learning materials to suit the user's capability (score of 1.815 from the scale of 1–4 where value 1 indicates very easy and value 4 indicates very hard). Tai et al. (2008) proposed an e-learning course recommendation based on artificial neural networks which were used to classify the learners based on groups of similar interests where learners were able to obtain course recommendations from the group's opinion. The experiment results showed that 80.7% of the learners felt that the course recommendation by the system was appropriate. Moreover, 75.8% of the learners thought that the personalized recommendation services provided by the system satisfied their requirements.

Table 1 summarizes the recommendation strategies proposed by the current researches on e-learning as well as their advantages and disadvantages. To date, none of the e-learning systems (as reported in Zaiane 2002; Khribi et al. 2009; Soonthornphisaj et al. 2006; Liu and Shih 2007; Liang et al. 2006; Tang and McCalla 2003; Kerkiri et al. 2007; Chen et al.

Table 1 Summary of the recommendation strategies, disadvantages, input, and output of the current researches

Researchers		Recommendation			
	Strategies	Advantages	Disadvantages	Input	Output
Zaiane (2002), Khribi et al. (2009), Liu and Shih (2007), and Tai et al. (2008)	Data mining techniques	Suitable to recommend the learning path (e.g., sequence of learning materials to read)	Unable to recommend independent items and/or classify the quality (usefulness) of items	Learner's activities/ access history, learners rating, item attributes, learners score	Related items/ documents, related links, learning activities, courseware module
Khribi et al. (2009), Soonthornphisaj et al. (2006), Liu and Shih (2007), Liang et al. (2006), Tang and McCalla (2003), Bobadilla et al. (2009), Tanet al. (2008), and Santos and Boticario (2008)	Collaborative filtering (User-oriented)	New items are able to be recommended	Items recommended may not be in the current learning context and recommendation rely heavily on user's ratings		
Khribi et al. (2009) and Liang et al. (2006)	Content-based filtering (Item-oriented)	Items are able to be recommended even though they have not received any ratings	Lacking in term of which items are good for recommendation and new items are not able to be recommended		
Liang et al. (2006), Tang and McCalla (2003), Kerkiri et al. (2007), Chen et al. (2005), Otair and Hamad (2005), Tai et al. (2008), Aijuan and Baoying (2008), and Zheng and Li (2008)	Clustering, Knowledge discovery, metadata, Item repository theory, Rule-based expert system, Artificial neural network, questionnaire	Able to solve complex computation (e.g., multi-criteria recommendation, clustering)	Involve high computation		

2005; Otair and Hamad 2005; Tai et al. 2008) have attempted to recommend the items based on the ratings of good learners as a recommendation method. This method helps to classify useful items based on the good learners' perception. Only the work done by Bobadilla et al. (2009) considered the use of learner score when calculating the weight for recommendation. Furthermore, none of the researches (reported in Zaiane 2002; Khribi et al. 2009; Soonthornphisaj et al. 2006; Liu and Shih 2007; Liang et al. 2006; Tang and McCalla 2003; Kerkiri et al. 2007; Chen et al. 2005; Otair and Hamad 2005; Tai et al. 2008; Bobadilla et al. 2009) have compared the learners' performances while testing the recommender system. The recommendation system proposed in the following section aims to address these issues.

System description

The system development of the proposed system can be divided into 2 phases: (i) development of the modeling which involves retrieving the document's keywords, calculation for the item's similarity, retrieving the good learners' ratings, and calculation for the good learners' average and prediction ratings, and (ii) development of the recommendation which involves the process of selecting the top-N recommended items that exceed the item's similarity threshold, and recommending good learners' rating for a particular learning material. The output of the recommendation phase is depicted in Fig. 1.

Note that, for every viewing item (refer to label 2, Fig. 1), learners will be provided with a good learners' rating (refer to label 1, Fig. 1) which is located in the top section of the slide (learning material). The rating gives the indication whether the viewing item is highly recommended by the good learners or it is less recommended. The rating feature is programmed as an Asynchronous Java Script and XML (AJAX) script that is independent

