AN INTERACTION-BASED COLLABORATIVE FILTERING APPROACH FOR PERSONAL LEARNING ENVIRONMENT

Syed Mubarak Ali, Imran Ghani
Software Engineering Department, Faculty of Computing, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia.
mubarakali.00@gmail.com
Software Engineering Department, Faculty of Computing, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia.
imransaieen@gmail.com

ABSTRACT: *E-learning has become very famous in recent years. This is the modern era of internet and technology and a lot of information is available on the internet. E-learning helps the learner in accessing the best information available for learning purpose. In this bundle of endless information, not every material is useful for every learner. So, in order to filter the available contents, a process called recommendation is defined which help users in recommending the most appropriate material for learning. One of the main issues with recommender systems is called the new-user cold-start problem. When a new user enters in a system, there is very little information present about the user and this causes inaccuracy in recommendation for the user. In this paper, we have proposed an approach which increases the recommendation accuracy for new-user cold-start problem. The evaluation of the proposed approach shows the improvement by 6% from the existing approach.*

1. INTRODUCTION

In recent years, e-Learning has become very famous because this type of learning system uses up-to-date and latest educational technologies to implement such a learning environment where information technology is integrated in curriculum which helps in creating much more effective learning environment than traditional learning system [1]. In comparison with traditional 'face-to-face' learning, e-Learning has got much attention and is quite famous. However, e-Learning has a lot of difference when compared with other e-Activities because e-Learning has its own way of information retrieval, knowledge management and pedagogical process [2].

Personal Learning Environment (PLE) is becoming very popular in the line of Technology Enhanced Learning (TEL). "PLEs refer to a set of learning tools, services, and artifacts gathered from various contexts to be used by the learners" [3]. Furthermore, according to Van Harmelen [4] PLE's help learners to enjoy ICT-based environments for learning activities so that a learner can connect to different networks in order to collaborate on shared outcomes and acquire necessary (professional and rich professional) competences through using the PLE.

Because there is so much information available on the Internet, it can be quite difficult for a learner to find the most appropriate contents for learning. This problem arises specially when a user is new to the learning and he has very little personal experience. In learning activities of a learner, recommendations are considered a good technique because it helps the user to choose the materials for learning. Recommendations can be very useful in different aspects of Personal Learning Environment for example, for finding relevant tools, get recommendations for learner to interaction in specific situations [5].

Recommendation or recommender systems are very popular since the past few years. It is defined as "systems that produce individualized recommendations as output or have the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options" [6-7]. Recommendation systems are used in e-commerce and social networking sites commonly and now this system has gained a lot of attention in the e-Learning community as well. There are different types of recommendation techniques but the most popular techniques are 'Collaborative Filtering', 'Content-based Filtering' and 'Hybrid Filtering' [6, 8-9].

In this paper, we are going to propose an approach which is based on collaborative filtering. This approach has increase the recommendation accuracy for the new-user cold start problem specifically in the personal learning environment. The approach is discussed in detail in section 3. In the next section, we have presented some extensive literature review related to PLE, recommender systems and the current problems regarding recommendation.

2. LITERATURE REVIEW

2.1 e-Learning

e-Learning has become an integral part of learning environment these days. This is the era of information technology and social media and there is lots of information available on the internet. e-Learning environment simply just makes these information available to the learner through the use of web technologies like internet, social networks, blog, wiki's, etc. E-Learning is different from other learning systems because of its difference in involvement of information retrieval, knowledge management and pedagogical process [10-11].

As the technology advances, there arises the need for personalization in the e-Learning environment. Because there is so much information freely available on the internet, not every learner can read these vast amounts of information or learning content. So, in order to keep the learners interest on the acquired goals, personalization needs to be done. In order to fulfill this requirement of personalization, an environment came to existence known as Personal Learning Environment (PLE).

2.2 Personal Learning Environment (PLE)

PLE can be described as the set of tools, services and communities that establish an individual educational platform for learners to use for their own direct learning and accomplish their educational goals and targets [12]. Learning Management Systems (LMS) are frequently compared with PLE but they are different due to the fact that LMS are course-centric whereas PLE's are learner-centric. It is not necessary for a PLE to intersect with an institutional LMS but an individual learner may integrate different components of LMS in the environments they created for themselves. For example, in a typical PLE, different blogs may be integrated where comments upon their learning style or post information related to their learning and this might reflect the information drawn from internet, sites like YouTube. As it been argued above, the PLE is not a specific software or application but it is more like an ideology about how learners achieve their task of learning. Although online environments are a major focus of PLE, this is not the limitation here [13].

