International Journal of Academic Research in Computer Sciences and Electrical Engineering
January 2015, Vol. 2, No. 1

The Improvement of Electronic Learning's Recommender Systems Performance

Mahnaz Najafi
Master of Computer Engineering, Software Major, Iran

Fereshte Khoshnam
Master of Computer Engineering, Software Major, Iran

Sepide Jahangiri
Master of Computer Engineering, Software Major, Iran

DOI: 10.6007/IJARCSEE/v2-i1/1524 URL: http://dx.doi.org/10.6007/IJARCSEE/v2-i1/1524

Abstract
Today, the virtual environment is becoming more and more widespread as far as the control and processing of information is almost impossible. Therefore, the need for a system that can overcome this matter is felt more than ever. The systems that suggest the best and most friendly cases from among huge numbers of different products and data, according to the specific characteristics of each user, are very popular. The recommender systems are intelligent systems in the internet which identify the interests and preferences of users and offer relevant information to them. This study aimed to introduce recommender systems, analyze their techniques in detail, study the role of these systems in e-Learning to improve the learning of users in a virtual learning environment, and enhance the quality of recommendations with involvement of users' information level in cooperative filtration algorithm to enhance the quality of users' learning.

Keywords: E-Learning, User Profiles, Cooperative Filtration, Recommender Systems.

Introduction
The electronic world is moving towards saturation of information. During the past decade, large volumes of data are stored on information servers and databases. Today, the amount of available data doubles every five years. Considering large variety of information, the access to suitable data seems necessary for proper decision making. In fact, along with increased options, the volume of information that must be processed to achieve the target and the amount of time and energy to reach the final data increases dramatically. In such environments, the systems with the ability to identify and update the interests and priorities of users and with the ability to index and store the information in a searchable manner are needed to predict and identify users requirements and direct them to find appropriate materials. To date, several methods have been proposed to solve the problem of accumulation of information. One of these methods is the use of search engines. But so

far, these motors have not been very successful in customizing search results and often offer same results for all users; while the users may have very different profiles and may consider different aspects of research results. Another practical method is the use of recommender systems. The recommender systems are the tools that may direct the users in electronic environments to find interested information, goods, and services. With the ability to recognize and predict the priorities of users, the recommender systems screen the information that is interested by user from among large volume of data and offer them to users to save their time and energy. On the other hand, these systems with the ability to analyze the past behavior of users also include the services and information which the users did not pay attention to them and offer the results to users.
The recommender system in e-learning environment is software that tries to recommend actions to learners based on previous learners actions. This recommendation may be an offline activity such as conducting an exercise, reading messages in a conference system, or running an online simulation.
It can be said that e-Learning recommender is different from other recommenders in many aspects, because:
The favorite items of users for learning may not be appropriate for them. For example, a user with no previous background in web mining techniques may only be interested in knowing the newest web mining techniques in E-commerce. However, customization is not only about selecting learning items, but it is also about their transfer and training methods.

1- The introduction of recommender systems and their types

In the simplest sense, the recommender systems are the intelligent information filtering systems in internet which try to identify user interests and priorities and recommend the information which is probably the user's favorite. An information filtering system is a system which removes the additional or unfavorable information before displaying them to user using semi-automated or fully automated methods. The main objective of these systems is information accumulation management. In other words, these systems try to guess the users' way of thinking and decrease their difficulty in understanding and decision-making which may be caused by too much information.
The recommender systems manage the information to provide customized content to users according to their interests. The efforts to provide a customized view of information to users have an effective role in encouraging people to choose high quality items (Ricci,2010).
The successful examples are the recommender systems of Google News, Amazon, and Netflix. The reflection of recommender systems' success can clearly be found in e-commerce.
According to the used knowledge source for their information, the recommendation techniques are divided into several types. In some systems, the knowledge of other users or the sociological information is added to database using human knowledge engineering. Thus, the following four techniques are created:

2.1: Content-based systems:
In this method, the items are obtained which the user has given them an acceptable rating. Then, the cases are searched which are similar to ones are rated by users and the most similar items are recommended. In general, in this method, the items are offered to users which have similarities with interested items of user. In fact, the focus is on finding similarities between items (Asanov,2011). Thus, this method requires two types of information: user profiles and items content.

