*Journal of Software Engineering and Applications*, **2011, 4, 411-416**
doi:10.4236/jsea.2011.47047 Published Online July 2011 (http://www.SciRP.org/journal/jsea)

# Proposing a New Metric for Collaborative Filtering

**Arash Bahrehmand[1*], Reza Rafeh[2]**

[1]Department of Computer Engineering, Islamic Azad University, Arak, Iran; [2]Deputy of Computer Engineering, Arak University, Arak, Iran.
Email: *stdmwmrule@gmail,com; r-rafeh@araku.ac.ir

Received May 21st, 2011; revised June 29th, 2011; accepted June 30th, 2011.

## ABSTRACT

*The aim of a recommender system is filtering the enormous quantity of information to obtain useful information based on the user's interest. Collaborative filtering is a technique which improves the efficiency of recommendation systems by considering the similarity between users. The similarity is based on the given rating to data by similar users. However, user's interest may change over time. In this paper we propose an adaptive metric which considers the time in measuring the similarity of users. The experimental results show that our approach is more accurate than the traditional collaborative filtering algorithm.*

***Keywords***: *Recommendation Systems, Collaborative Filtering, Similarity Metric*

## 1. Introduction

In recent years, collaborative filtering has become a widely adopted algorithm in recommendation systems. "Collaborative" means that a group of people with the same interests define their preferences in order to set up the system. CF[1] is used to create "recommendation systems" that can, for instance, enhance your experience on a Web site by suggesting music or movies that users might like. This algorithm uses the similarity of other users who have the same history as target user to recommend some items that may be related to his personal interest. The history holds all feedbacks which can be achieved from users by the system. In fact, feedbacks simulate the behavior of users. By analyzing the behavior of users, the similarity values between target user and other members would be recognized and based on these similarities, the items that he has not purchased yet but his similar users have, will be recommended to him.

The better you can measure the similarity of users, the better result you achieve in predicting the user's behavior. In the field of social psychology there are many candidate aspects for similarity. But out of the infinite number of qualities that each person has, some characteristics are more important in determining the similarity of two persons. In general, the similarity depends on the system point of view on interactions of users. However, more knowledge about personalities of an individual helps us to make more confident decisions on how similar are humans. In recommendation systems it is also subject to the extent to which the system is able to capture. Most of the collaborative filtering algorithms have focused on rating information of users. They believe, ratings determine the level of agreement or disagreement of users for a specific item. But the user interest may change over the time. By increasing the number of interactions and interconnections of people, internet, the probability of transitory of interests, personal and social behavior is not far-fetched, therefore; maybe what people want now is likely to variation (is not what they have explored in the past).In general, people have various reactions to events happen in the society. The reflections of these reactions can be observed in e-commerce. Also, the interpretation of each action is related to its time. For instance, suppose a science fiction movie has added to the list of an online movie show site. Users, who are fan of these kinds of movies, would watch and rate it as soon as possible, but those who are not very interested in mentioned type of movie maybe watch it later or when there is no one left in the list which is fascinating from their point of view. Hence we can infer, this group of users is different from pervious users and one of the most important factors which help us to recognize this issue is measuring the

[1]Collaborative Filtering

level of time distances in interactions of users. An intelligent recommender system must be able to adapt itself with current requirements of users. On the other hand, the order of interactions can be a source of semantics which help us in simulating the mentally evolution process of users.

In this research we have presented a new similarity metric which combine different aspect of user behavior in order to better understand who are more similar to each other and who are different from other. Experiment results demonstrate that our contribution improves the accuracy of predictions and evaluate what approach is more effective to satisfy users.

The rest of the paper is organized as follows. In Section 2, we review previous research in the field of collaborative filtering and analyze most significant metrics which are applied in recommendation system. In Section 3, we propose a novel method to measure the similarity of users by considering different effective aspects such as time. In Section 4 the experimental results are shown to evaluate the proposed approach. And Section 5 concludes the paper.