As discussed above, there is lot of information available on website, social networks, blog, wiki's, etc. Now, the learner is only interested in information which is most relevant to him. If a leaner get lots of information, in which there is little information of use for the learner, the learner will lose interest. Looking at another scenario, we have come to learn that sometimes the learner's have very little or no personal experience regarding learning. So, in this case, recommendation comes to aid the learner by filtering the available information, customize and recommend only the relevant information according to the features and needs of the learner. 'Good' recommendations are a vital and important task in a successful PLE. 'Bad' leads to the failure of a PLE and learner's trust whereas 'Good' recommendation ensure the user's trust in the PLE [14].

2.3 Recommender Systems

Recommender systems are a type of information filtering system that gives advice on products, information, or services that a user may be interested in. They assist users with the decision making process when choosing items with multiple alternatives [15]. All recommender systems, whether they are in the e-commerce world or any other application, all have three things in common: they all take in inputs; they all have a goal; and they all produce and output [16]. The basic function of recommender systems is to provide recommendations to the users, based on items or information related to their interests and in some cases, to provide guesses or ratings to each item which the user may prefer [15].

2.3.1 Issues in Recommender Systems

After going through the literature, a few major challenges in the recommender systems were found. These problems pose serious challenge on the use and performance of recommendation systems. Few major problems are discussed below:

a)- Cold Start: This is very common problem for new user or new item. When a new user or new item is registered in a system, usually there is insufficient information and ratings for them and this affects the recommendation algorithm to predict them or recommend them. Every recommendation technique uses ratings or past history, in order to recommend effectively and efficiently. Now, if the item or user has not enough ratings or history, it is very hard for the algorithm to process them and this one of the major challenges of recommender systems. There are two types of cold-start problem; new-user cold-start problem and new-item cold-start problem [17-19].

b)- Data Sparsity: Since users may not rate some items, the user-item matrix may have many missing ratings and be very sparse. Therefore, finding correlations between users and items becomes quite difficult and can lead to weak recommendations. Many users don't rate every item they like, so it is not necessary that if a user has not rate an item means that the user did not like the item. This is big issue in recommender systems [18, 20-21].

c)- When there are result sets which are too similar to one another, this can lead the users to lose interest due to the reduced diversity of results. The user needs diverse results which are relevant to them but are also dissimilar [22].

2.4 Types of Recommender Systems

There are different types of recommender systems like collaborative filtering recommender systems, content-based filtering recommender systems, hybrid filtering recommender systems, knowledge-based filtering recommender systems, demographic-based filtering recommender systems and utility-based filtering recommender systems [6, 8-9, 23]. Here we are going to discuss collaborative filtering recommender system according to the scope of this paper.

2.4.1 Collaborative Filtering Recommender System

Collaborative filtering recommender system [9, 17, 20, 24-27] uses collaborative filtering approach to provide the user with recommendations based on what other users with comparable interests or preferences might have liked in the past. It leverages the similarity between users to make recommendation because their preferences are sometimes correlated.

Collaborative Filtering systems are often classified as "memory-based collaborative filtering" and "model based collaborative filtering". In the early research phase, researchers' used a memory-based approach to predict items. In memory-based approach the items are predicted based on the past ratings by users on an entire collection of item. Due to the limitations with the memory-based approach, model-based approach came into existence. In model-based approach, a model of created which is then trained by giving a collection of ratings and this model is then used for prediction. Although this approach overcomes some of the limitations of memory-based approach, it also has its own limitations.

So in conclusion, we can say that in this modern PLE can play a very crucial role in the field of e-learning. As discussed above PLE use different technologies and integrate them to provide a complete platform for the learner. So recommendation comes in handy for the selection of most relevant information for a particular user. But recommender systems come with their own limitations and advantages, as discussed above. So in this research, we have presented an approach to improve the accuracy of recommendation for new-user cold-start problem. In the next section we are going to present our approach and an in depth discussion about the approach.