2-2: Cooperative filtration systems:
Since the content-based methods were not always able to find the desired documents and items of users and providing keywords may describe the content of documents well, they had poor performance in understanding the use of these keywords or quality of these documents. The filtration process or evaluation of items using the opinion of others is called cooperative filtration. The technology collects the opinion of many related communities in the web and filters the data inherently (Baraglia,2004).
The collaborative filtration makes automated the process of recommendations, taking into account the existence of similar beliefs among users. The cooperative recommendations algorithms are divided into two main memory-based and model-based categories (Khoshgoftaar ,2009)

Memory-based algorithms:
Each user is part of a group that is similar to him/her. By defining neighbors for active users, their unknown favorite items may be recommended to them. To define the neighbors for active users, the similarities are first calculated with weights sim i, j. The weight may be the distance or correlation between two users or two items i and j. Then, the ranking of unknown items is estimated based on the average of all users' ratings weight. Finally, N best recommendations are given to users. The Pearson correlation coefficient, cosine distance, and adjusted cosine distance are used to calculate the similarities. The distance may be between two users or two items. Accordingly, two types of memory-based algorithms are formed: algorithms based on user (user-user) and algorithms based on item (item-item). The similarity calculation should be conducted for providing recommendations; then, the unknown rankings should be estimated (Khoshgoftaar ,2009).

The similarity calculation:

The Pearson correlation coefficient is a method for calculation of similarities (Sarwar,2001 ;Yang,2000):

$$sim_{i,j} = \frac{\sum_{m\in(i\cap j)} (r_{i,m} - \overline{r_i})(r_{j,m} - \overline{r_j})}{\sqrt{\sum_{m\in(i\cap j)} (r_{i,m} - \overline{r_i})^2} \sqrt{\sum_{m\in(i\cap j)} (r_{j,m} - \overline{r_j})^2}} \quad (1)$$

Another method for calculation of similarity is cosine distance (Sarwar,2001 ;Yang,2000):

$$sim_{i,j} = \frac{\sum_{m\in(i\cap j)} r_{i,m} r_{j,m}}{\sqrt{\sum_{m\in(i\cap j)} r_{i,m}{}^2 r_{j,m}{}^2}} \quad (2)$$

And the adjusted cosine method is as follows (Sarwar,2001 ; Yang,2000):

$$sim_{i,j} = \frac{\sum_{m\in(i\cap j)} (r_{i,m} - \overline{r_i})(r_{j,m} - \overline{r_j})}{\sqrt{\sum_{m\in(i\cap j)} (r_{i,m} - \overline{r_i})^2 \sum_{m\in(i\cap j)} (r_{j,m} - \overline{r_j})^2}} \quad (3)$$

In the formulas (1) and (2) and (3), the sim i, j is the distance between two items or two users i and j. If i and j are two users, then the iΩj is the set of all items that the users i and j have rated them.
The ri and rj is the average of user i and user j ratings on the same items of iΩj. The ri is the ranking of user i for item m and rj is the ranking of user j for item m. If i and j are both items, then iΩj is the set of all users who have rated items i, j. The r i and rj is the ranking average of items i and j on iΩj users. The ri, m is the ranking of user m for item i and rj, m is the ranking of user m for item j.

Estimation of recommendation:
Following similarity calculation, the users or items that are most similar should be used to estimate the recommendation. The formula (4) is used to make a recommendation for user u in item i in user-user method (Khoshgoftaar, 2009;Sarwar,2001).

$$prediction_{u,i} = \frac{\sum_{n\in neighbors} (r_{n,i} - \overline{r_n}) sim_{u,n}}{\sum_{n\in neighbors} |sim_{u,n}|} + \overline{r_u} \quad (4)$$

In this formula, the neighbors in recommendation system are the set of closest users to user u. The obtained distance between user u and the rest of users are ordered in a list. Then, a set of users with minimum distance are selected as neighbor in recommendation system. The $\overline{r_u}, \overline{r_n}$ is the average of rankings for users u and n in all items except item i.

The recommendation systems in e-Learning and the conducted in this area:

Due to the development of new communication technologies, many key changes have been occurred in various aspects of human life; however, the education is not an exception. In recent years, there have been huge advances in educational systems and new technologies have been introduced such as web-based training. Today, many people benefit from e-learning applications. However, the large variety of learners in the web have created new challenges for traditional learning model "one for all" in which a single set of learning resources is provided for all users.

The users may have different interests. Even with common interests, they may have different skill levels. Therefore, they should not all be treated uniformly (Yang, 2000).

The recommender system in e-learning environment is software that tries to recommend actions to learners based on previous learners actions. This recommendation may be an offline activity such as conducting an exercise, reading messages in a conference system, or running an online simulation (Loll, 2009).

There are two main parts in the design of such system:

A learning module that is trained by past access patterns and extracts a shared or individual access model.

The consultation module which uses learned model for user's recommended actions.

There are many methods to implement this process such as data clustering, associated rule mining, cooperative filtration, and etc. (Sikka, 2012).

Today, students often encounter a lot of learning online materials. Unlike the spent time for learning of this material, the students are tempted to devote more time for search and filtering. Therefore, they will obtain the information that better meet their needs in terms of value or scientific priorities. The limited time of learning may prevent from effective placement of learning concepts. This is important considering the fact that many of them may end up irrelevant concepts. One of the possible ways to overcome this problem is the use of recommendation systems.

