The Effect of Incorporating Good Learners' Ratings in e-Learning Content-based Recommender System

Khairil Imran Ghauth and Nor Aniza Abdullah
Faculty of Information Technology, Multimedia University, Malaysia // Faculty of Computer Science and Information Technology, University of Malaya, Malaysia // khairil-imran@mmu.edu.my // noraniza@um.edu.my

ABSTRACT
One of the anticipated challenges of today’s e-learning is to solve the problem of recommending from a large number of learning materials. In this study, we introduce a novel architecture for an e-learning recommender system. More specifically, this paper comprises the following phases i) to propose an e-learning recommender system based on content-based filtering and good learners’ ratings, and ii) to compare the proposed e-learning recommender system with exiting e-learning recommender systems that use both collaborative filtering and content-based filtering techniques in terms of system accuracy and student’s performance. The results obtained from the test data show that the proposed e-learning recommender system outperforms existing e-learning recommender systems that use collaborative filtering and content-based filtering techniques with respect to system accuracy of about 83.28% and 48.58%, respectively. The results further show that the learner’s performance is increased by at least 12.16% when the students use the e-learning with the proposed recommender system as compared to other recommendation techniques.

Keywords
E-learning, Recommendation system, Good learners’ ratings, Content-based filtering, Collaborative filtering

Introduction
Web-based learning environments are becoming very popular nowadays as a means of delivering lectures or simply as a place to share notes. This has led to the creation of huge amounts of digital learning materials that are either used as mandatory or supporting materials in the web-based learning environments (MERLOT, 2009). The arising problem with the existence of such huge amounts of learning materials is how to recommend quality items to learners when they have limited time to view and study all the learning materials (Nachmias & Segev, 2003). We are motivated to solve this problem by proposing a new e-learning recommender system that is able to recommend quality items to learners. This in turn can improve the student’s performance.

A recommender system is a tool that supports users in identifying interesting items especially among large numbers of items. Among the popular approaches used in recommender systems are using either a collaborative filtering or a content-based filtering technique. Collaborative filtering identifies the interesting items from other similar users’ opinions by calculating the nearest-neighbor from a rating matrix. New items that are of interest to the nearest-neighbor and that have not been rated by other users with similar interest will be recommended to them. In contrast, content-based filtering uses features of items to infer recommendations. Hence, items with similar content to the current viewing item will be recommended to the active user (Felfernig et al., 2007). As like in other domains, recommender systems in e-learning can differ in many ways depending on what kind of object is to be recommended (i.e., course to enroll, learning materials, and etc.) and whether the context of learning is considered important (Soonthornphisaj et al., 2006; Liang et al., 2006; Tang & McCalla, 2003). For example, when a learner is reading on C++ arrays, the learner is expecting to get recommended items that are able to support and enhance their understanding of the current learning context, namely the topic of arrays in the programming language of C++. The recommendations would be different if the learner is reading about a business topic of study where the context may depend on several other fields such as finance and accounting.

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. Helping and support between peers can happen in many ways such as teaching, demonstrating, or sharing materials. Topping (2005) uses the term “peer helper” for someone who is considered among the “best students” and acts as a surrogate teacher, in a linear model of the transmission of knowledge, from 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 where by exhibiting good learners’ behavior (i.e., focus on highly rated items) can increase students’ performance.

The objective of this study is to propose a new e-learning recommender system based on the content-based filtering and good learners’ ratings techniques. Our proposed method ensures that the recommended items will remain in the current learning context. The good learners’ ratings are used in this study as rating recommendation that will help other learners to focus and choose the best learning materials. A good learner is a learner that scores more than 80% in the post-test. The terms learning materials, items, and documents are used interchangeably through the whole paper.

In summary, a new e-learning recommender system is proposed in this paper that uses the combination of content-based filtering and good learners’ ratings. We evaluate the proposed e-learning recommender system by comparing the recommender system accuracy and the decision-support accuracy with several techniques of recommendation in e-learning. In addition, we compare the student performance of several groups that make use of (i) an e-learning system without a recommender system, (ii) an e-learning system with the proposed recommender system, (iii) an e-learning recommender system with content-based filtering, and (iv) an e-learning recommender system with collaborative filtering.