## 2. Related Work

For a long time recommender systems did not deal with changing user's interest in their predictions. For example, similarity of users was constant in different time periods and training datasets were considered as a static resource. In memory-based collaborative filtering, predictions are based on preferences of neighbor users or items. A widespread approach in memory-based collaborative filtering is the k-nearest neighbor algorithm. In [1] two imputed neighborhood based collaborative filtering algorithms are proposed that improved the performance of recommender system in very sparse rating data. Another popular technique is the top-N recommendations which recommends a set of N top-sorted items which probably will be interesting items by a target user [2,3]. However hybrid approaches are recognized as a way to alleviate the scaling and sparsity problem without any optimized solution when the content of environment changes dynamically over the time.

In [4] the history period is divided into intervals with different length the similarity is computed in each interval separately. But the way of determining initial time interval and length of each time interval, considering the conditions of each system is not described.

A critical step in K-nearest algorithm is the neighborhood formation which means finding a special subset of the community for target user by recognizing others with analogous behavior to act as recommenders. In order to, every pair of users profile to be computed, to measure the degree of similarity, $Sim_{u,v}$, shared between all pairs a and b [5]. The core objective of similarity metrics is evaluating potential relationship between users with a numeric value.

Jaccard similarity has focused on the amount of overlap between users based on the amount of shared items in each pair. However, the value of the user's rate is neglected [6]. Equation (1) describes this metric in more details in which $R_{u,i}$ determines the rating of user u for item *i*.

$$Sim_{u,v} = \frac{\left|R_{u,i} \cap R_{v,i}\right|}{\left|R_{u,i} \cup R_{v,i}\right|} \quad (1)$$

One of the most widespread methods in measuring the similarity of two users is Pearson Correlation Coefficient (PCC) which aims to analyzing the behavior of user, considering ratings input by each user [7]. As you see in Equation (2) ratings are normalized by subtracting the user's mean rating, $\overline{R_u}$ , because each user may have a particular rating scale. I is the set of common items between user *u* and *v*. Despite the popularity of PCC, it has been the subject to a number of modifications that will be discussed later in a more exact way.

$$Sim_{u,v} = \frac{\sum_{i\in I}\left(R_{u,i} - \overline{R_u}\right)\left(R_{v,i} - \overline{R_v}\right)}{\sqrt{\sum_{i\in I}\left(R_{u,i} - \overline{R_u}\right)^2}\sqrt{\sum_{i\in I}\left(R_{v,i} - \overline{R_v}\right)^2}} \quad (2)$$

Another similarity metrics is cosine similarity as in (3) that two users are compared as two vectors, while the similarity between them depends on their cosine angle [8]. The main application of cosine similarity is evaluating the similarity of two documents by tracing each document as a vector of word frequencies and computing the cosine of the angle formed by the frequency vectors [9].In collaborative filtering, users or items which are corresponding to documents and ratings are used instead of word frequencies.

$$Sim_{u,v} = \frac{\sum_{i\in I} R_{u,i} R_{v,i}}{\sqrt{\sum_{i\in I} R_{u,i}^{\ 2}}\sqrt{\sum_{i\in I} R_{v,i}^{\ 2}}} \quad (3)$$

Shanle Ma [10] presented a hybrid recommender system to cope with the interest drifting problem that used a time-sensitive-function. One of the most related works to our study is [5] that investigated on the performance of collaborative filtering over time in which the similarity of users changes with time. But it was almost an analytical research but a clear way to cope relevant problem was not presented.

In all mentioned approaches, there is no difference in similarity of users considering items rated either in the same timestamps or dissimilar timestamps. In other words, the behavior of users is not monitored w.r.t time. There is still an open issue of measuring similarity in which comparing prediction accuracy demonstrates that one similarity measure can outperform another one based on the particular dataset.

## 3. Proposed Metric

There is a public standard notation to describe CF problems. The input can be structured as an M × N interaction matrix, A, associated with M user $U = \{u_1, u_2, \cdots, u_{\mathrm{m}}\}$ and N items $I = \{i_1, i_2, \cdots, i_{\mathrm{n}}\}$. In this matrix, $a_{k,j}$ can take the value of either 0 or 1, which 1 represents an observing transaction between $u_k$ and $i_j$ (for example, $u_1$ has purchased $i_1$.) and 0 represents the absence of any transaction. By using this matrix the level of user's interest for different items can be shown. For analyzing the behavior of users in a more exact way we need to know the time that each interaction is occurred. As we mentioned earlier, one important factor on the performance of similarity metrics is the ability of the recommender system to record the different type of feedback from users. There are several types of feedbacks which can be captured in these systems such as rating, listening habits for music recommendation and comments in e-commerce web site.