3. PROPOSED APPROACH

In this section, we have discussed our proposed approach. Fig. 1 explains the working of our proposed approach.

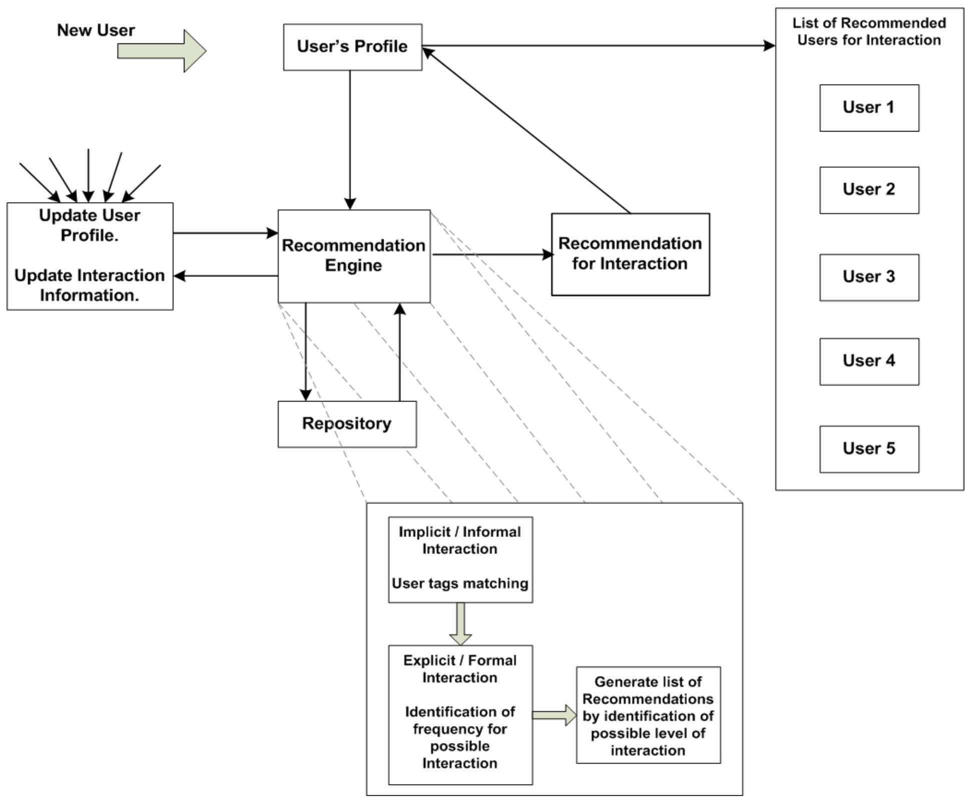

Fig. 1. Interaction-based Collaborative Filtering Approach

The main objective of this research work is to improve the accuracy of recommendation for new-user cold-start problem. So, this approach focuses on new-user when the new user enters in the system. As the initial step, when the user creates a profile, the system shows some tags to the user and user choose few tags that matches best to the user criteria. The profile of that user is then saved in the repository and with the help of these tags the recommendation is generated for the user.

As shown in Figure 1, when the user profile is completed, this profile is then sent to the recommendation engine. In this recommendation engine, the tags are weighted and after few calculations, a list of the recommended users is then given to the new-user and this new information is then updated in the user profile. There is a repository where all the files and preferences are stored. There is one important component whose work is to continuously update the user profile and interaction details. With the help of this component, the accuracy of recommendations can increase drastically. The more history, ratings and interaction details are there, the more system can generate accurate results.

The recommendation engine has further two main processes involves in it. The first process is called implicit or informal interaction and second is called explicit or formal interaction. In the implicit interaction, a list of recommended users is generated by calculating the tags similarity between the new user and the existing users. This can be done by the using the method called 'Levenshtein Distance' [28]. This list is then passed to the explicit interaction process. In this process, the list which is passed to this process is then filtered further to get a more accurate list. This filtration is done by identifying the frequency of tags that are common among the new user and the users in the generated list. So, the more frequency there is, the more this user is appropriate for the new user.

4. EXPERIMENTAL EVALUATION

In this section, we have discussed in detail about the dataset used in the experiment and the experimental setup. We also have discussed the results.

