Measuring learner's performance in e-learning recommender systems

Khairil Imran Ghauth
Multimedia University

Nor Aniza Abdullah
University of Malaya

A recommender system is a piece of software that helps users to identify the most interesting and relevant learning items from a large number of items. Recommender systems may be based on collaborative filtering (by user ratings), content-based filtering (by keywords), and hybrid filtering (by both collaborative and content-based filtering). Recommender systems have been a useful tool to recommend items in many online systems, including e-learning. However, not much research has been done to measure the learning outcomes of the learners when they use e-learning with a recommender system. Instead, most of the researchers were focusing on the accuracy of the recommender system in predicting the recommendation rather than the knowledge gain by the learners. This research aims to compare the learning outcomes of the learners when they use several types of e-learning recommender systems. Based on the comparison made, we propose a new e-learning recommender system framework that uses content-based filtering and good learners' ratings to recommend learning materials, and in turn is able to increase the student's performance. The results show that students who used the proposed e-learning recommender system produced a significantly better result in the post-test. The results also show that the proposed e-learning recommender system has the highest percentage of score gain from pre-test to post-test.

Introduction

Nowadays, learners are often overwhelmed with the large amount of learning materials available online. Despite having to spend time learning the materials, learners are lured into spending more time on browsing and filtering to identify information that suits their needs better, either in terms of knowledge value or preferences. Limited learning time can hinder learners in locating useful learning materials, as often they may end up getting irrelevant materials (Nachmias & Segev, 2003). One of the possible ways to overcome this problem is by using recommender systems.

A recommender system is a software tool that supports users in identifying interesting items, especially among large numbers of items. The popular approaches used in recommender systems are collaborative filtering, content-based filtering, and hybrid filtering. Collaborative filtering identifies the interesting items from other similar users' opinions by calculating the nearest neighbor (i.e. top-N users that have a similar rating pattern) from a rating matrix. New items that are of interest to the nearest neighbor and that have not been rated by the users 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, Friedrich & Schmidt-Thieme, 2007). Hybrid filtering on the other hand combines both content-based filtering and collaborative filtering techniques to produce a recommendation (Adomavicius & Tuzhilin, 2005). Recommender systems in e-learning can differ in many ways depending on the kind of object to be recommended, such as course to enrol, learning materials, and so forth, and whether the context of learning is considered important (Liang, Weining & Junzhou, 2006; Soonthornphisaj, Rojsattarat & Yim-ngam, 2006; Tang & McCalla, 2003).

While recommender systems have become a popular method of suggesting items, collaborative and peer learning systems have also emerged as an effective way of learning (Cecez-Kecmanovic & Webb, 2000; Topping, 2005). Topping (2005, p. 631) 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, p. 631) 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) states that people can learn by observing the behaviour of others and the outcome of those behaviours. Furthermore, the theory also mentions that other people will most likely exhibit the behaviour if the outcome is positive. This theory strongly supports the idea of learning from good learners, whereby exhibiting good learners' behaviour (i.e. focusing on highly rated items) can increase performance.

Our proposed recommender system produces a recommendation based on the combination of content-based filtering and good learners' ratings. The good learners' rating feature in our proposed recommender system promotes collaboration among learners to help each other during the learning process. The term good learner that is used in this study can be defined as a learner who has studied the learning materials and obtained a score of more than 80% in the post-test. Some of the works described below in the sections headed 'E-learning recommender system using good learners' ratings' and 'Method' have been reported in Ghauth & Abdullah (2009) and Ghauth & Abdullah (2010). In Ghauth & Abdullah (2009), the work on the recommendation process was explained in general and there had been no experiment conducted at that time. The experiment between the proposed recommender system and content-based filtering was described in Ghauth & Abdullah (2010). The article (Ghauth & Abdullah, 2010) also focused on the development of the proposed recommender system emphasising both the system accuracy and learner's performance. In contrast, the work described in this paper focuses solely on the learner's performance and we have extended the experiment by comparing outcomes from the proposed recommender system with outcomes from both collaborative filtering and hybrid filtering.