The remaining part of this paper is organized as follows. ‘Related Work’ section will discuss the current work of e-learning recommender systems. ‘Learning Materials Recommendation Framework’ section will discuss the process involved in the modeling and recommendation phase. In each phase, the mathematical models that are used for calculations are presented. ‘Experimentation and Results’ section will discuss the data set, experiment setup, metrics used for measurement, and the result of the experiment. Note that some of the works described in ‘Learning Materials Recommendation Framework’ section and ‘Experimentation and Results’ section have been reported in Gauth & Abdullah (2009), Gauth & Abdullah (2010a), and Gauth & Abdullah (2010b). The main differences is that this paper focuses on both student’s performance and recommender system accuracy as well as comparing the proposed recommender system results (i.e., student’s performance and recommender system accuracy) with various types of recommender system techniques. Finally, ‘Conclusion and Future Works’ section provides the concluding remarks along with suggestions for future work.

**Related Work**

Recent trends in e-learning recommender systems show that most of the researchers use the data mining approach and the information retrieval technique as the recommendation strategies (Zaiane, 2002; Liang et al., 2006; Kerkir et al., 2007). Zaiane (2002) proposed the use of web mining techniques to build agents that could recommend online learning activities or shortcuts in a course website based on learners’ access histories to improve course navigation as well as assist with the online learning process. Khribi et al. (2008) compute 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. Soonthornphaisaj et al. (2006) apply the collaborative filtering approach to predict the most suitable documents for the learner. New learning materials are able to be recommended to learners with a high degree of similarity. They were also proposing a new e-learning framework using web services that has the ability to aggregate recommended materials from other e-learning web sites and predicts more suitable materials for learners. Liu et al. (2007) designed a material recommendation system based on association rule mining and collaborative filtering. The system is implemented by integrating the techniques of LDAP and JAXB to reduce the load of development of search engine and the complexity of the content parsing for improving the learning performance of learners. Liang et al. (2006) applies the knowledge discovery techniques, and a combination of content-based filtering and collaborative filtering to make personalized recommendations for a courseware selection module. Their experiment shows that the algorithm used is able to reflect users’ interests with high efficiency. Tang et al. (2003) proposed an evolving web-based learning system that is 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 use a clustering technique to cluster the learners before using collaborative filtering to calculate learners’ similarities for content recommendation. Kerkiri et al. (2007) proposed a framework that exploits 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) proposed a personalized e-learning system based on Item Repository Theory which estimates the abilities of online learners and recommends appropriate course materials to learners. The experiment shows that the system can precisely provide personalized course material recommendations based on learners’ abilities and accelerate learners’ learning efficiency and effectiveness. Otair et al. (2005) proposed a framework for an expert personalized e-learning recommender system by using a rule-based expert system that can help learners in finding learning materials that best suits their needs. Tai et al. (2008) proposed e-learning course recommendation based on artificial neural network (ANN) and data mining techniques. ANN is used to classify the learners based on groups of similar interests and learners can obtain course recommendations from the group’s opinion. They used a data mining technique to elicit the rules of the best learning path.

From the previous related works, none of the recommender systems have attempted to use content-based filtering together with good learners’ ratings as learning materials recommendation method. The benefits of having both content-based filtering and good learners’ ratings as recommendation methods is that it will ensure the recommended items remain in the current learning context and quality materials are able to be recommended. We have chosen the method proposed by (Soonthornphisaj et al., 2006) to be compared with our proposed e-learning recommender system in terms of system accuracy and student’s performance since they used an information retrieval technique for recommendation that is similar to the one in our proposed work.

Learning Materials Recommendation Frameworks

Figure 1 shows the good learners’ recommendations strategy framework. Before the recommendation process begins, the instructors have to upload the learning materials and provide keywords for the learning materials. The learning materials can be either mandatory learning materials or additional learning materials. All the learning materials will be converted into image to preserve the fidelity of the original presentation of the learning materials. Our system provides an easy tool for instructors to upload the image and provide keywords. The keyword is used to describe the learning materials and it consists of information about the author, title, and keywords. Once the learning materials are stored in the server, the recommender system will query for the keywords of the learning materials. The keywords are then used to calculate the items’ similarity. Good learners’ ratings are used as rating recommendations on every learning material. If the good learners’ ratings are available for a particular item, the average rating of the good learners will be used as rating recommendation. Otherwise, the recommender system will predict the good learners’ ratings for the item. Finally the learner is able to view the learning materials. For each viewing learning material, it will be recommended with similar learning materials and good learners’ ratings. ‘Modeling Phase’ will describe the process which involves creating a content profile builder and a rating profile builder. Meanwhile, ‘Recommendation Phase’ will describe the process which involves selecting top-N similar items and calculating the good learners’ ratings for a particular item.