We analyze relation of two users in four phases and then combine these factors to compute the final result as the similarity among them. Actually, the main purpose of this step is calculating the difference between user's interests. Four factors are important for the final weight:

a) The number of common items.

b) The degree of user's interest in common items.

c) Timestamp of rated items.

d) Rating sequences.

However it is unclear which factor plays more important role. Considering conditions and goals of our business, each factor may have various effects on the final weight between users. For example, if there is no appropriate rating definition to evaluate the user's interest for a specific item, the first factor plays more important role. If the system and users have a dynamic behavior in various time intervals, the third factor becomes more important.

In the following, we explain how each factor is contributed in the proposed metric.

### 3.1. The Number of Common Items

The number of common items between users u and v, $NCI_{u,v}$, is first factor which is considered in our calculation. To normalize the value we use $MaxCI_{u,v}$. This function returns the number of all possible common items [11]. If one rates 10 items and the other one rates 5 items, they can have up to 5 common items. Therefore, Proportion in (4) is a normal value (range from 0 to 1) which represents the first similarity factor.

$$\text{Proportion} = \frac{NCI_{u,v}}{MaxCI_{u,v}} \tag{4}$$

### 3.2. The Degree of User's Interest in Common Items

The Equation (5) is the sum of rate differences between two users (*u* and *v*) and *CI* indicates the set of items which are in common between them.

$$SORD_{u,v} = \sum\nolimits_{i \in CI} \left| R_{u,i} - R_{v,i} \right| \tag{5}$$

To measure rating difference factor, Equation (6) is applied as $MaxRateDiff_{u,v}$ which determines the maximum difference that may occur between *u* and *v* w.r.t. item's rate. For example, if numerical feedback of the system is 5-star, we have:

$$MaxRateDiff_{u,v} = 4 * NCI_{u,v} \tag{6}$$

In order to measure the similarity of a pair by considering the first two factors, we use the following equation:

$$BaseWeight_{u,v} = \text{Proportion} * \left( 1 - \left( \frac{SORD_{u,v}}{MaxRateDiff_{u,v}} \right)^{1/EOR} \right) \tag{7}$$

EOR implies the effect of item's rate factor and is greater than one. By increasing the value of *EOR*, the difference between users becomes more important.

### 3.3. Timestamps of Rated Items

The users' similarity increases when they have more common time distance. The Equation (8) implies sum of time distance among *u* and *v* and is $T_{u,i}$ the time of rating *i* by user *u*.

$$SOTD_{u,v} = \sum\nolimits_{i \in CI} \left| T_{u,i} - T_{v,i} \right| \tag{8}$$

Consequently for the time factor we use (9) as an adaptive decay function. In this equation *α* is exploited for the purpose of adaptability which depends on the system's time scale (*α* > 0).

$$DF\left(SOTD_{u,v}\right) = \frac{\alpha}{\alpha + \left( \frac{SOTD_{u,v}}{NCI_{u,v}} \right)} \tag{9}$$

Through a learning process we can tune α to make the method more accurate. We will discuss this subject in Section 4.

### 3.4. The Sequence of Rated Items

In this section the distance between users w.r.t. interactions order is computed. If items rated by the user are kept in a queue, each item has a priority number. Therefore the distance can be calculated according to the following equation.

$$SOPD_{u,v} = \sum\nolimits_{i \in CI} \left| P_{u,i} - P_{v,i} \right| \qquad (10)$$