Previous research

Recent trends show that most of the researchers use data mining approaches and information retrieval techniques as their recommendation strategies (Kerkiri,

Manitsaris & Mavridou, 2007; Liang et al., 2006; Zaiane, 2002). Zaiane (2002) proposed the use of a web mining technique 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, Jemni and Nasraoui (2009) devised an online automatic recommendation system 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. Soonthornphisaj et al. (2006) applied 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 predict more suitable materials for learners. Liu and Shih (2007) designed a material recommendation system based on association rule mining and collaborative filtering. The system is implemented by integrating the techniques of LDAP (Lightweight Directory Access Protocol) and JAXB (Java Architecture of XML Binding) to reduce the development load of the search engine and the complexity of the content parsing for improving learning performance of learners.

Liang et al. (2006) applied a knowledge discovery technique, and a combination of content-based filtering and collaborative filtering to generate personalised recommendations for a courseware selection module. Their experiment shows that the algorithm used is able to reflect users' interests with high efficiency. Tang and McCalla (2003) proposed an evolving web-based learning system that is able to find relevant content on the web, personalise and adapt the content based on the system's observation of its learners and the accumulated ratings given by the learners, without the need for learners to directly interact with the open web. They use a clustering technique to cluster learners into a subclass according to the learning interest 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 personalised learning resources. Their experiment proved that the use of reputation metadata augmented learner's satisfaction by retrieving those learning materials which were evaluated positively.

Chen, Lee and Chen (2005) proposed a personalised 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 provide precisely personalised course material recommendations based on learners' abilities, and accelerate learners' learning efficiency and effectiveness. Otair and Hamad (2005) proposed a framework for an expert, personalised e-learning recommender system by using a rule-based expert system that can help learners in finding learning materials that best suit their needs. Tai, Wu and Li (2008) proposed an e-learning course recommendation based on artificial neural networks (ANN) and data mining techniques. ANN is used to classify 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.

In the previous literature, most of the researchers were focusing more on the algorithms and techniques to be used in the recommendation, without an emphasis

upon the effect on learner's knowledge gain. There was no research carried out to compare thoroughly the effectiveness of the e-learning recommender systems in improving students' performance. Furthermore, none of the researchers have attempted to use good learners' ratings as recommendation techniques. This study aims to address the above mentioned issues.

Research aims

This study adds to the body of knowledge on e-learning recommender systems in two ways. First, it extends the current, content-based e-learning recommender system framework by incorporating good learners' ratings. Second, learning outcomes for students who used the proposed e-learning recommender system are measured and compared to learning outcomes for students who used other types of e-learning recommender systems.

E-learning recommender system using good learners' ratings

The recommendation process begins after the annotated learning materials and the related keywords have been uploaded to a database by an instructor. The keywords are then retrieved by the recommendation engine for the document weight calculation. The document weight calculation calculates term frequency in both local (that is frequency of the term in the document itself) and global documents (that is frequency of the term in the whole document stored in the database), and the product between the local and the global term frequency. The resulting weight becomes the input for the cosine similarity calculation. This calculation creates a vector that represents a document in an n-dimensional term space. The relevancy rankings between the documents are determined by measuring the angle between the vectors. The smaller the angle the higher the similarity values between the two documents. The items' similarity values are stored in the item similarity database.
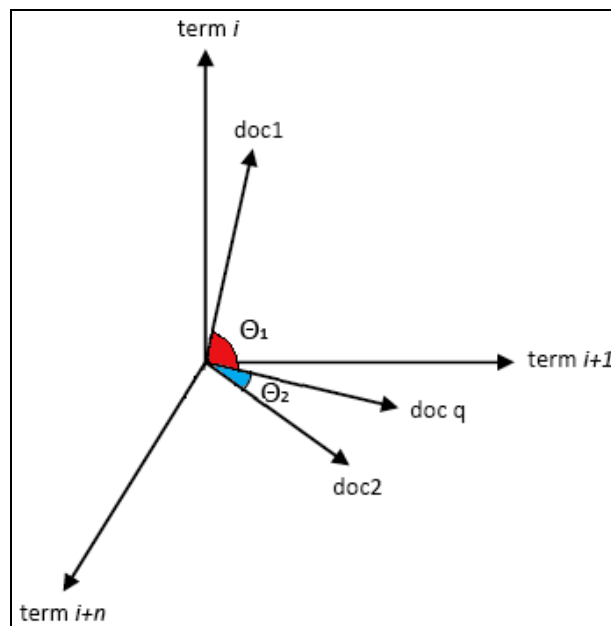
Figure 1: Similarity between documents

Figure 1 shows an example on how the similarities between documents are determined. The query document (doc q) has higher similarity with document 2 (doc 2) compared to document 1 (doc 1) since the angle from vector doc q to vector doc 2 is smaller compared to the angle from vector doc q to vector doc 1. The similarity values between documents are used to recommend a set of similar items and during the calculation of good learners' prediction rating.

