Available at: www.sabauni.net
Journal home page: www.sjitn.sabauni.net

S.J.I.T.N
Saba Journal of Information Technology and Networking

Recommendations Generating Using Multi Attribute Decision Making

Hael Al-bashiri* , Amer Al-badarnehand,Ismail Hmeidi

Jordan University of Science and Technology (JUST), Department of Computer Science

Article info

Article history :
Accepted March, 2014

Keywords:
Decision Making

Abstract

In the last few years, the amount of information that is available on the Internet to the users grows rapidly. Due to this, the searching process for information or certain items and decision-making in most cases became difficult and very complex. The recommendation system technique is very important to help users how to deal with information overload. In this paper, we present collaborative filtering and Multi Attribute Decision Making (MADM) method that increases the accuracy of the system. The (MADM) technique is implemented to generate the recommendations. This technique is based on all ratings of most similar users that lead to improve the accuracy of traditional collaborative filtering approach.

The experimental evaluation has shown that the proposed recommendation system outperforms the traditional system in terms of accuracy by (12-15) %.

* Corresponding author:Hael Al-bashiri
E-mail address:hailealbashiri@gmail.com

Introduction

In our daily life, we make our decisions in most cases based on recommendations of people, newspapers, or the Internet (e.g., book reviews, movie critics, restaurant rating). However, as the amount of information that is available on the Internet grows very fast, So searching and making decisions about information becomes difficult. We need new technologies to help Internet users to how deal with information overload.

Recommendation systems can assist Internet users how to deal with this problem. These systems provide what you need according to what you chose at a previous time. The main objective of the recommender systems is to provide tools that help users to control the information search and gather actions of other users.

The recommender systems have important application areas that focus on considerable recent academic and commercial interests. They are widely used by many commercials and nonprofit web sites to help users to choose items based on users' preferences. These systems assist to overcome the problem of increasing information overload by providing customers with suggestions or recommendations based on their likes (good ratings) and dislikes (bad ratings) relative to the other customers. Big purchasing websites like eBay [18] and Amazon [17] represent some of the businesses that have essential recommendations into their shopping experience. Recommendations have become an integral aspect of these e-commerce platforms and are used to personalize the shopping experience [3].

A recommender system is a computer-based system that gives advice on items, services, or information based on pre-collected data such as past users' activities. These items/services are not yet accessed or purchased and may be users' interests. Recommendation list is generated based on analyzing historical information of other users' interests.

Recommendation systems are implemented by creating profile of a target user and comparing it with closest users' profiles which are stored in the database. The recommendation systems can be generally classified as [4]:

Collaborative Filtering (CF): recommendation systems: propose items to a target user depending on information about similarities among other users' preferences .*Content-based recommendation systems*: items are recommended to a target user based on similarity between their content and content of items which user has rated in the past time. *Hybrid recommendation systems*: These systems combine between both collaborative and content-based approaches in order to improve system accuracy and performance by avoiding limitations of them [5].

In collaborative filtering instead of using the items content itself, the items ratings are used to generate recommendations. One of the well-known collaborative filtering techniques is the k-nearest-neighbor (KNN).KNN compare this preference history with a preference of other users in order to find the K most similar users. Similarity calculation is based on the rating of items [6].

This paper is organized as follows. Introduction about the recommendation systems and problem space will be explained in section 2. Relevant recommendation approach and overview of collaborative recommendation systems will be discussed in section 3.The methodology and our approach will be presented in section 4. Experimental results are discussing in section 6. This paper will be concludes with section 7.

Problem Space

The essential problem in information filtering is calculating whether a certain item is likely to interest a user or not. The outcome of such a computation is either Boolean, yes or no, or a score that corresponds the scale to which a person may wish that item. Such a score helps to determine if an item can be suggested to a target user or not.

In this paper, we will explore the space of neighborhood-based and explain the MADM method that we have used.

Collaborative filtering using MADM method can be consisted of six steps.