4.1 Introduction to the Dataset

The dataset used in the evaluation of the proposed approach is Movielens dataset which contains the cold start users (users with less than 20 votes). This dataset contains 10000054 ratings and 95580 tags applied to 10681 movies by 71567 users. The minimum rating is 0.5 and maximum rating is 5 with the step size of 0.5. The mean, median and standard deviation of ratings is 3.61, 3.72 and 1.06 respectively.

For our experiments, we need users and tags information. Since cold-start users (users with less than 20 votes) are not accounted in Movielens dataset, so in order to perform experiments related to cold-start users, we have removed votes from the dataset. Indeed, we have removed randomly between 5 and 20 votes of those users who have rated between 20 and 30 items. Similarly, the tags for these users have been removed as well. So now those users who have rated between 2 and 20 items and have few tags selected are referred to as cold-start users. A screenshot of the data tables which are used in this experiment can be seen in Figure 2.

Table 1. Main parameters of the Movielens dataset

Number of Users	71567
Number of Movies	10681
Number of Tags	95580
Number of Ratings	10000054
Minimum Ratings	0.5
Maximum Ratings	5

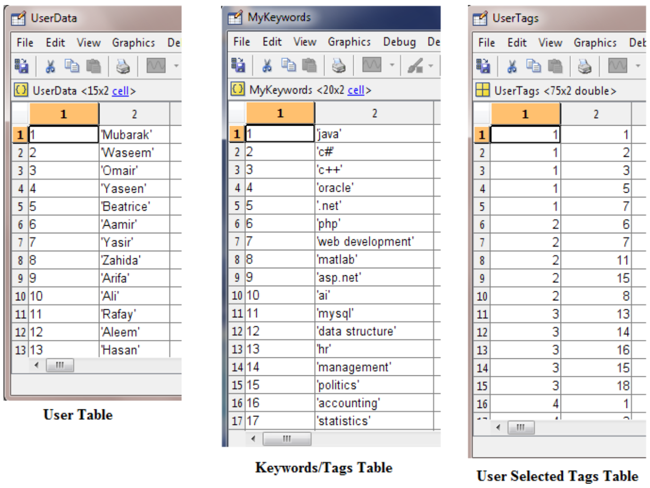

Fig. 2. Data Tables Structure

4.2 Experimental Setup

The experiment has two parts. In the first part, when a new user enters in the system, a list of different keywords to select 5 tags. These tags are saved in the user profile and this profile is then used to generate first list of recommendation for the user.

In the second part, we assume that the user has updated the profile by visiting other links, other users' profile and some other articles. All these information is saved in the form of tags and keyword. This profile is then used to generate the most suitable list of 5 users. Step by step description is given below:

Step 1 – A new-user is entered in the system. A list of few tags is shown to user and the user selects few tags from that list. These selected tags are then saved in the user profile. This profile is constantly updating. Whenever user interacts with the system for example, when user clicks on some article or visit some other users' profile.

The input is a new user, the user selects few tags from a list of tags which is data and the output is the profile which is saved for every user.

Step 2 – Now the first phase of recommendation is done through informal or explicit interaction. This can be achieved with the help of tags that user has selected, when the user is entered in the system. By matching the tags with other users' tags using Levenshtein Distance [28] algorithm. This algorithm calculates the difference between two strings. By using this algorithm, we search 50 users which are best related to current user. A detailed explanation of Levenshtein Distance algorithm is given in table 2.

The input for this step is the user profile. From the profile, the system retrieves all the tags and start comparing the tags with other users' selected tags, so the profile is the data for this step. The output is the list of 50 users recommended for the new user.

Step 3 – For the second phase of recommendation, namely explicit or formal interaction, we assume that user has selected some tags or selected some users and this information is saved in the profile. Now, the profile of user contains more than 5 tags and keywords which are most relevant to user. We retrieve all the tags and keywords from the users' profile and once again we run these tags through Levenshtein Distance algorithm to match all the tags to other users' tags.

The input for this step is the user profile. From this profile, again all tags information is retrieved and once again computed with other users profile in order to find the closest match for the new user. This is the data for this step. A list of 5 users is then recommended to the new user, which are the most appropriate candidate for the interaction. This information is updated in the profile as output.