Figure 2 depicts the overall process in our recommendation strategy framework.
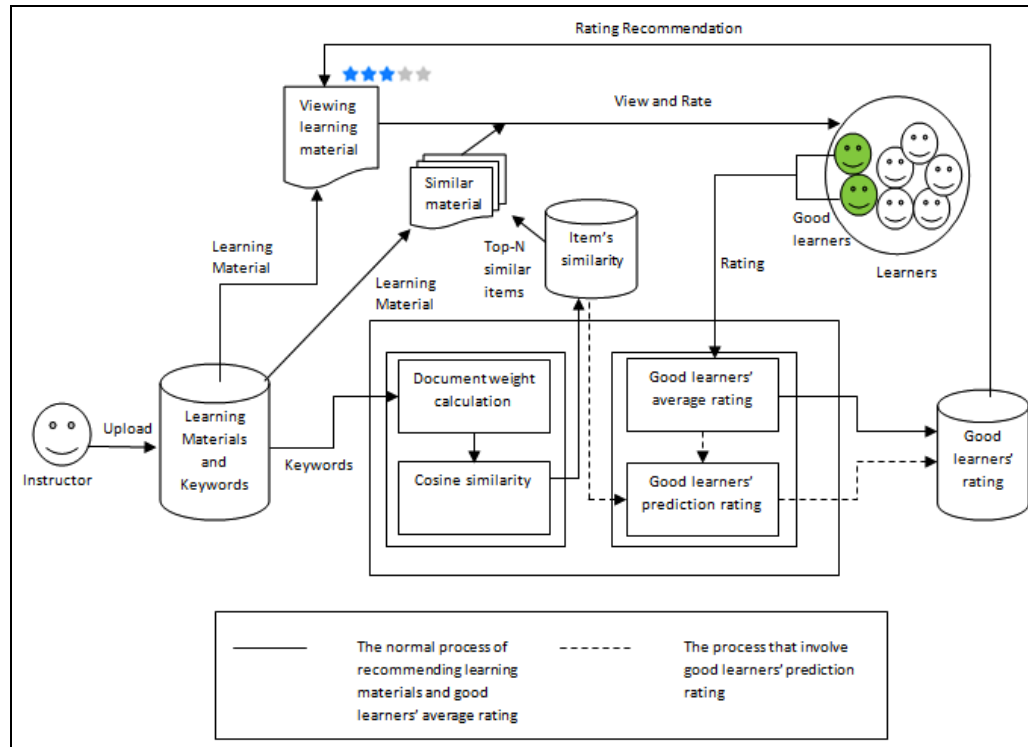

Figure 2: The good learners' recommendation strategies framework

The good learners' rating calculation starts by gathering the initial rating from the good learners (refer to 'Procedure' subsection on how the initial ratings are gathered for the purpose of the experiment). This initial rating is important to avoid the cold-start problem whereby the recommendation cannot be produced, due to insufficient ratings or absence of ratings that is usually faced by collaborative filtering technique (Adomavicius & Tuzhilin, 2005). If the ratings exist for a particular item, the good learners' average rating will be calculated by dividing the sum of all good learners' ratings by the number of good learners that have rated that particular item. The good learners' average rating are then stored in the rating database and will be used for rating recommendation and for the calculation of good learners' prediction rating. The mathematical equation to calculate the good learners' average rating is given as follows.

$$R_{i,j} = \frac{\sum_{i=1}^{N} r_{i,j}}{N_j} \tag{1}$$

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 based solely on good learners' ratings.