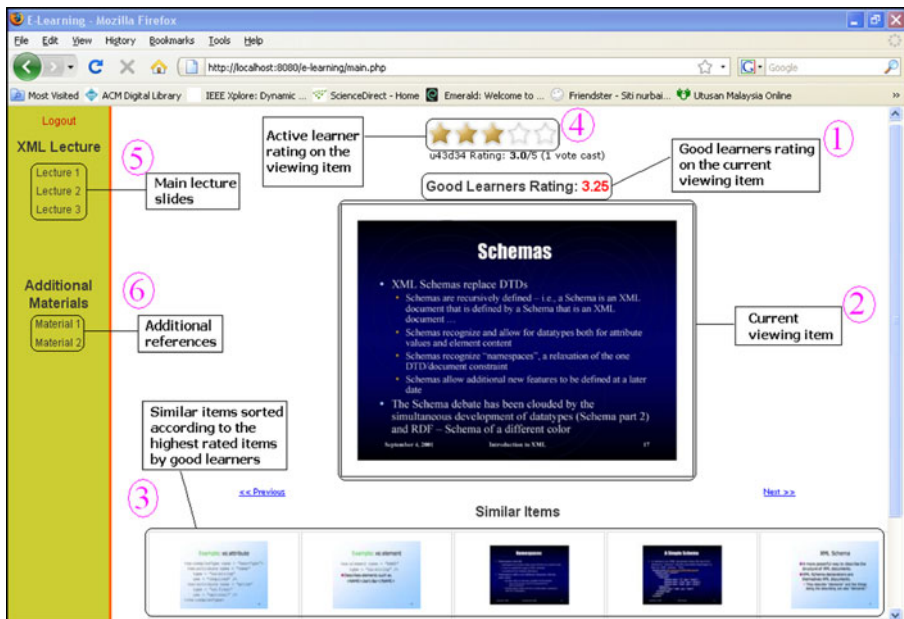


Fig. 1 Screen shot of the proposed e-learning recommender system

of any learning management system and can be easily embedded into any web page. Additionally, other similar items (refer to label 3, Fig. 1) are also recommended to the learners. These similar items are sorted according to their good learners' ratings, and are placed at the bottom of the page as shown in Fig. 1. Label 4 indicates the active learner's rating on the viewing item while the list of lectures and other additional materials which contain a set of slides are shown using label 5 and 6, respectively. The calculation and recalculation (e.g., adding, updating, and deleting keywords and ratings) for item's similarity and good learners' rating are done offline to avoid disrupting the learning process. The rest of the 'System description' section will describe the architecture and framework of the system in detail, as well as the calculation for item's similarity and good learners' average and prediction ratings.

Architecture

Figure 2 shows the overall system architecture of the proposed e-learning recommender system. The *instructor* is responsible to create learning materials (documents) using the *converting* and *authoring tools*. The *converting tool* is used to convert the documents, such as presentations and word documents into image files. Converting such documents into images preserves the fidelity of the original documents and saves time to re-write the documents using other *authoring tools*. On the other hand, the *authoring tool* is used to embed the image into a predefined html template page, set the link between html pages, and to provide keywords for the html page. All the documents are stored in a *document repository*. *Learners* are able to view and rate the viewing document. The viewing document is retrieved from the *document repository*. The *content profile builder* calculates the document weight that will be used later to calculate the items' similarity. The *rating repository* stores all the good learners' ratings. Finally the *rating profile builder* will query

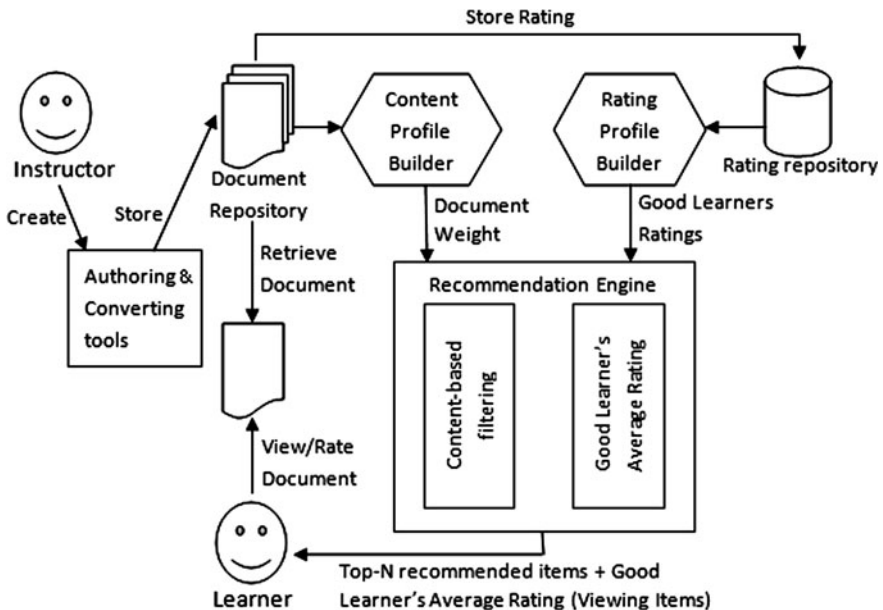


Fig. 2 Overall system architecture of the proposed e-learning recommender system

the rating repository for the ratings. Both the document weight and the ratings are then used by the *recommendation engine* to calculate the similarity of the items and the good learners' ratings.

To describe more specifically, the proposed e-learning recommender system can be divided into three parts. The first part discusses the recommendation framework to introduce the overall process flow. This is followed by the vector space model to recommend a similar content item. Finally, a recommendation based on good learners' ratings is presented together with the good learners' ratings prediction.