1. Represent User-Item matrix.
2. Compute the similarities between all users and target user.
3. Select *n* users that have the highest similarity to create a neighborhood.
4. Construct Decision matrix using all neighbors' items (candidate items) with neighbors' ratings.
5. Apply TOPSIS method on decision matrix to normalize ratings and ranking all candidate items.
6. Chose a set of ranked items as recommendations.

Overview Of Collaborative Recommendation Systems

Many researchers in recommendation system area focused on rating constructions. Specifically, the estimate ratings for an item, which is unrated from user, based on past rating of the target user and other similarity users' rating. Many techniques are developed to predict ratings for items that cannot be accessed or purchased from user. And recommend set of them to target user according to the highest predicted rating, which is the most common and preferred approaches.

Collaborative recommendation systems recommend items to target user depending on the similarity between the current user preferences and other similar users. The collaborative filtering solves most of the problems that are found in the content-based approach. The recommendation of a collaborative filtering system depends on the similarity of users' preferences rather than similarity of items' content [13]. Close users are grouped together using some methods, such as k-nearest neighborhood [14]. The collaborative filtering approach suggests a set of items, which are liked to other items that the user's group prefers.

However, this approach has some shortcomings: it is difficult to give a good recommendation to users who have evolving preferences or strange tastes. The second problem appears when the information is not sufficient to find similarities between users' preferences. Another problem is if a new user or item entered into database can lead into a weak recommendation [9], this is known as cold star problem.

There are a lot of researches that can be classified as collaborative filtering system such as the system in [13]. Many collaborative recommendation systems have been developed and proposed in the literature. Some of these systems use correlation-based model. Some other algorithms employ a Bayesian network model [8][19], while others use association rules [10, 11].

User-Based Recommendation: Estimated predictions in this approach are divided into two steps. The first, computes the similarities among the target user and all other users using the Pearson Correlation Coefficient [12, 7, and 8].

$$W(u,v) = \frac{\sum_{a\in I}(R_{u,i}-R'_u)(R_{v,i}-R'_v)}{\sqrt{\sum_{a\in I}(R_{u,i}-R'_u)^2 \sum_{a\in I}(R_{v,i}-R'_v)^2}} \qquad 1$$

Where *I* denoted to all items that rated by both *v* compared users and *u* target user, *R(u,a)* is the voting rated by user u on item *i*. and R'_u is the average rating for user *u*.

Then, in the second step *k* closest users to user *u* are taken in order to calculate the prediction rating for target user *u* on item *i* using equation 2. This formula shows how a prediction *P(u,i)* is computed. Where *P(u,i)*represents the predicted view for target user u about item *i*. *R(v,i)*represents the rating of item *i* by user *v*. R'_v is the average rating for user *v*.

$$P_{u,i} = R'_u + \frac{\sum_{i=1}^{k}(R_{v,i}-R'_v)W(u,v)}{\sum_{i=1}^{k} W(u,v)} \qquad 2$$

Item-based KNN: the computing item-item similarities are used to calculate predictions [15]. The equation 3 is used to compute similarities between two items:

$$w(i,j) \frac{\sum_{u\in U}(R_{u,i}-R'_u)(R_{u,j}-R'_u)}{\sqrt{\sum_{u\in U}(R_{u,i}-R'_u)^2 \sum_{u\in U}(R_{u,j}-R'_u)^2}} \qquad 3$$

Where, *U* represents all users who have rated both *i* and *j* items. Then apply the equation 4 to generate the prediction *P(u,i)*. This approach can use threshold for *k* similar items in here rather than all.

$$P_{u,i} = \frac{\sum_{j=1\ to\ K}\left(w(i,j)*(R_{u,j})\right)}{\sum_{j=1\ to\ K}(|w(i,j)|)} \qquad 4$$

Where *k*=1,2,..*l* and *l* represent all items which are taken from the neighborhood.