The good learners' rating prediction will be calculated when the good learners' ratings do not exist for a particular item. To calculate the prediction rating, the recommendation engine will retrieve both the item's similarity and its corresponding good learners' average rating, and divide the product between them with the sum of the items' similarities. The prediction ratings are then stored in the rating database. The formal calculation is shown as follows.

$$P_i = \sum_{n=1}^{N} \frac{sim(d_i, d_n) * R_n}{sim(d_i, d_n)} \tag{2}$$

where $sim(d_i, d_n)$ is the similarity between item *i* and item *n* and R_n is the good learners' average rating on item *n*. Note that once the document has received ratings from good learners, the prediction rating will be overwritten with the good learners' average rating.

The final stage in the recommendation process is to recommend the good learners' rating for the viewing item and to recommend top-N (i.e. items with the highest similarity) similar items. For this purpose, the recommendation engine will query the item's similarity from the item similarity database and based on the item's similarities (i.e. that exceed a threshold value), the top-N documents will be retrieved from the database. Concurrently, the good learners' rating for the viewing item is retrieved from the rating database to be displayed to the learners.

The sample screen shot of the working system is shown in Figure 3, in which the good learners' rating is shown at the top of the viewing item. The rating indicates the good learners' opinion about the item. The good learners' ratings were also being used to sort the similar items which are placed at the bottom of the viewing item. As we mentioned earlier, the number of similar items are determined by a threshold in which the top-N similar items must exceed the threshold value before the items are sorted out as to the top most rated items by the good learners.

Method

Participants

The participants were divided into 5 groups according to the tutorial sections that the students had registered in. Group 1 (G1) consisted of 21 students who used the e-learning without any recommender system, Group 2 (G2) consisted of 21 students who used the e-learning with a content-based recommender system, Group 3 (G3) consisted of 29 students who used the e-learning with collaborative filtering recommender system, Group 4 (G4) consisted of 26 students who used the e-learning with hybrid filtering recommender system, and Group 5 (G5) consisted of 24 students who used

the e-learning with the newly proposed recommender system. The difference between the systems used by G4 and G5 is that the hybrid recommender system used by G4 used all the users' ratings to produce recommendation, while the recommender system used by G5 only used the good learners' ratings in producing recommendation. All of the participants are second year students of software engineering.
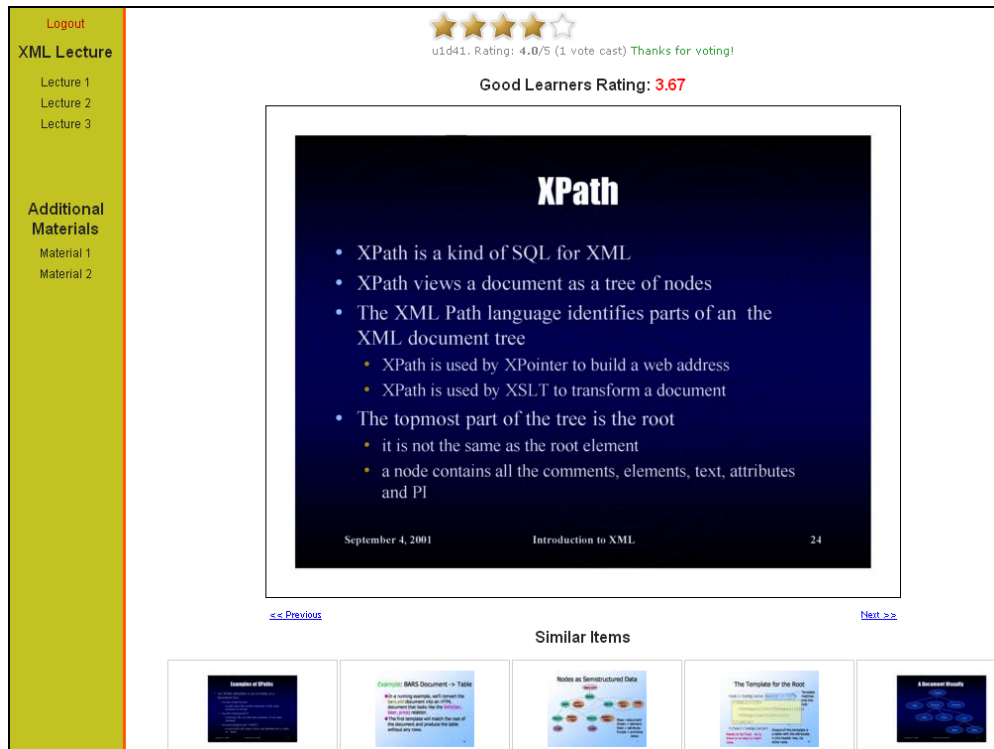