![Figure 1: The good learners’ recommendation strategies framework](image-url)
**Modeling Phase**

The modeling phase involves creating the content profile builder and the rating profile builder. The content profile builder is responsible to query the document repository for the item’s keywords that is to be used for calculating the document weight. The document weight is then used to calculate the document similarity. The document weight $w_{i,j}$ is calculated using the term frequency/inverse document frequency (TF-IDF) with normalized frequency as shown in equation (1). In TF-IDF, all the terms are treated as independent terms. The equation is defined as follows.

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

where $f_{i,j}$ denotes the frequency a term $i$ occurs in document $j$. The $\text{max}_z f_{z,j}$ is the maximum frequency among all the $z$ keywords that appear in document $j$. The $D$ is the total number of documents that can be recommended to the learners. The $d_j$ 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 computed using equation (1) is used to calculate the similarity value between the two items. The cosine similarity value is defined as follows.

$$\cos(w_c, w_s) = \frac{w_c \cdot w_s}{\| w_c \| \| w_s \|} \quad (2)$$

where $w_c$ and $w_s$ are treated as a vector of content based profile of user $c$ and the content of document $s$. Both $\| w_c \|$ and $\| w_s \|$ are the magnitude of the vector $w_c$ and $w_s$. Similarities between the documents are measured by measuring the angle between the two vectors where a smaller angle indicates a higher similarity.

Meanwhile, the rating profile builder is responsible to query the good learners’ ratings from the rating repository. The ratings will be used in recommendation phase (refer to ‘Recommendation Phase’ section) to calculate the good learners’ ratings for a particular item.

**Recommendation Phase**

The recommendation phase involves selecting the top-N similar items calculated earlier [as calculated using equation (2)] in the modeling phase and recommending each item along with its good learners’ ratings. The good learners’ average rating is obtained by calculating the average rating of good learners’ ratings on a 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 $j$. The $N_j$ is the total number of good learners that rated item $j$. Note that the calculation for good learners’ average rating on a particular item is solely based on good learners’ ratings.

In the case where the item has not received any rating from the good learners, then the item will be recommended with good learners’ prediction rating that is calculated as follows.

$$p_i = \sum_{n} \frac{\text{sim}(d_i, d_n) \cdot R_n}{\text{sim}(d_i, d_n)} \quad (4)$$

where $\text{sim}(d_i, d_n)$ is the similarity between item $i$ and item $n$ [as calculated using equations (1) and (2)] and $R_n$ is the good learners’ average rating on item $n$ [as calculated using equation (3)].
Experimentation and Results

Although the data sets to test the recommender system are available at MovieLens (2009), the data included in (MovieLens, 2009) is less suitable for testing on the e-learning recommender systems, since the data sets are based on movie ratings. Therefore, for this experiment, we use three main sets of PowerPoint slides on different topics of XML and also two additional sets of PowerPoint slides as additional references that comprise 131 separate items. All the slides are converted into images and each of the slides is described by author-defined keywords. The images are embedded into an HTML page and can be accessed by an authorized learner.