Methodology

We apply MADM to recommend a set of movies that are important for the active user. The main steps of our work are:

- *Data Pre-processing:* is the important step in the data mining process in order to prepare data for another processing task to make data more easily and effectively processed. Data's gathering methods are often loosely controlled, resulting in out-of-range values, missing values, etc. This leads to produce misleading results after performed an analysis on dataset. Thus, the quality of data is important before running an analysis.
- *Representation Data Matrix*: The users and items can be set as a collection of numerical ratings into a user-item matrix.
- *Neighborhood Formation*: This step is the most important in the recommendation process. Neighborhood formation required computing the similarity between target user and other users within the user-item matrix. Similarity will be utilized to produce a recommendation for a target user.

  Neighborhood formation is performed in follows steps:

  i. Compare the similarity between all users with the target user within the user-item matrix.
  ii. Take K users that have the highest similarity to create a neighborhood.
- *Recommendation Generation:* In our system we employ multi attribute decision making approach in order to provide recommendations to the target user. MADM method describes any type of measurements that use a set of criteria to rank a set of alternatives. The output of MADM method is a set of alternative ranked that helps the user to whichever of these options is better than other.

The technique that is used in solving decision-making problems in our work is TOPSIS method (Technique for Order Preference by Similarity to the Ideal Solution) [2].

The TOPSIS technique is composed of the following steps:

Step 1: Construct normalized decision matrix R.

In this step, we transform various attribute dimensions into non-dimensional attributes, which allows comparisons across criteria. The normalized value of the decision matrix can be any linear-scale transformation to make value of rating between zero and one.

Normalizing score ratings is as follows:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum x_{ij}^2}} \qquad 5$$

Where i=1… m and j= 1… n.

Step 2: Build the weighted normalized decision matrix.

We construct the weighted normalized decision matrix by multiplying each column of the normalized decision matrix by its related weight. The weight is determined directly using ranking method. In this method, the criteria are simply ranked in perceived order of importance by decision-makers: $c_1 > c_2 > c_3 > \ldots > c_i$. The weights are non-negative.

The elements of weighted normalized decision matrix are computed using equation 6.

$$v_{ij} = w_j * r_{ij} \qquad 6$$

Step 3: Determine the positive ideal solution A* and negative ideal solutions A'.

- Positive ideal solution:

$A^* = \{ v_1^*, \ldots, v_n^* \}$, where $v_j^* = \{ \max (v_{ij})$ if $j \in J$; $\min (v_{ij})$ if $j \in J' \}$.

- Negative ideal solution:

$A' = \{ v_1', \ldots, v_n' \}$, where $v' = \{ \min (v_{ij})$ if $j \in J$; $\max (v_{ij})$ if $j \in J' \}$.

Step 4: Calculate the separation measures from the positive and the negative ideal solutions.

The separation measure from positive ideal solutions is:

$$S_i^* = \left[\Sigma \left(v_j^* - v_{ij} \right)^2 \right]^{½} \qquad 7$$

Where i=1… m.

Similarly, the separation measure from the negative ideal alternative is:

$$S'_i = \left[\Sigma \left(v_j' - v_{ij} \right)^2 \right]^{\frac{1}{2}} \qquad 8$$

Where i=1… m.

Step 5: Calculate the relative closeness to the ideal solution C_i*.

$$S'_i = \left[\Sigma \left(v_j' - v_{ij} \right)^2 \right]^{\frac{1}{2}} \qquad 9$$

Where $0 < C_i^* < 1$.

Step 6: Rank the alternative based on C_i^*,

Datasets

In our experiments the Movie Lens dataset that contains 100,000 movie ratings from 943 users on 1682 movies is used. The releasing time spans from 1922 to 1998 and the user's rating scale ranges from one star (in the worst case) and five stars (in the best case). Each movie has been rated at least once, and each user has rated at least twenty movies.

Users and items are numbered consecutively from one. The density of this version is around 6%. Here, the density is calculated as the fraction of ratings over the possible number of ratings[16].

In this work, we used three splitting data cases:

- First case: divided data into 80% training data and 20% testing data.
- Second case: divided data into 90% training data and 10% test data.
- Third case: we take ten items from each user as a test.