In above equation, $P_{u,i}$ is the priority of item *i* in *u*'s item queue. For example in **Figure 1,** $P_{u_1,D}$ =4 In order to determine the distance order (interaction sequence) of $u_1$from $u_2$ and $u_3$ in **Figure 1, (**10) is applied. As **Figure 1** shows, 7 items are in common among all users. The common items queue is illustrated as a link list. According to (10), we have $SOPD_{u_1,u_2}$ =24 and $SOPD_{u_1,u_3}$ =6, therefore u$_3$ is more similar to $u_1$ in comparison with $u_2$ w.r.t. interactions order. A normal value for calculating the impact of time in final result is needed. In (11), $NSOPD_{u,v}$ is a normal value (rang from 0 and 1) that uses $MSOPD(n)$ function in (12) which implies the maximum dissimilarity of interactions sequence of a user to another one when they have n common items.

$$NSOPD_{u,v} = \frac{SOPD_{u,v}}{MSOPD(n)} \qquad (11)$$

$$MSOPD(n) = \begin{cases} n^2 - n - 2\left(\dfrac{n}{2}\right)^2 \text{ if } n \text{ is odd} \\ (n/2)^2 \text{ otherwise} \end{cases} \qquad (12)$$

Using linear combination which is shown in (13), the distance between two users can be computed in w.r.t. rated items timestamp and order of common items.

$$\begin{aligned} &TF_{u,v} = \alpha\left(DF\left(SOTD_{u,v}\right)\right) + \beta\left(1 - NSOPD_{u,v}\right) \\ &\alpha + \beta = 1 \end{aligned} \qquad (13)$$

Finally, the similarity between two users by considering all four factors can be measured through (14) in which EOT determines how the value of $TF_{u,v}$ can effect on the final result.

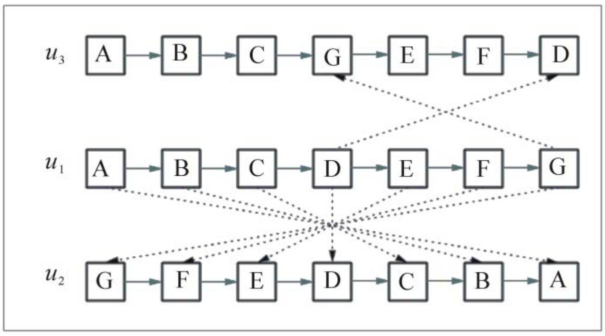


**Figure 1. $u_1$, $u_2$ and $u_3$ have purchased same items (A$\cdots$G) in different orders**.

$$Sim_{u,v} = BaseWeight_{u,v} * \left(EOT + \left((1 - EOT) * TF_{u,v}\right)\right) \qquad (14)$$

In above equation, 0 < $EOT$ < 1 and $0 < Sim_{u,v} < 1$

## 4. Evaluation

To evaluate the performance of our proposed metric, we need to apply it in creating recommendation process and compare the result with classic collaborative filtering algorithm[2]. Therefore the Weighted Sum of Other Ratings [7] is used to generate recommendations. The value of $Sim_{u,v}$ is calculated based on our metrics in the previous step. In this formula U is the collection of k-most similar user to u.

$$R_{u,i} = \overline{R}_u + \frac{\sum_{v \in U} \left(R_{v,i} - \overline{R}_v\right).Sim_{u,v}}{Sim_{u,v}} \qquad (15)$$

Statistical Accuracy Metrics evaluate the accuracy of a system by comparing the numerical recommendation rate on training dataset against the real user provided value for the user-item pairs in the test dataset [12,13]. In this study Mean Absolute Error (MAE) is used as one of the widely adopted metric for evaluation. Where *n* is the total number of ratings over all users, $p_{i,j}$ is the predicted rating for user *i* on item *j*, and $r_{i,j}$ is the actual rating. A lower MAE means a better prediction [14].

$$MAE = \frac{\sum_{\{i,j\}} \left| p_{i,j} - r_{i,j} \right|}{n} \qquad (16)$$

We have chosen MovieLens dataset as one of the most prevalent and reliable datasets for evaluation and comparison of recommender systems. MovieLens was collected through the MovieLens project, and was distributed by GroupLens Research Group at the University of Minnesota. However, each vote is accompanied by a time stamp, but for our own purpose, it is not an appropriate source for evaluating our research. For example, most of the time stamps belong to the time that the system asked the opinion of the user, not to the first time the user has seen the movie. Therefore, a user may rate 100 items in one day. As it can be seen in **Figure 2** the result of analyzing Movielens dataset shows that 55% of users interacted with the system in only one day.