Step 4 – Finally, the system recommends a list of five users to the current user, based on the tags similarity among them. This information is again saved in the profile. As the user profile stores more tags and keywords related to the user, the more accurate the recommendation gets.

The input is the profile and the list of 5 users is retrieved from the profile which is the data. The output is the list of users, which is then recommended to the user, which is the output of this step.

A pictorial description of experiment is given in Fig. 2.

Table 2. Levenshtein Distance Algorithm

- Set *n* to be the length of *x*, and *m* to be the length of *y*
- Create a matrix with *m* rows and *n* columns and initialize the first row and column to *0…m* and *0…n* respectively.
- Examine each of the characters of *x* and *y* to 1 to *n* and 1 to *m*.
- If *x[i] = y[i]*, the characters are equal and the transformation cost is *0*. If *x[i] != y[i]*, the characters are not equal and the transformation cost is *1*.
- The value of cell *d[i, j]* is set to the minimum of {*d[i-1, j] + 1 (the cell above + 1)*}, {*d[i, j-1] + 1 (the cell to the left + 1)*}, or {*d[i-1, j] + cost (the cell diagonally above and to the left)*}.
- Step 3-5 is repeated until the distance score is found in cell *d[n, m]*.

4.3 Results and Discussion

In this section, we have discussed about the results of the evaluation of our proposed approach. Movielens 10M dataset is used as a dataset to evaluation our approach. Our approach corroborates the good results for new-user cold start issue.

The graph of precision accuracy of our approach can be seen in Figure 3. The experiment is designed to run 10 times and gives the accuracy each time. We calculate precision on based of that accuracy and plot the graph of all times. It shows an improved precision for our proposed approach, the graph of the comparison is given in Figure 5.