Experimental Results

In our experiments, recommended items are classified as interest to the target user or not

Fig. 1 shows AME value for each of the previous splitting cases. In Figure 1, we note that the AME values in the first and second cases are relatively closed, but the AME value in the third case is worst one due to that interest. User interest items are the items that rated by this user.

To evaluate the system, we used the recall measurement. If all the recommended items do not exist in the test set then the recall value is 0.0, which indicates that the accuracy of the system is very weak, if system finds all users' interests then the value of recall is 1.0, which indicates that the system has a great accuracy.

Our evaluation strategy based on comparing our recommendation system that applies the MADM approach to generate recommendations for a target user to a traditional recommendation system that uses predicted approach to generate recommendations.

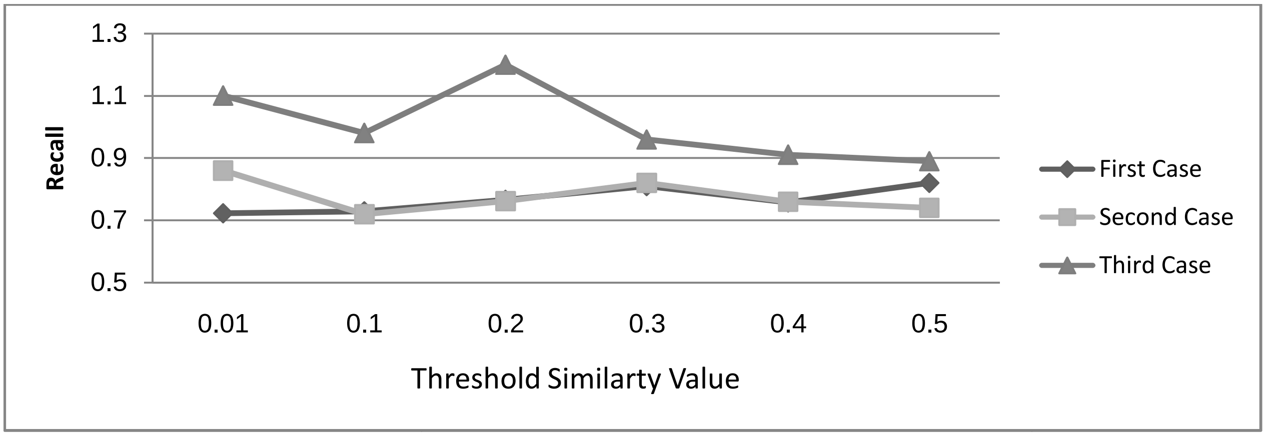

Fig. 1. AME for Different Splitting Dataset with Several Threshold Similarity Value.

the number of tested items for each user smaller than two previous cases. Also the AME value at threshold similarity value = .1 is better than AME value at threshold similarity value = .01, .2, .3, .4 and .5 in the all cases.

Thus, the Best Threshold Similarity Value (BTSV) is .1.