As we mentioned before our approach can adopt itself to different circumstances and limitations of the target system. Because of this adaptively, the algorithm should be customized to our qualifications and concepts of our business. For achieving this purpose some parameters such as EOR, EOT and *α* should be set up. **Figure 3** shows the effect of each parameter with 200 neighbors

[2] In classic approach, PCC is applied as the similarity metric and Weighted Sum of Other's Ratings as the recommendation method.

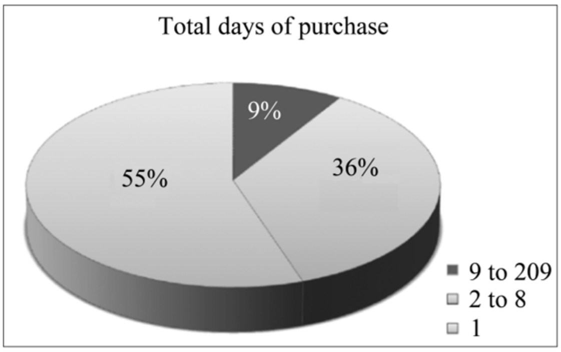
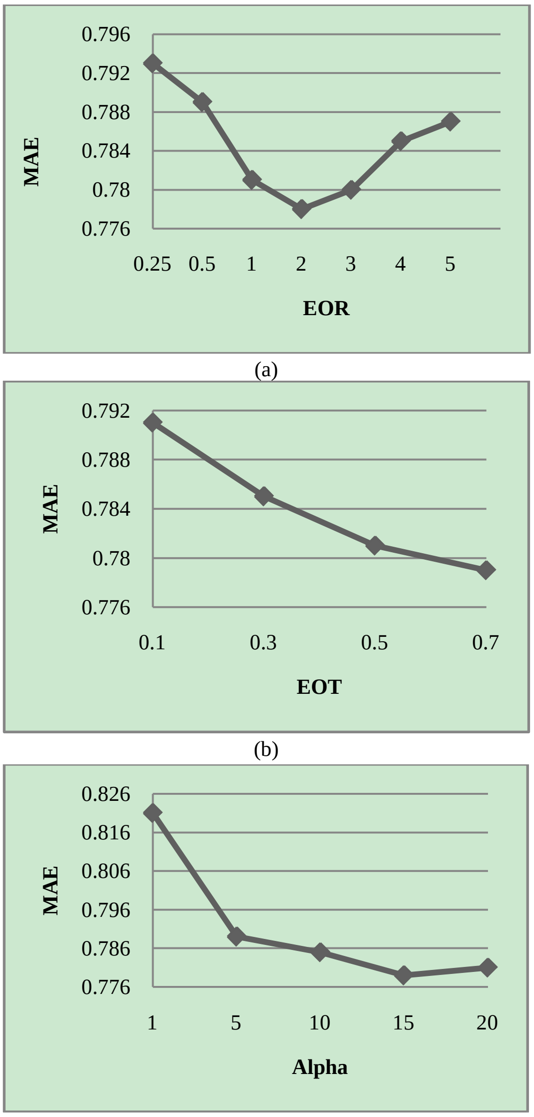


**Figure 2. In MovieLens dataset 55% of people have bought items only in one day.**



(a)

(b)

(c)

**Figure 3. Evaluation of the proposed approach, k = 200. (a) $\alpha$ = 10, EOT = 0.7 and 0.25 ≤ EOR ≤ 6. (b) $\alpha$ = 10, EOR = 2, 0.1 ≤ EOT ≤ 0.7. (c) EOT = 0.5, EOR = 2 and 1 ≤ $\alpha$ ≤ 15.**

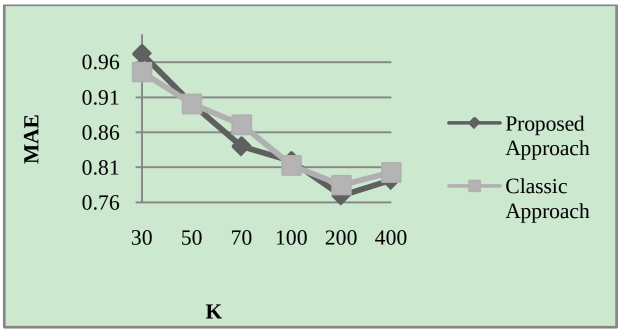


**Figure 4.Comparing the proposed approach with the classic approach.**

(k). This process helps us finding the best values of each parameter for our data set.