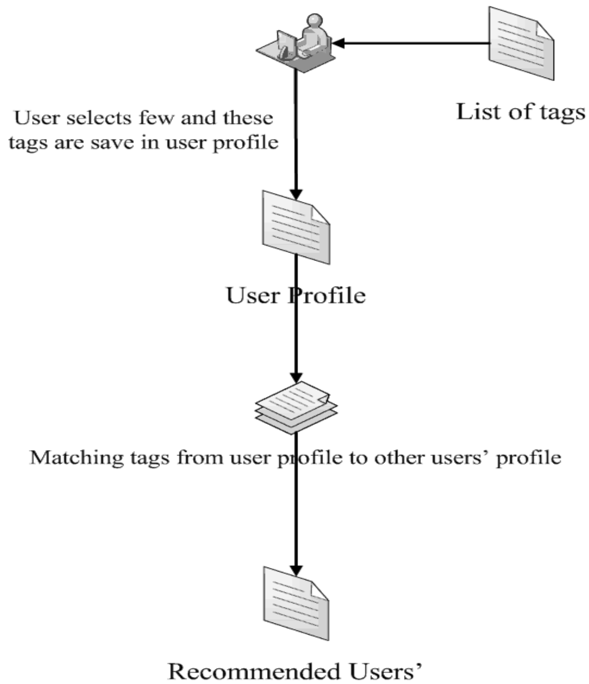

Fig. 2. Description of flow of experiment

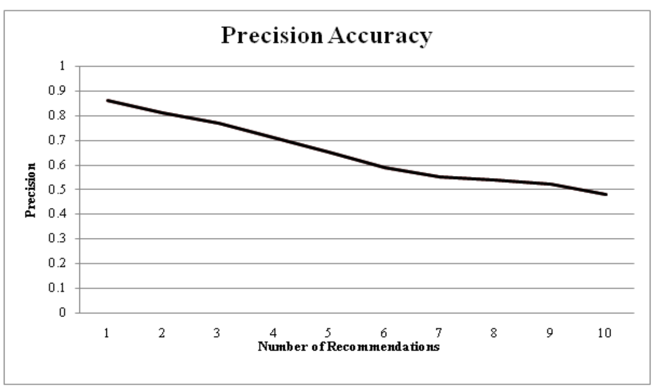

Fig. 3. Precision Accuracy

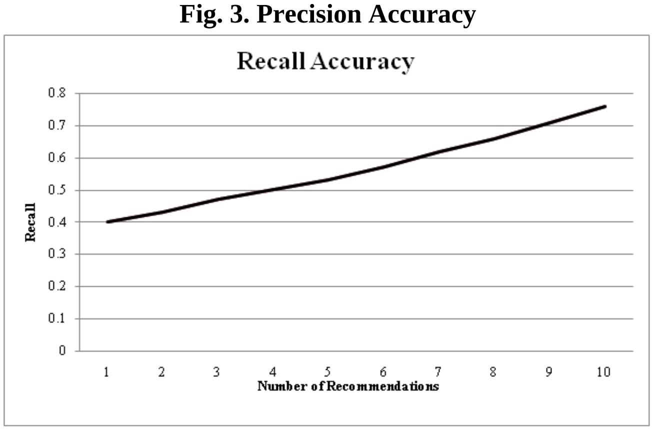

Fig. 4. Recall Accuracy

Similarly, we took the recall of our proposed approach which is shown in Figure 4. Each time the experiment runs and we get an accuracy value, we took recall and then plot the graph of all ten results we get. The recall has shown slight improvement.

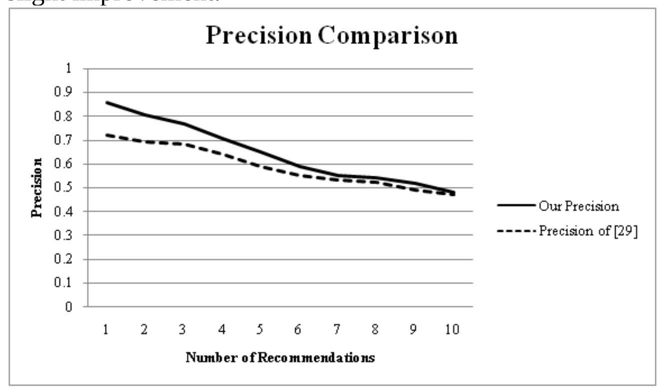

Fig. 5. Precision Comparison with [29]

The Figure 5 shows the comparison of precision of our approach and [29]. Our approach shows improvement of about 14% in 1st recommendation, 12% in 2nd recommendation, 9% in 3rd recommendation and so on. By taking average, we calculated overall 8% improvement when compared with [29].

5. CONCLUSION

E-learning has become an integral part in learning for students in this era of modern of technology and internet. As the technology advances, so does the problem. A very common problem in this regard is the problem of new-user cold start. It is the problem in which when a new user enters in a system, the system has to recommend different contents to user. But because the user is new and there is not much information about the student, therefore the system fails to recommend most appropriate content to the user. In this research paper, we proposed an approach to improve the recommendation accuracy for new–user cold start problem. We have discussed about our approach and the experimental evaluation in detail in sections 3 and 4 respectively. The result shows the improvement in the accuracy of recommendation for the new-user cold start problem.

REFERENCES

[1] S. Duo and L. X. Song, "An E-Learning System based on Affective Computing," in *International Conference on Applied Physics and Industrial Engineering*, 2012, pp. 1893-1898.

[2] B. Wu, *et al.*, "Experience Effect in E-Learning Research," *Physics Procedia*, vol. 24, pp. 2067-2074, 2012.

[3] F. Henri, *et al.*, "Understanding PLE as an Essential Component of the Learning Process," in *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2008, pp. 3766-3770.

[4] M. V. Harmelen, "Design Trajectories: Four Experiments in PLE Implementation," *Interactive Learning Environment,* vol. 16, pp. 35-46, 2008.

[5] F. Modritscher, "Towards a Recommender Strategy for Personal Learning Environments," in *Proceedings of 1st Workshop on Recommender Systems for Technology Enhanced Learning*, pp. 2775-2782. 2010.

[6] R. Burke, "Hybrid Recommender Systems: Survey and Experiments," *User Modeling and User-Adapted Interaction,* vol. 12, pp. 331-370, 2002.

[7] I. Ghani, *et al.*, "A Role-Oriented Content-based Filtering Approach: Personalized Enterprise Architecture Management Perspective," *International Journal of Computer Science and Information Security,* vol. 8, 2010.

[8] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the Start-of-the-Art and Possible Extensions," *IEEE Transactions on Knowledge and Data Engineering,* vol. 17, 2005.