The experiments were conducted on 95 university students (second year students of Software Engineering) from four different classes who were undertaking the Web Services course. Group 1 (G1) consists of 21 students from the first class that use the e-learning without a recommender system, group 2 (G2) consists of 21 students from the second class that use the e-learning with a content-based recommender system (G2), group 3 (G3) consists of 24 students from the third class that use the e-learning with the proposed recommender system, and finally, group 4 (G4) consists of 29 students from the fourth class that use the e-learning with a collaborative filtering recommender system (G4) as proposed by (Soonthornpisaj et al., 2006). The experiments on G1 and G2 were conducted in week 1 (a week before experiment on G3), since we need to get the good learners’ ratings from G1 and G2 to be used in the proposed recommender system. The experiment on G3 was conducted on week 2. On the other hand, the experiment on G4 was conducted on week 3 (a week after the experiment on G3), as the experiment on G4 needs the ratings from a large number of users (i.e., users from G1 until G3) to calculate the users’ similarity and to minimize the cold-start problem. Cold-start problem is a problem where the items cannot be recommended due to insufficient ratings received (Herlocker et al., 2004). Initially, no items will be recommended to the users in G4 until the users have rated a few items and the similarity values between users have been updated. The system will periodically calculate the users’ similarities as the similarity value will change each time the user provides a rating or re-rates a learning material. Before the students can use the assigned systems, they will have to sit for the same pre-test. It is used to assess the pre-knowledge of the students before they start the formal learning using the assigned systems. The students are then given one week to study and rate the learning materials based on the usefulness of the materials. The learning process is then followed by a post-test. The questions for the pre-test and the post-test are arranged in different orders and the save function is set to disabled for the pre-test and the post-test web pages as it helps to avoid the possibility of cheating. All the assessments (pre-test and post-test) are conducted during a formal class in a monitored environment. Furthermore, the experiments are conducted on four different classes, instead of one class, to minimize the possibility of collaborating between the students from different groups. The experiments are unknown to the students until they are given the URL address of the assigned e-learning system. All of these precautions are taken into account since the experiment on G3 is conducted after the tests on G1 and G2 are over, and the experiment on G4 is conducted after the experiments on G1, G2, and G3 are over. The results obtained from the pre-test and the post-test are used to measure student’s performance.

We use Mean Absolute Error (MAE) to measure the accuracy of the recommender systems. MAE can be defined as the rating deviation between the predicted rating and the user-given rating (Hernandez del Olmo & Gaudioso, 2008). The smaller the MAE value indicates that the rating prediction is closer to the user-given ratings and the recommender system has a high accuracy. The formula is given as follows:

\[
MAE = \frac{\sum_{i=1}^{N} |p_i - r_i|}{N}
\]  

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 ratings \( p_i \) and \( r_i \).

Figure 2 shows the MAE value of the three types of recommender systems. We receive a total of 9803 ratings for all the items. It is clear that the proposed system has the smallest MAE value that is less than 0.5. In contrast, the collaborative filtering technique has the highest MAE value which is more than 2.5. The content-based filtering technique yields a better MAE value than the collaborative filtering technique where the MAE value is about 0.8. Based on the obtained results, we can say that the proposed method has a better accuracy in terms of rating deviation compared to the content-based filtering and the collaborative filtering techniques.
Mean Absolute Error (MAE)

### Note
CF (collaborative filtering); CBF (content-based filtering); CBF-GL (content-based filtering with good learners rating strategies)

Figure 2: The Mean Absolute Error for all recommendation techniques

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

\[
\text{Precision} = \frac{tp}{tp + fp} \\
\text{Recall} = \frac{tp}{tp + fn} \\
F - \text{measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

(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, Recall, and F-measure ranges from 0 to 1, where 0 indicates the worst value and 1 indicates the best value.

Figure 3: The Precision for all recommendation techniques

Figures 3, 4, and 5 show the obtained Precision, Recall, and F-measure of the recommender systems. It is obvious that the proposed e-learning recommender system has the highest precision with the expense of recall. In contrast, the recommender system that uses collaborative filtering technique has the highest recall with the expense of
precision. The F-measure reveals that the proposed e-learning recommender system has slightly a higher F-measure value compared to the recommender system that uses the content-based filtering technique. Meanwhile, the recommender system that uses the collaborative filtering technique has the worst F-measure value. From the obtained results, we can say that the recommender system that uses the collaborative filtering technique has the worst decision-support accuracy and the proposed e-learning recommender system has the best decision-support accuracy.

**Figure 4:** The Recall for all recommendation techniques

**Figure 5:** The F-measure for all recommendation techniques

**Figure 6:** The pre-test and post-test average mark
The student’s performance is measured by calculating the average mark of each group and the t-score for the average mark among each pair of groups to determine the significance. We use a two-tailed test for the pre-test and a one-tailed test for the post-test. We assume in our hypothesis that in a pre-test, there is no significant difference in the average mark among the groups. For the post-test hypothesis, we assume that learners who use the proposed system will have a higher average mark compared to other groups. To further justify the student’s performance, we calculate the average percentage of mark increments from the pre-test to the post-test for each group.