Figure 3: A screenshot of the e-learning recommender
system using good learners' ratings

Procedure

Students were asked to participate in this study as part of the requirement for the 'Web Services' course. The course requires students to have knowledge of XML and this study is used to measure the students' pre-knowledge and the knowledge gained after self-learning. Students were given one week after the pre-test had been conducted to complete the learning of a selected XML chapter. The learning materials comprised 5 sets of *PowerPoint* slides, in which the slides are converted into images and embedded into a website. During the process of learning, the students were encouraged to provide ratings for the learning materials, based on their usefulness.

Since the recommender system used by G5 required the good learners' ratings, the experiment on G5 was conducted after the completion of experiments on G1 and G2. The good learners' ratings from G1 and G2 were then used as the input ratings for the recommender system used by G5. As the time frame was different when conducting the experiments, we have taken precautions to avoid cheating and collaboration among the students. Firstly, the pre-test and post-test were conducted in a monitored

lab, and the web pages were set to disable the save function. Use of removable hard disks and thumb drives were not allowed during the tests, and access to the tests (i.e. pre-test and post-test) website was only via a local area network (LAN), and the server was shut downe after the tests were completed. This was to ensure that none of the questions were accessible and viewable by other groups.

Pre-test and post-test

The pre-test consisted of ten multiple choice questions. Basic questions about XML such as the definition of a well-formed and a valid XML document were asked during the pre-test. Twenty minutes were given to the students to answer the pre-test questions. The post-test consisted of fifteen multiple choice questions. The post-test questions covered some advanced knowledge in XML whereby the students needed to understand the concept of XML very well in order to be able to answer the questions. Among the questions which were asked during the post-test were some about spotting the syntax error in DTD and schema, and determining the namespace for prefix in an XML document. Students were given thirty minutes to complete the post-test.

Results and analysis

We measured the learning outcome by calculating the mean score obtained from the pre-test and the post-test for each group, and compared the mean score between the groups to check for the significance of the difference. For the pre-test, we used a two-tailed test since we assumed that there is no significant difference among the groups. In contrast, we used a one-tailed test when determining the significance of the difference for the post test as we assumed that students who used the e-learning with the proposed recommender system would have a better post-test score compared to other groups.

Table 1 and Table 2 summarise the scores obtained by groups of learners who used different types of recommender systems during the learning process. The pre-test marks show that there were no significant differences at $p < 0.05$ between marks obtained by all the groups.

Table 1: The mean score and standard deviation obtained from pre-test and post-test

Group	Pre-test			Post-test		
	N	M	SD	N	M	SD
G1	21	40.48	12.44	21	59.05	15.68
G2	21	35.71	15.02	21	58.41	14.13
G3	29	34.48	13.78	29	50.11	15.80
G4	26	37.31	11.51	26	58.21	13.93
G5	24	36.67	15.79	24	67.22	14.96

In contrast, there were significant differences when we compared the post-test marks between the groups. Obviously, G5 has obtained the highest post-test mark and the difference between the post-test mark obtained by G5 and other groups was significant (G1-G5: $t(43) = 1.79$, $p = 0.04$; G2-G5: $t(43) = 2.02$, $p = 0.02$; G3-G5: $t(51) = 4.02$, $p = 0.0001$; G4-G5: $t(48) = 2.21$, $p = 0.02$) at $p < 0.05$ with effect sizes greater than 0.5. The results also revealed that the post-test mark obtained by G3 (G3 obtained the lowest mark in the post-test) has a significant difference (G1-G3: $t(48) = 1.98$, $p = 0.03$; G2-G3: $t(48) = 1.92$, $p = 0.03$; G3-G4: $t(53) = 2.01$, $p = 0.02$; G3-G5: $t(51) = 4.02$, $p = 0.0001$) at $p < 0.05$ when compared to other groups.