After achieving the best performance of our approach for MovieLens dataset ($\alpha$ = 10, *EOT* = 0.7 and *EOR* = 10), the accuracy of the proposed algorithm is compared with classic collaborative filtering algorithm. In **Figure 4**, the graph shows the progression of the MAE using the PCC and Our metrics when the number of neighbors, k, ranges from 30 to 400. The evaluation demonstrates the best accuracy is achieved with k=200.

## 5. Conclusions and Future Works

In this article, we proposed a new metric using the time factor as an important parameter for measuring the similarity of users. Our experiments show that our approach is more accurate than the classic one especially when users' interests change over the time.

We are currently working on tuning the required parameters such as *EOR* and *EOT* for a given dataset. In the future, we intend to include implicit factors as user comments in our approach.

## REFERENCES

[1] X. Su and T. Khoshgoftaar, "Imputed Neighborhood Based Collaborative Filtering," *Proceedings of the* 2008 *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Sydney, 9-12 December 2008, pp. 633-639.

[2] M. Deshpande and G. Karypis, "Item-Based Top-N Recommendation Algorithms," *ACM Transactions on Information Systems*, Vol. 22, No. 1, 2004, pp. 143-177. doi:10.1145/963770.963776

[3] G. Karypis, "Evaluation of Item-Based Top-N Recommendation Algorithms," *Proceedings of the* 10*th International Conference on Information and Knowledge Management*, Atlanta, 5-10 November 2001, pp. 247-254.

[4] L. He and F. Wu, "A Time-Context-Based Collaborative Filtering Algorithm," *IEEE International Conference on Granular Computing*, Nanchang, 17-19 August 2009, pp. 209-213.

[5] N. Lathia, “Evaluating Collaborative Filtering over Time,” Ph.D. Thesis, University College London, London, 2010.

[6] M. Charikar, “Similarity Estimation Techniques from Rounding Algorithms,” *Annual ACM Symposium on Theory of Computing*, Montreal, 19-21 May 2002, pp. 380-388.

[7] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom and J. Riedl, “Grouplens: An Open Architecture for Collaborative Filtering of Netnews,” *ACM Conference on Computer Supported Cooperative Work*, New York, 22-26 October 1994, pp. 175-186.

[8] B. M. Sarwar, G. Karypis, J. A. Konstan and J. Riedl, “Analysis of Recommendation Algorithms for E-Commerce,” *ACM Conference on Electronic Commerce*, New York, 17-20 October 2000, pp. 158-167.

[9] G. Salton and M. McGill, “Introduction to Modern Information Retrieval,” Facet Publishing, New York, 1983.

[10] S. Ma, X. Li, Y. Ding and M. E. Orlowska, “A Recommender System with Interest-Drifting,” 8*th International Conference on Web Information Systems Engineering*, Nancy, 3-7 December 2007, pp. 633-642.

[11] T. Segaran, “Programming Collective Intelligence,” O'Reilly Media, Sebastopol, 2007.

[12] K. Goldberg, T. Roeder, D. Gupta and C. Perkins, “Eigentaste: A Constant Time Collaborative Filtering Algorithm,” *Information Retrieval*, Vol. 4, No. 2, 2001, pp. 133-151. doi:10.1023/A:1011419012209

[13] J. L. Herlocker and J. A. Konstan, “Evaluating Collaborative Filtering Recommender Systems,” *ACM Transactions on Information Systems*, Vol. 22, No. 1, 2004, pp. 5-53. doi:10.1145/963770.963772

[14] X. Su and T. Khoshgoftaar, “A Survey of Collaborative Filtering Techniques,” Advances in Artificial Intelligence, Vol. 2009, 2009, pp. 1-20. doi:10.1155/2009/421425