The recommendation system is the software that supports users in identifying the most interesting items. The famous approaches which are used in recommendation systems are cooperative filtration, content- based filtration, and hybrid filtration. The cooperative filtration identifies interesting items according to the opinions of other users using the calculation of nearest neighbors (for example, top-N users which have identical ranking pattern) from a ranking matrix. The new items are those that have most benefit to nearest neighbors. In contrast, the content-based filtration uses the properties of items for extracting the recommendations. Therefore, the items with similar content to user's current evaluation items will be recommended to active user. The hybrid filtration is mixed with content-based filtration and cooperative filtration methods to provide a recommendation. The recommender systems in e-learning may have various types depending on the type of recommended concept such as registration period, learning concepts, the importance of learning concept, and etc. (Ghauth, 2010).

3.2 The proposed method:

One of the main ideas in the field of recommendation systems which are based on equality between users is the possibility of users' access to services and the consideration of recommendations' characteristics. A normal recommendation system provides recommendations for users based on the factors supplied by them.

The similar behavior among users is enough based on equal performance between users and it is consistent with the majority of recommendation systems. For example, there is no reason to believe that a user is more qualified to comment on movies, travel, blogs, and etc. However, this is important in e-learning environment.

In cooperative filtration model which is suggested in the e-learning recommender system, we work from two-dimensional matrix of ratios R in users U for these items and add a two-dimensional matrix C for scores of users U and T-tests. However, each user's scientific level in this education system will be determined by grading items and the scores in t-test which shows their awareness of training materials. The T-score of level tests may be high and includes the scores of automated corrected tests, the scores of passed subjects, passed trainings, and etc. However, each user's scientific level in this education system will be determined by grading items and the scores in t-test which shows their awareness of training materials.

The first proposed method:

A number of metrics may be created to assess the importance of knowledge of user x in recommendations which is received from him/her. In this study, the use of simple and symmetrical metric is considered which can be created using function f.

$$f = \begin{cases} c_x - c_y & , c_x > c_y \\ 0 & c_x \leq c_y \end{cases} \qquad (5)$$

In the equation (5), the Cx is the knowledge of user x and Cy is the knowledge of user y. The new scale of similarity between users x and y which is important for us may be defined as equation (6). The first part of the equation refers to the importance of grades. However, the second part considers similarities between users based on their proportions using one of the similarity comparison methods such as Pearson, cosine coefficient, adjusted cosine, and some traditional metrics.

$$imp(x,y) = [\frac{1}{T}\sum_{t=1}^{T} f\,(cxt, cyt)] * sim(x,y) \qquad (6)$$

The second proposed method:

In the second proposed method, according to equation (7), we used the distance between the scores of two users for different courses.

$$sim_{(i,j)} = \frac{\sum_{t=1}^{T} \frac{\max_t - \min_t}{G_{i_t} - G_{j_t}}}{\sum_{t=1}^{T} \max_t - \min_t} \quad (7)$$

In equation (7), T is the number of courses, maxt is the highest score in lesson t, min t is the lowest score in lesson t, and Git and Gjt are the scores of users i and j in lesson t.

Conclusion

In general, it can be said that one new equation was defined to measure the similarity between users. It increases the importance of recommendations which are created by users depending on their level of knowledge. To include the knowledge of users in cooperative filtration, it is necessary to develop new metrics which coordinate other information about each user's scores based on the current metrics. The confirmation of new metrics requires a change in traditional methods which are used to evaluate the overall error of system when these metrics are used. The error mean, error absolute mean, error square mean, and other overall accuracy evaluations of system should indicate the user's knowledge level. Although there are many experiments have been conducted on e-learning database, the designed equations and used methods can be used in a similar fashion in different electronic recommender systems.

References:

Ricci, F., Rokach, L., Shapira, B. (2010). "Recommender SystemsHandbook",Springer New York Dordrecht Heidelberg London.

Asanov, D. (2011). “Algorithms and Methods in Recommender Systems” , Berlin Institute of Technology , Berlin , Germany.

Baraglia, R., Silvestri, F. (2004). “An online recommender system for large web sites”, in: Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence , pp. 199–205 ,

Su, X., Khoshgoftaar, T. (2009). "A Survey of Collaborative Filtering Techniques" Advances in Artificial Intelligence , vol. 2 , pp. 1-20.

Sarwar, B., Karypis, G. and Konstan, K. (2001) “Itembased collaborative filtering recommendation algorithms” , WWW '01 Proceedings of the 10thinternational conference on World Wide Web

Yang, Qu, Y. and Huang, X. (2000) “Survey of Recommendation System and Algorithms”. EE 380L: DataMining.

Sikka, R., Dhankha, R. (2012) ”A Servey Paper on E-Learning Recommender System”, International Journal of Computer applications , vol.47 , No.9 , pp. 888-975.