Recommendation framework

The recommendation process starts with the calculation of the similarity value among items. The *attributes table* is used for similarity value calculation between the items and the similarity value is stored in the *item-item matrix*. The similarity calculation is calculated using a vector space model as elaborated in '[Vector space model](#)' section. Top-N recommended items that exceed the similarity value threshold will be used to calculate good learners' predicted ratings. The *rating matrix* that consists of $m \times n$ (m is the total number of users and n is the total number of items) elements, initially stores the good learners' predicted ratings and the predicted ratings will be replaced once there are real ratings from the good learners' for the particular items. The good learners' predicted ratings are then stored in the *rating log*. Figure 3 depicts the overall recommendation process.

The *evaluation table* which stores the good learners' evaluation marks will query the rating matrix for good learners' ratings. The average ratings of good learners are then used for item recommendation. The *process stack* includes a set of predefined instructions that will be executed during one of the following events: a new item is inserted, an existing item is deleted, an item attribute is updated, an item is rated, and an item is re-rated.

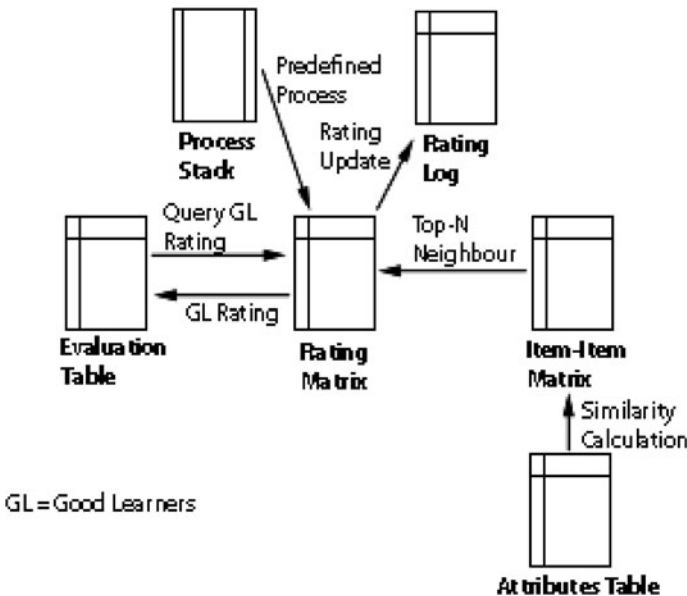


Fig. 3 The recommendation process flow

Vector space model

One of the main objectives in this study is to recommend items that are in the learning context, and to achieve this, we propose a recommender system that will only recommend the items that have similar content. We use vector space model (Eleni and John 2008) to calculate the similarity between the items. Similarities between the items are determined by measuring the angle between the two vectors, where a smaller angle indicates bigger similarity. Vector space model involves two main phases of calculation: (i) calculate the weight and (ii) calculate the cosine similarity.

The weight $w_{i,j}$ is calculated using the term frequency/inverse document frequency (TF-IDF) with normalized frequency as shown in Eq. 1. In TF-IDF, all the terms are treated as independent terms. The equation is defined as follows.

$$w_{i,j} = \frac{f_{i,j}}{\max_z f_{z,j}} \times \log\left(\frac{D}{d_i}\right) \quad (1)$$

where $f_{i,j}$ denotes the frequency of term i occurring in document j , $\max_z f_{z,j}$ is the maximum frequency among all the z keywords that appear in document j , D is the total number of documents that can be recommended to the learners, and d_i is the number of documents that contain term i . The normalized frequency ensures that the long documents with high occurrence terms will not have high impact on the weight thus it helps to reduce the possibility of keyword spamming (Castells et al. 2007).

The weight obtained from Eq. 1 is used to calculate the similarity value between the two documents. The relevancy rankings of the documents are measured based on the deviation angles between two document vectors that can be calculated using cosine similarity as follows.

$$\cos(\vec{w}_c, \vec{w}_s) = \frac{\vec{w}_c \cdot \vec{w}_s}{\|\vec{w}_c\| \|\vec{w}_s\|} \quad (2)$$

where \vec{w}_c and \vec{w}_s are treated as a vector of content based profile of user c and a vector of content of document s . Both $\|\vec{w}_c\|$ and $\|\vec{w}_s\|$ are the magnitude of the vectors \vec{w}_c and \vec{w}_s . Note that the value of the dot product between the two vectors can become 0 if there is no overlap keyword between them, whereas the value for the magnitude of the vectors is unlikely to become 0 unless keywords are not provided for the documents.

Good learners recommendation strategy

We use good learners' ratings as rating recommendation in order to help the learner to choose high quality items. The good learners' recommendation strategy starts by querying the *evaluation table*. If the good learners exist, then the system will query the good learners' ratings from the *rating repository* to calculate the good learners' average rating for the particular item. The mathematical equation is defined as follows.

$$R_{i,j} = \frac{\sum_{i=1}^N r_{i,j}}{N_j} \quad (3)$$

where $r_{i,j}$ is the rating of good learner i on item j , and N_j is the total number of good learners that rated item j . Note that the calculation for good learners' average rating is solely based on good learners' ratings.