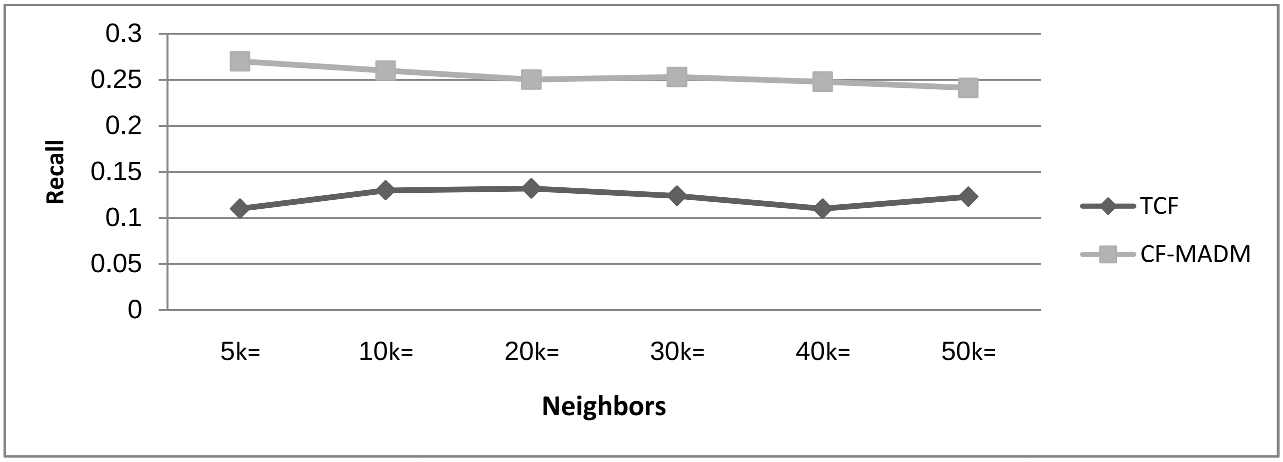

Fig. 2. Recall for CF-MADM and TCF.

Fig. 2 shows a comparison between the proposed method and traditional method in term of recall. Experiment conducted on the first case. The number of users is varied between 150, 100, and 50 randomly selected users. For each number of users we varied number of retrieved items between 50 and 100 retrieved items. After we done all the experiments we compute the recall average value for all cases and depicted it in the Fig. 2.

The results show that our approach performs better than the traditional one by approximately 12%. This improvement is due to that our approach is based on all ratings from most similar users for each candidate item. Unlike traditional approach that depends on prediction rate that used only the weight between target and similar

Fig. 3 shows a comparison between the first, second and third cases in term of recall. The number of users is 100, randomly selected users. For each number of users we varied number of retrieved items between 50 and 100 retrieved items. After we done all the experiments we compute the recall average value for all cases and depicted it in the Figure 3.The recall is affected by the number of items in testing data. Thus, you can note that the third case recall is better than the two previous cases recall due to that the number of tested items for each user is smaller than two previous cases user and ratings between them.

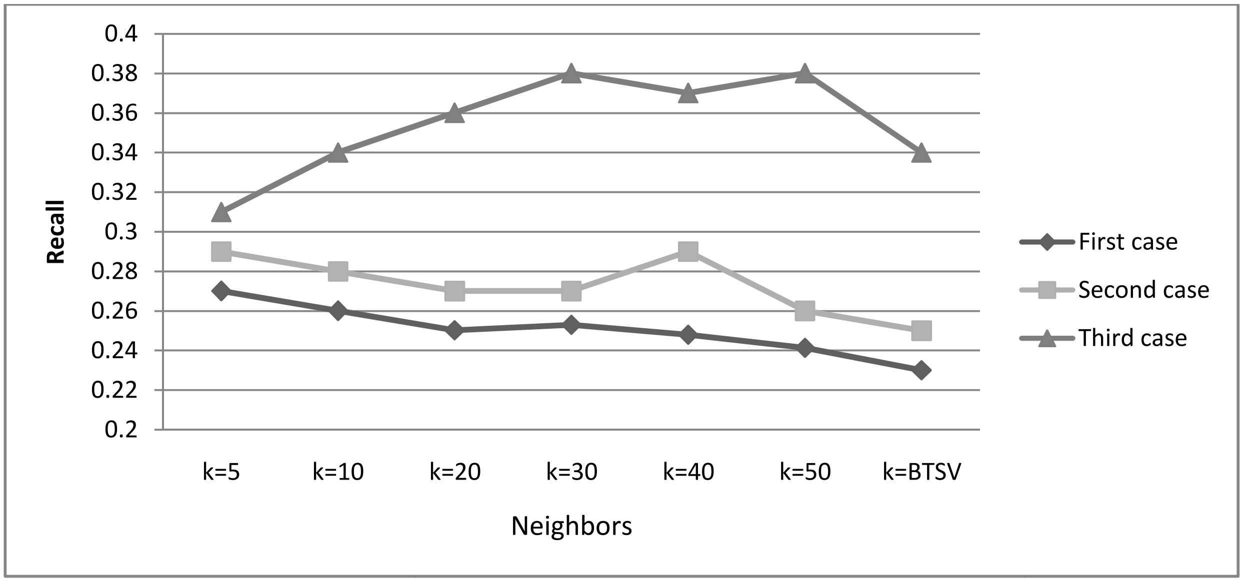

Fig. 3. Recall for First, Second and Third Cases.