[9] X. Zhou, *et al.*, "The State-of-the-Art In Personalized Recommender Systems for Social Networking," *Artificial Intelligence Review,* vol. 37, pp. 119-132, 2012.

[10] S. Duo and L. X. Song, "An E-learning System based on Affective Computing," *Physics Procedia,* vol. 24, pp. 1893-1898, 2012.

[11] E. Y. Huang, *et al.*, "What type of learning style leads to online participation in the mixed-mode e-learning environment? A study of software usage instruction," *Computers & Education,* vol. 58, pp. 338-349, 2012.

[12] M. A. Chatti, *et al.*, "LaaN: Convergence of Knowledge Management and Technology-Enhanced Learning," *Learning Technologies, IEEE Transactions on,* vol. 5, pp. 177-189, 2012.

[13] T. Valtonen, *et al.*, "Perspectives on personal learning environments held by vocational students," *Computers & Education,* vol. 58, pp. 732-739, 2012.

[14] U. Kirschenmann, *et al.*, "Demands of modern PLEs and the ROLE approach," *Sustaining TEL: From Innovation to Learning and Practice,* pp. 167-182, 2010.

[15] A. Gunawardana and G. Shani, "A Survey of Accuracy Evaluation Metrics of Recommendation Tasks," *Journal of Machine Learning Research,* vol. 10, pp. 2935-2962, 2009.

[16] E. Vozalis and K. G. Margaritis, "Analysis of recommender systems algorithms," in *Proceedings of the 6th Hellenic European Conference on Computer Mathematics and its Applications (HERCMA-2003), Athens, Greece*, 2003.

[17] J. Bobadilla, *et al.*, "A collaborative filtering approach to mitigate the new user cold start problem," *Knowledge-Based Systems,* vol. 26, pp. 225-238, 2012.

[18] C. Birtolo, *et al.*, "Personalized Suggestions by means of Collaborative Filtering: A Comparison of two different Model-based Techniques," in *In Proceedings of 3rd World Congress on Nature and Biologically Inspired Computing NaBIC '11*, 2011, pp. 444-450.

[19] N. Koenigstein, *et al.*, "Yahoo! Music Recommendations: Modeling Music Ratings with Temporal Dynamics and Item Taxonomy," in *Proceedings of the fifth ACM Conference on Recommender systems RecSys '11*, 2011, pp. 165-172.

[20] J. L. Herlocker, *et al.*, "Evaluating Collaborative Filtering Recommender Systems," *ACM Transactions on Information Systems (TOIS),* vol. 22, pp. 5-53, 2004.

[21] M. López-Nores, *et al.*, "Property-based collaborative filtering for health-aware recommender systems," *Expert Systems with Applications,* 2012.

[22] C. Yu, *et al.*, "Recommendation diversification using explanations," in *Data Engineering, 2009. ICDE'09. IEEE 25th International Conference on*, 2009, pp. 1299-1302.

[23] Y. Jiao and G. Cao, "A Collaborative Tagging System for Personalized Recommendation in B2C Electronic Commerce," presented at the International Conference on Wireless Communications, Networking and Mobile Computing, 2007, 2007.

[24] J. Bobadilla, *et al.*, "A collaborative filtering similarity measure based on singularities," *Information Processing & Management,* vol. 48, pp. 204-217, 2012.

[25] J. B. Schafer, *et al.*, "Collaborative Filtering Recommender Systems," *Lecture Notes in Computer Science,* vol. 4321, pp. 291-324, 2007.

[26] N. Barbieri and G. Manco, "An Analysis of Probabilistic Methods for Top-N Recommendation in Collaborative Filtering," in *Proceedings of the 2011 European Conference on Machine Learning and Knowledge Discovery in Databases ECML PKDD '11*, 2011, pp. 172-187.

[27] H. N. Kim, *et al.*, "Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation," *Electronic Commerce Research and Applications,* vol. 9, pp. 73-83, 2010.

[28] L. Yujian and L. Bo, "A normalized Levenshtein distance metric," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 29, pp. 1091-1095, 2007.

[29] C. Wang, *et al.*, "Tag Recommendation Based on Collaborative Filtering and Text Similarity," *International Journal of Education and Management Engineering (IJEME),* vol. 2, p. 7, 2012.