If the good learners' ratings for a particular item do not exist, then the system will predict the good learners' ratings. The rating prediction is calculated using the similarity value retrieved from the *item-item matrix* and good learners' average rating queried from the *rating repository*. The mathematical equation is defined as follows.

$$P_i = \sum_{n=1}^N \frac{sim(d_i, d_n) \times R_n}{sim(d_i, d_n)} \quad (4)$$

where $sim(d_i, d_n)$ is the similarity between item i and item n [as calculated using Eqs. 1 and 2] and R_n is the good learners' average rating on item n [as calculated using Eq. 3]. For example, assuming that user 1 and user 2 are the only good learners, and the ratings of user 1 for item a and item b are 4 and 5, and the ratings of user 2 for item a and item b are 3 and 4. Thus the good learners' average rating for item a and b is 3.5 and 4.5, respectively. Let's say there is a new item c added to the system and has not received any ratings from the good learners yet, while the similarity value between item a and c is 0.8, and item b and c is 0.9. Thus the good learners' prediction rating for item c is 4.03.

Methodology

The experimental evaluation section is divided into four parts as follows: (1) 'Data sets', (2) 'Evaluation metrics', (3) 'Experimental procedure', and (4) 'Experimental results'.

Data sets

Although the data sets to test recommender systems are available at MovieLens (2009), the data is less suitable to test on e-learning recommender systems since the data sets are based on movie ratings. Furthermore, there are no standard data sets used previously to test e-learning recommender systems (Khribi et al. 2009; Liang et al. 2006). For the experiment, we use three main sets of PowerPoint slides on different topics of XML and two additional sets of slides as additional references. The total number of slides from the main and additional reference is 131 and each of the slides is treated as a separate item. All the slides are converted into images that are embedded in separate HTML pages and each of these slides is described by giving a title and some keywords.

Evaluation metrics

In this experiment, we measure both the learners' performance and the system's performance.

Learners' performance

Learners' performance is measured by calculating the mean score and the standard deviation for both the pre-test and the post-test for every group. The mean score and the standard deviation are then used to calculate the significant between the marks obtained from each group by using the t -test. We use a two-tailed test for the pre-test as we assumed that there is no significant difference between the pre-test marks among all the groups that use the e-learning without a recommender system, the e-learning with a content-based recommender system, and the e-learning with a content-based and good learners' ratings

recommendation method. For the post-test, we use a one-tailed test as a significant measure as we assumed that the group of learners that use the e-learning with a content-based and good learners' ratings recommendation have a higher post-test score compared to the group of learners that use the e-learning without a recommender system, and the group of learners that use the e-learning with a content-based recommender system. To further analyze the learners' performance, we present the number of students in a particular mark range and the percentage of marks' increments from pre-test to post-test for each group.

System's performance

The system's performance is measured by calculating the Mean Absolute Error (MAE) which determines the difference between the predicted rating and the real rating from the user. This measurement is important since it determines the accuracy of the rating prediction from the user given rating (Hernandez del Olmo and Gaudioso 2008). The formula is given as follow:

$$\text{MAE} = \frac{\sum_{i=1}^N |p_i - r_i|}{N} \quad (5)$$

where p_i is the predicted rating for item i , r_i is the user given rating for item i , and N is the total number of the pair rating p_i and r_i .

Besides measuring the MAE, we also calculate the Precision and Recall of the recommender systems to measure the decision-support accuracy that indicates how effectively predictions help a user to select high-quality items from the item set (Kunaver et al. 2007). The mathematical formula for Precision and Recall are given as follow

$$\begin{aligned} \text{Precision} &= \frac{tp}{tp + fp} \\ \text{Recall} &= \frac{tp}{tp + fn} \end{aligned} \quad (6)$$

where tp stands for true positive, fp stands for false positive, and fn stands for false negative. We set the threshold for determining true positive to 0.7 meaning that if an item is rated 0.7 or higher, it is considered to be accepted by the user. The value for Precision and Recall ranging from 0 to 1, where 0 indicates worst value and 1 indicates best value.

Experimental procedure

Students (a batch of second year students majoring in software engineering) from three different classes participated in this experiment. They were divided into groups according to the class that the students were registered in. Group 1 (G1) consisted of 21 students of the first class, Group 2 (G2) consisted of 21 students of the second class, and Group 3 (G3) consisted of 24 students of the third class. G1 used the e-learning without a recommender system, G2 used the e-learning with a content-based recommender system, and G3 used the proposed e-learning recommender system to study the learning materials. However, the G3's experiment was conducted after G1's and G2's, because the good learners' ratings which were required as initial input for the proposed e-learning recommender system were to be generated during the experiments on G1 and G2 and by their sampled students (Note that, although the e-learning systems used by G1 and G2 had no rating features, the rating features were included in both e-learning systems specially for the purpose of this experiment, so that the good learners' ratings were captured from G1 and G2, and subsequently