Table 2: The mean comparison between groups of learners

Group	Pre-test (two-tailed test)				Post-test (one-tailed test)			
	t	df	p	d	t	df	p	d
G1-G2	1.1208	40	0.2691	0.3544	0.1389	40	0.4451	0.0439
G1-G3	1.5818	48	0.1203	0.4566	1.981	48	0.0267	0.5719
G1-G4	0.9055	45	0.3700	0.2670	0.1943	45	0.4234	0.0579
G1-G5	0.8898	43	0.3785	0.2714	1.7872	43	0.0405	0.5451
G2-G3	0.3	48	0.7655	0.0866	1.915	48	0.0308	0.5528
G2-G4	0.4136	45	0.6811	0.1233	0.0486	45	0.4808	0.0145
G2-G5	0.2081	43	0.8361	0.0635	2.0222	43	0.0247	0.6168
G3-G4	0.8212	53	0.4152	0.8212	2.0065	53	0.0250	0.5512
G3-G5	0.5391	51	0.5922	0.1510	4.0192	51	0.0001	1.1256
G4-G5	0.1647	48	0.8699	0.0475	2.2054	48	0.0161	0.6366

Besides comparing the mean, we also measured the percentage gained from pre-test to post-test for each group to determine the performance of the students when they used the assigned e-learning system.
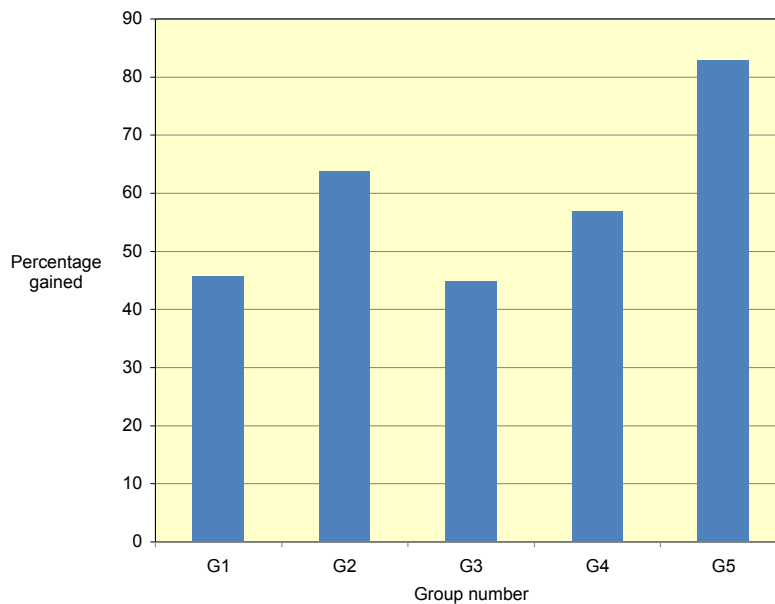

Figure 4: The percentage of mark gained from pre-test to post-test

As Figure 4 depicts, G5 has the highest percentage of mark gain from pre-test to post-test of about 83%. In contrast, G3 and G1 have the lowest percentage of mark gain of about 45%.

Conclusion

Recommender systems are widely used in online systems including e-learning systems, but their benefits to learners are still being debated. This study provides empirical evidence which clearly demonstrates the value of user's ratings as a collaboration tool in helping other learners by suggesting suitable items. The study compares the learning outcome of several groups of students who used different types

of e-learning recommender systems. The findings indicate that the incorporation of good learners' ratings in the content-based recommender system has a significantly positive impact on the learning outcome of the students by at least 13.8%. They outperform other groups of students who used several other types of e-learning recommender systems.

This study has shown that there are clear benefits in using the proposed recommender system in an online learning system. However, in order to maximise the benefits, more research is needed, through which the effectiveness of the proposed method can be further determined. First, the proposed recommender system relies on content-based filtering to recommend similar items, thus the accuracy of recommendations depends on the keywords assigned to each item. A poor choice of keywords may lead to poor recommendations to similar items (Adomavicius & Tuzhilin, 2005). In this study, the keywords were assigned manually by the instructor to each item since the number of learning materials used was relatively small. An automatic keyword extraction can be used when the number of items is large (Ercan & Cicekli, 2007). However, the recommended items may differ between the case where human assigned keywords are used and the case where automatically extracted keywords are used, and as the similarities between items are heavily dependent on the number of the matched keywords, this may affect the recommendation accuracy.