Conclusions

This paper proposes a new generation method for collaborative filtering algorithm. Our new approach is based on all neighbors' ratings that lead to accuracy improvement. The experimental study shows that this approach has a better accuracy measurement (12%-15%) in generating recommendations using MADM method compared to generating recommendations using prediction method.

In the future work there are many possible directions, which include the following:

- Applying our experiments using different and larger dataset.
- Applying MADM method with other similarity methods and comparing between them.
- Applying MADM method with several similar methods to rank the most similar users.

References

[1]. Y. J. Lai, C. L. Hwang, Fuzzy multiple objective decision making methods and applications, New York: Springer-Verlag, 1994.

[2]. C. L. Hwang, K. Yoon, Multiple attributes decision making methods and applications, Springer: Berlin Heidelberg, 1981.

[3]. Sarwar, B. M., Karypis, G., Konstan, J. A., and Riedl, J. Application of Dimensionality Reduction in Recommender System–A Case Study. In Proceedings Of The AcmWebkdd 2000 Web Mining For E-Commerce Workshop, 2000b.

[4]. Vozalis, E, K Margaritis. Analysis of Recommender Systems' Algorithms. HERCMA. 2003; 732-745.

[5]. T. Y. Tang, P. Winoto, K. C. C. Chan. On the temporal analysis for improved hybrid recommendations. Web Intelligence, Proceedings. IEEE/WIC International Conference on, 2003.

[6]. Christakou, C., L. Lefakis, et al. A Movie Recommender System Based on Semi-supervised Clustering. omputational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on.

[7]. P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl. GroupLens: An open architecture for collaborative littering of netnews. In CSCW '94: Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, 175-186, Chapel Hill, North Carolina, United States, 1994, ACM Press.

[8]. Billsus D, Pazzani MJ. Learning collaborative information filters. Proceeding of the 15th International Conference on Machine Learning, Madison, Wisconsin, 1998.

[9]. Zan Huang, Hsinchun Chen, Daniel D. Zeng. Applying Associative Retrieval Techniques to Alleviate the Sparsity Problem in Collaborative Filtering. ACM Transaction on Information Systems (TOIS). 2004; 22(1): 54-88.

[10]. Chamgchien S, Lu T. Mining association rules procedures to support on-line recommendation by customers and products fragmentation. Expert systems with applications. 2001: 20, 325-335.

[11]. Liu D, Shih Y. Integrating AHP and data mining for procedures recommendation based on customer lifetime value. Information and management. 2005; 42, 387-400.

[12]. J. Herlocker, J. Konstan, A. Borchers, J. Riedl. An algorithmic framework for performing collaborative lettering. In Proceedings of the 1999 Conference on Research and Development in Information Retrieval (SIGIR-99). 1999.

[13]. M. Balabanovic, Y. Shoham. Fab. Content-based, collaborative recommendation. Communications of the ACM. 997; 40(3):66–72.

[14]. Teng SH. Yao FF. K-nearest neighbor clustering and percolation theory. In Workshop on Algorithms and Computation. Hong-Kong 1993; 192-211.

[15]. B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In WWW '01: Proceedings of the 10th International Conference on World Wide Web, 285-295, Hong Kong, 2001, ACM Press.

[16]. Collaborative recommendation system [accessed 2012 May]. Available from URL http://www.grouplens.org/.

[17]. Recommendation system [accessed 2012 June]. Available from URLhttp://www.amazon.com/.

[18]. Recommendation system [accessed 2012 June]. Available from URLhttp://www.ebay.com/.

[19]. Breese JS, Heckerman D, Kadie C. Empirical analysis of predictive algorithms for collaborative filtering. Proceeding of 14th Conference on Uncertainty in Artificial Intelligence; 1998 July; Madison, WI.