fed to the proposed e-learning recommender system as initial good learners' ratings for the experiment on G3). The students for all groups were needed to register with the assigned e-learning systems and sit for the same pre-test which consisted of 10 multiple choice questions before they were able to login to view the learning materials. The students were given 1 week to study all the learning materials using the assigned e-learning systems before sitting for the post-test which consisted of 15 multiple choice questions. During the process of learning, students of all groups were encouraged to rate the learning materials which they found useful in helping them to understand the learning topic better. The ratings done by G1 and G2 at the other end were used by the proposed system as initial good learners' ratings, and the ratings done by G3 were used for updating good learners' ratings of the proposed system concurrent to the experiment. The questions for pre-test and post-test were arranged in different orders and the save function was set to disable for the pre-test and post-test web pages to avoid the possibility of cheating by the groups. All the assessments were conducted during a formal class in a monitored environment. Furthermore, the experiments were conducted on three different classes, instead of one class, to minimize the possibility of collaborating between the students from different groups. The experiments were unknown to the students until they were given the URL address of the assigned e-learning system. All of these precautions were taken since the experiment on G3 was conducted after the test on G1 and G2. Figure 4 summarizes the experimental procedure.

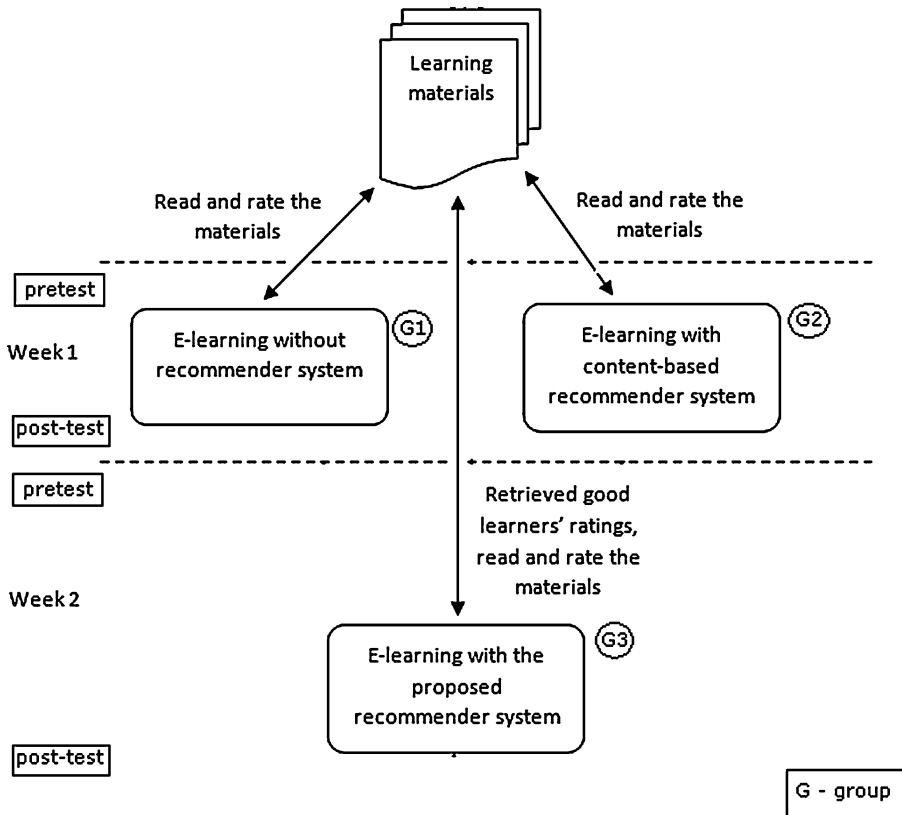


Fig. 4 Summary of the experimental procedure

Experimental results

The results section is divided into two parts: (i) ‘Student performance’ and (ii) ‘System performance’.

Student performance

The obtained results show that the mean score for the pre-test marks for all 3 groups (G1: $M = 40.48$, $SD = 12.44$; G2: $M = 35.71$, $SD = 15.02$; G3: $M = 36.67$, $SD = 15.79$) are relatively the same; with G1 having the highest mean score for the pre-test. The t -test results further revealed that there is no significant difference at $p < 0.05$ between the pre-test marks among all groups (G1 and G2: $t = 1.12$, $df = 40$, $p = 0.2694$, $d = 0.346$; G1 and G3: $t = 0.904$, $df = 43$, $p = 0.3710$, $d = 0.268$; G2 and G3: $t = 0.207$, $df = 43$, $p = 0.8370$, $d = 0.062$).