![Standard Deviation](image)

*Figure 7: The pre-test and post-test standard deviation*

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test Mean</th>
<th>Pre-test Standard Deviation</th>
<th>Post-test Mean</th>
<th>Post-test Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>40.48</td>
<td>12.44</td>
<td>59.05</td>
<td>15.67</td>
</tr>
<tr>
<td>G2</td>
<td>35.71</td>
<td>15.02</td>
<td>58.41</td>
<td>14.13</td>
</tr>
<tr>
<td>G3</td>
<td>36.67</td>
<td>15.79</td>
<td>67.22</td>
<td>14.96</td>
</tr>
<tr>
<td>G4</td>
<td>34.48</td>
<td>13.78</td>
<td>50.11</td>
<td>15.79</td>
</tr>
</tbody>
</table>

*Table 1: The mean scores and standard deviations*

![Percentage of Mark Increment from Pre-test to Post-test](image)

*Figure 8: The percentage of mark increment from pre-test to post-test*

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test (two-tailed test)</th>
<th>Post-test (one-tailed test)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t</td>
<td>Df</td>
</tr>
<tr>
<td>G1-G2</td>
<td>1.12</td>
<td>40</td>
</tr>
<tr>
<td>G1-G3</td>
<td>0.90</td>
<td>43</td>
</tr>
<tr>
<td>G1-G4</td>
<td>1.61</td>
<td>48</td>
</tr>
<tr>
<td>G2-G3</td>
<td>0.21</td>
<td>43</td>
</tr>
<tr>
<td>G2-G4</td>
<td>0.30</td>
<td>48</td>
</tr>
<tr>
<td>G3-G4</td>
<td>0.53</td>
<td>51</td>
</tr>
</tbody>
</table>

*Note: * indicates statistically significance
The average mark and standard deviation for all groups are summarized in figure 6, figure 7, and Table 1 respectively. Table 2 summarizes the t-score (t) and the degree of freedom (df) between each group pairs for the pre-test and the post-test average marks. The results show that there is no significant difference at p < 0.05 between each group for the pre-test average mark. On the other hand, the post-test average mark obtained by G3 has a significant difference when compared with G1, G2, and G4 with a t-score value of 1.78 (significant at p < 0.05), 2.03 (significant at p < 0.025), and 4.04 (significant at p < 0.0005) respectively. Furthermore, G3 has the highest percentage of mark increment from the pre-test to the post-test which is about 45% as shown in figure 8. In contrast, G4 obtained the lowest percentage of mark increment from the pre-test to the post-test which is slightly above 30%.

Conclusion and Future Work

In this paper, we have discussed various techniques of e-learning recommender systems and we have proposed a new e-learning recommender system framework based on the content-based filtering and good learners’ ratings that offers better system accuracy and increases student’s performance. The implementation of content-based filtering ensures that the related items are in the learning context and the good learners’ rating serves as a guideline for other learners to choose from and focus on the learning materials. A comparative study among other e-learning recommender systems has been conducted for performance benchmarking. The experimental results show that the system’s accuracy of the proposed system is increased by 83.28% when compared to collaborative filtering technique and 48.58% when compared to content-based filtering technique. The student’s performance has also increased by at least 12.16%. These have showed that the introduction of good learners’ rating into recommender systems has improved both the system accuracy and increased the learner’s performance.

Even though the results have shown that the proposed system has produced a better accuracy and increase learner’s learning performance, there are several works that can be done in the future to further justify and enhance our work. First of all, the experiments were conducted on Software Engineering students whose computer literacy is expected to be high. Furthermore, the experiments were using a predefined set of learning materials prepared by one instructor. It will be interesting to see the performance of other group of learners with different level of education using the proposed system to study various subjects. This will help to justify whether the proposed system is suitable to be used by students from different level of knowledge and from various topics domain. In order to improve the accuracy of the proposed system, we plan to incorporate the good learners’ ratings with other types of content-based filtering algorithms. The content-based filtering algorithm has a direct impact on the calculation of good learners’ rating prediction and the better prediction accuracy may improve the recommendation. Since the number of learning materials used in the experiments is considered small, the task of assigning keyword was done by the instructor. The use of automatic keyword extraction should be considered if the number of learning materials is large. The accuracy of the system may be affected since the keywords assigned by the instructor may differ from the extracted keywords. As our proposed system disregard the user’s knowledge in recommending the items, we plan to incorporate the knowledge based method (Knutov, De Bra and Pechenizkiy, 2009) within the recommendation framework. This is to ensure that the learners are recommended with highly rated items by good learners and the items recommended are suit to the learner’s knowledge level. Lastly, we plan to make the proposed e-learning recommender system as a light program module that can be integrated into any web-based learning management system.

References