Another important factor that has a direct impact on the recommendation of the proposed recommender system is the amount of good learners' ratings. Our proposed recommender system is prone to the 'cold start' problem, in which the system is not able to calculate or predict the good learners' rating for the items if the good learners' ratings are unavailable. Some researchers have suggested the use of hybrid filtering to overcome the 'cold start' problem (Lekakos & Giaglis, 2007), hence the incorporation of good learners' ratings with hybrid filtering technique needs further research. Finally, it is also important to study the range of contexts in which recommender systems may be relevant, as that will help to ascertain whether the usage of recommender systems in e-learning can be extended to other subject fields and whether it is suitable for formal or informal learning (Drachsler, Hummel & Koper, 2009).

Acknowledgments

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.

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.

Cecez-Kecmanovic, D. & Webb, C. (2000). Towards a communicative model of collaborative web-mediated learning. *Australian Journal of Educational Technology*, 16(1), 73-85. http://www.ascilite.org.au/ajet/ajet16/cecez-kecmanovic.html

Chen, C. M., Lee, H. M. & Chen, Y. H. (2005). Personalized e-learning system using item repository theory. *Computers & Education*, 44(3), 237-255.

Drachsler, H., Hummel, H. & Koper, R. (2009). Identifying the goal, user model and conditions of recommender systems for formal and informal learning. *Journal of Digital Information*, 10(2), 1-17. http://journals.tdl.org/jodi/article/view/442/279

Ercan, G. & Cicekli, I. (2007). Using lexical chains for keyword extraction. *Information Processing & Management*, 43(6), 1705-1714.

Felfernig, A., Friedrich, G. & Schmidt-Thieme, L. (2007). Recommender systems. *IEEE Intelligent Systems*, 22(3), 18-21.

Ghauth, K. I. & Abdullah, N. A. (2010). Learning materials recommendation using good learners' ratings and content-based filtering. *Educational Technology Research and Development*. DOI: 10.1007/s11423-010-9155-4

Ghauth, K. I. & Abdullah, N. A. (2009). Building an e-learning recommender system using vector space model and good learners average rating. *Ninth IEEE International Conference on Advanced Learning Technologies (ICALT 2009)*, Latvia, 194-196.

Kerkiri, T., Manitsaris, A. & Mavridou, A. (2007). Reputation metadata for recommending personalized e-learning resources. *Proceedings of the Second International Workshop on Semantic Media Adaptation and Personalization*, Uxbridge, 110-115.

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. http://www.ifets.info/journals/12_4/4.pdf

Lekakos, G. & Giaglis, G. M. (2007). A hybrid approach for improving predictive accuracy of collaborative filtering algorithms. *User Modeling and User-Adapted Interaction*, 17(1-2), 5-40.

Liang, G., Weining, K. & Junzhou, L. (2006). Courseware recommendation in e-learning system. *Advances in Web Based Learning – ICWL2006*, Springer Berlin/Heidelberg, 10-24.

Liu, F. & Shih, B. (2007). Learning activity-based e-learning material recommendation system. *Proceedings of the Ninth IEEE International Symposium on Multimedia Workshops (ISMW'07)*, Beijing, China, 343-348.

Nachmias, R. & Segev, L. (2003). Students' use of content in web-supported academic courses. *The Internet and Higher Education*, 6(2), 145-157.

Otair, M. A. & Hamad, A. Q. A. (2005). Expert personalized e-learning recommender system. (EPERS). In *Proceedings of The First International Conference on E-Business and E-Learning EBEL 2005*. Amman, Jordan.

Soonthornphisaj, N., Rojsattarat, E. & Yim-ngam, S. (2006). Smart e-learning using recommender system. *Computational Intelligence*, Springer-Verlag, Berlin, Heidelberg, 518-523.

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. [verified 12 Oct 2010] http://www.emeraldinsight.com/journals.htm?issn=0264-0473&volume=26&issue=3&articleid=1728644&show=pdf

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.

Zaiane, O. R. (2002). Building a recommender agent for e-learning systems. *Proceedings of the International Conference on Computers in Education (ICCE'02)*, Auckland, New Zealand, 55-59.

Khairil Imran Ghauth, Faculty of Information Technology
Multimedia University, 63100 Cyberjaya, Malaysia.
Email: khairil-imran@mmu.edu.my

Nor Aniza Abdullah, Faculty of Computer Science and Information Technology
University of Malaya, 50603 Kuala Lumpur, Malaysia.
Email: noraniza@um.edu.my