For the post-test, G3 ($M = 67.22$, $SD = 14.96$) has the highest mean score compared to G2 ($M = 58.41$, $SD = 14.13$) and G1 ($M = 59.05$, $SD = 15.68$). Based on the t -test score calculation for post-test marks, there is a significant difference at $p < 0.05$ between the post-test marks obtained by G3 when compared to G1 ($t = 1.78$, $df = 43$, $p = 0.0821$, $d = 0.533$) and G2 ($t = 2.03$, $df = 43$, $p = 0.0486$, $d = 0.605$). However, there is no significant difference at $p < 0.05$ between the post-test marks obtained by G1 and G2 ($t = 0.14$, $df = 40$, $p = 0.8894$, $d = 0.042$).

Referring to Fig. 5, all groups produce the same number of good learners (marks above 80%) while G1 produces the highest number of students who score below 50%. Significantly, G3 has the highest percentage of mark-increments from pre-test to post-test at about 45%, followed by G2 with 38% of mark-increments and G1 with 31% of mark-increments.

System performance

The total number of ratings received for all the 131 items is 6704. It is clear that the MAE value obtained by the e-learning with the content-based filtering and good learners’ ratings (CBF-GL) technique is lower than the e-learning with the content-based filtering (CBF) technique, thus the CBF-GL technique has a better prediction accuracy (as shown in Fig. 6). It is also obvious that the CBF-GL technique obtained the highest precision with the expense of recall. In contrast, the CBF technique has slightly higher recall compared to the CBF-GL technique (as shown in Figs. 7 and 8).

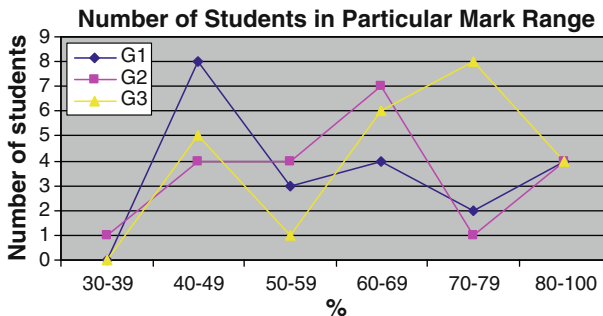


Fig. 5 The number of students in a particular mark range

Fig. 6 Comparison between the MAE of CBF and CBF-GL technique

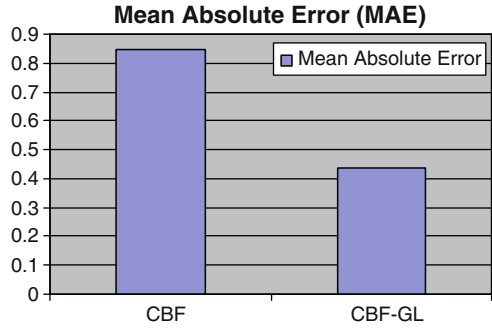


Fig. 7 Comparison between the precision of CBF and CBF-GL technique

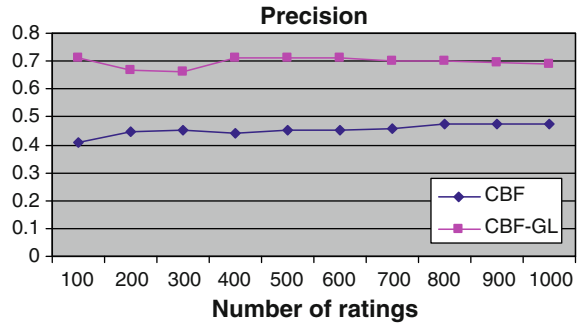
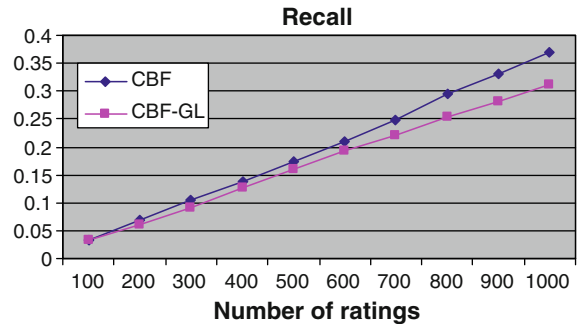


Fig. 8 Comparison between the recall of CBF and CBF-GL technique



Implications

The use of good learners' ratings as rating recommendation has helped the learners to choose high quality items and thus increase the learning performance. Moreover, incorporating the good learners' ratings into a content-based recommender system has improved the rating prediction accuracy. While the proposed system has been successfully tested on a real learning environment, it is worth to consider few issues pertaining to the system development for further improvement. Our proposed system used a vector space model (VSM) to calculate the item's similarity where the VSM itself has several limitations (Eleni and John 2008). VSM requires an intensive calculation for the terms and the similarity between documents that leads to a longer processing time. All the vectors need to be recalculated each time a new term is added, deleted, or updated in the term space. A proper

scheduling for the recalculation need to be planned so that it will not disrupt the learning process as well as it provides the latest rating recommendation.

The choice of the keywords is the determining factor for the level of similarity among documents. However, devising an effective method that picks out suitable keywords is not easy, and sometimes similar documents are having low similarity due to improper description. A standard vocabulary or ontology for keywords should be used to have more precise matching between the documents. Plenty of keywords and the most common words should be avoided to describe a particular learning material as it will decrease the similarity value (it leads to many topic-dimensions, and the document becomes multi-topic which means it is not focused on a certain topic).

On top of the issue of providing precise keywords to the documents, the major problem is that the effort and time required in providing keywords to a very large number of documents. An automatic keyword extraction (Matsuo and Ishizuka 2004) can be used to speed up the process of extracting keywords from documents. The trade-off of using automatic keyword extraction is the keyword precision. A higher precision can be achieved by using several combinations of automatic keyword extraction methods and using the overlap keywords as input to the VSM.

The implementation of the rating feature has imposed an extra task to the learners by having not only to study the learning materials but also to judge the quality (usefulness) of the learning materials (by giving an explicit rating). Some learners may judge a very useful learning material as 4 (from a scale of 1–5) but others may judge it as 5. An implicit rating (Yoon and Jae 2004) such as the time spent on reading the materials and the page visit frequency can be used to replace the explicit rating when the amount of learning materials is huge. The downside of implicit rating is that the system is assuming the user's preferences based on the user's behaviors which can lead to a false perception.

Conclusions and future works

In this paper, we propose a recommender system for e-learning that is able to recommend similar items to the viewing item and also to recommend items based on good learners' ratings. Furthermore, the accuracy of the system as well as its impact on student's performance has been put to test thoroughly. The results show that the proposed e-learning recommender system performed better in terms of a small rating deviation and has better precision as compared to an e-learning with a content-based recommender system. Moreover, this study shows that the rating feature in e-learning plays an important role as it can help to improve the performance of the learners. The proposed method shows an increment of 12.16% with the effect size of 0.6 and 13.11% with the effect size of 0.53 for the post-test results when compared to the e-learning without a recommender system and the e-learning with a content-based recommender system, respectively. Our work is helpful to distance education learners since they are among those who have limited contact hours with the instructors and other students.

There are several works that can be done in the future to further justify and enhance our work. First of all, the system was tested on Software Engineering students whose computer literacy is expected to be high. Thus the use of ratings as a recommendation tool is easily understandable by the students. It would be appropriate to test the system with other groups of students that have a different level of computer literacy (e.g., sociology students) or a different level of education (e.g., primary school students and secondary school students). It is important to note that the learning process must be followed by assessments (post-tests

to categorize the students in which the good learners' ratings will be used as recommendation). Also the testing was done on a predefined set of learning materials and a single type of learning material (lecture slides) that have been determined by the course lecturer. It would be interesting to see the accuracy of the system and the student's performance when different sets and types of learning materials are used. To further justify and increase the value of the good learners' ratings and its effect on learning, we plan to incorporate the good learners' ratings into other types of content-based filtering algorithms. The content-based filtering algorithm has a direct impact on the rating recommendation since one of the variables to calculate the good learner's predicted rating depends on the content similarity (which is calculated using content-based filtering algorithm). Currently, the authors are working on automation of the rating feature so that it can be easily integrated into newly added learning materials by the lecturer or any supportive learning materials uploaded by the students. This rating feature enhancement is important to alleviate the burden of having to provide the learning materials, as well as using the authoring and converting tools to convert the learning materials and embed the rating feature, that currently lies on the instructors. All these processes will be automated in the future. The authors are also working on the ontology-based live tagging system so that the predefined and new keywords can be easily added to the learning materials by a drag-and-drop mechanism to describe the learning materials in a structured format. As our goal is to focus on the rating feature rather than the learning management system (LMS) itself, it would be feasible to integrate the proposed system into an existing LMS like Moodle (2009) as a plug-in. This integration will yield an accurate result since the existing users are already familiar with the LMS (need less time to learn other features of the LMS) and they only need to adapt to the rating feature as a recommendation tool. The large numbers of users using Moodle will also help to promote the proposed system and increase its usage. Finally, it is our intention to compare our proposed content-based filtering system with good learners' ratings against content-based filtering systems with all learners' ratings (as previously proposed by other researchers) to determine how much gains can be achieved when using only the ratings of good learners.

References

- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 734–749.
- Aijuan, D., & Baoying, W. (2008). Domain-based recommendation and retrieval of relevant materials in e-learning. In *IEEE international workshop on semantic computing and applications 2008 (IWSCA '08)* (pp. 103–108).
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bobadilla, J., Serradilla, F., Hernando, A., & MovieLens. (2009). Collaborative filtering adapted to recommender systems of e-learning. *Knowledge-Based Systems*. doi:10.1016/j.knosys.2009.01.008.
- Castells, P., Fernandez, M., & Vallet, D. (2007). An adaptation of the vector-space model for ontology-based information retrieval. *IEEE Transaction on Knowledge and Data Engineering*, 19(2), 261–272.
- Chen, C. M., Lee, H. M., & Chen, Y. H. (2005). Personalized e-learning system using item repository theory. *Computers & Education*, 44(3), 237–255.
- De Bra, P., Smith, D., & Stash, N. (2006). The design of AHA! In *Proceedings of the seventeenth conference on hypertext and hypermedia* (pp. 133–134). Odense.
- Eleni, M., & John, K. (2008). Utilizing vector space models for user modeling within a learning environments. *Computers & Education*, 51(2), 493–505.
- Felferning, A., Friedrich, G., & Schmidt-Thieme, L. (2007). Recommender systems. *IEEE Intelligent Systems*, 22(3), 18–21.

- Hernandez del Olmo, F., & Gaudioso, E. (2008). Evaluation of recommender systems: A new approach. *Expert Systems with Applications*, 35(3), 790–804.
- Kerkiri, T., Manitsaris, A., & Mavridou, A. (2007). Reputation metadata for recommending personalized e-learning resources. In *Proceedings of the second international workshop on semantic media adaptation and personalization* (pp. 110–115). Uxbridge.
- Khribi, M. K., Jemni, M., & Nasraoui, O. (2009). Automatic recommendations for e-learning personalization based on web usage mining techniques and information retrieval. *Educational Technology & Society*, 12(4), 30–42.
- Kim, Y., & Baylor, A. L. (2006). A social-cognitive framework for designing pedagogical agents as learning companions. *Educational Technology Research and Development*, 54(6), 569–596.
- Kunaver, M., Pozrl, T., Pogačnik, M., & Tasic, J. (2007). Optimisation of combined collaborative recommender systems. *International Journal of Electronics and Communications*, 61(7), 433–443.
- Liang, G., Weining, K., & Junzhou, L. (2006). Courseware recommendation in e-learning system. In *Advances in web based learning—ICWL2006* (pp. 10–24). Berlin/Heidelberg: Springer.
- Liu, F., & Shih, B. (2007). Learning activity-based e-learning material recommendation system. In *Proceedings of the ninth IEEE international symposium on multimedia workshops (ISMW'07)* (pp. 343–348). Beijing, China.
- Matsuo, Y., & Ishizuka, M. (2004). Keyword extraction from a single document using word co-occurrence statistical information. *International Journal of Artificial Intelligence Tools*, 13(1), 157–169.
- Moodle. (2009). Retrieved March 17, 2009, from <http://moodle.org/>.
- MovieLens Data Sets. (2009). Retrieved May 02, 2009, from <http://www.grouplens.org/node/73>.
- Otair, M. A., & Hamad, A. Q. A. (2005). Expert personalized e-learning recommender system. <http://medforist.grenoble-em.com/Contenus/Conference%20Amman%20EBEL%2005/pdf/21.pdf>.
- Santos, O. C., & Boticario, J. G. (2008). Recommendation strategies for promoting eLearning performance factors for all. In *The 6th workshop on intelligent techniques for web personalization & recommender systems*. Illinois, USA.
- Soonthornphisaj, N., Rojsattarat, E., & Yim-ngam, S. (2006). Smart E-learning using recommender system. In *Computational intelligence* (Vol. 4114, pp. 518–523). Berlin/Heidelberg: Springer.
- Tai, D. W., Wu, H., & Li, P. (2008). Effective e-learning recommendation system based on self-organizing maps and association mining. *The Electronic Library*, 26, 329–344.
- Tan, H., Guo, J., & Li, Y. (2008). E-learning recommendation system. In *The 2008 international conference on computer science and software engineering* (Vol. 5, pp. 430–433).
- Tang, T. Y., & McCalla, G. (2003). Smart recommendation for an evolving e-learning system: Architecture and experiment. *International Journal on E-learning*, 4(1), 105–129.
- Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, 25(6), 631–645.
- Yoon, H. C., & Jae, K. K. (2004). Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce. *Expert Systems with Applications*, 26(2), 233–246.
- Zaiane, O. R. (2002). Building a recommender agent for e-learning systems. In *Proceedings of the international conference on computers in education (ICCE'02)* (pp. 55–59). Auckland, New Zealand.
- Zheng, Y., & Li, L. (2008). A three-dimensional context-awareness model for peer recommendation in the e-learning context. *International Journal on E-Learning (IJEL)*, 7(1), 153–168.

Khairil Imran Ghauth graduated his degree in Bachelor of Information Technology (Hons.) from Multimedia University, Malaysia in 2005. He obtained the degree of Master of Information Technology from the University of Melbourne in the following year. He is currently lecturing in Multimedia University and doing his Ph.D. in University of Malaya, Malaysia in the area of E-learning Recommender System.

Nor Aniza Abdullah received the B.Sc. degree in Computer Science from the University of Malaya, Malaysia, the M.Sc. degree in Interactive Multimedia from the University of Westminster, London, and the Ph.D. degree in Computer Science from the Southampton University, United Kingdom. She is currently a senior lecturer at the Faculty of Computer Science and Information Technology in University of Malaya, Malaysia.