New Collaborative Filtering Based on Similarity Integration and Temporal Information

Keunho Choi
Business School, Korea University
(keunho@korea.ac.kr)

Gunwoo Kim
Department of Business Administration, Hanbat National University
(gkim@hanbat.ac.kr)

Donghee Yoo
Department of Electronics Engineering and Information Science, Korea Military Academy
(donghee.info@gmail.com)

Yongmoo Suh
Business School, Korea University
(ymsuh@korea.ac.kr)

As personalized recommendation of products and services is rapidly growing in importance, a number of studies provided fundamental knowledge and techniques for developing recommendation systems. Among them, the CF technique has been most widely used and has proven to be useful in many practices. However, current collaborative filtering (CF) technique has still considerable rooms for improving the effectiveness of recommendation systems: 1) a similarity function most systems use to find so-called like-minded people is not well defined in that similarity is computed from a single perspective of similarity concept; and 2) temporal information that contains the changing preference of customers needs to be taken into account when making recommendations. We hypothesize that integration of multiple aspects of similarity and utilization of temporal information will improve the accuracy of recommendations. The objective of this paper is to test the hypothesis through a series of experiments using MovieLens data. The experimental results show that the proposed recommendation system highly outperforms the conventional CF-based systems, confirming our hypothesis.

Received : July 27, 2011 Revision : August 03, 2011 Accepted : August 16, 2011
Type of Submission : Excellent Paper of Spring Conference Corresponding author : Yongmoo Suh

1. Introduction

Personalized recommendation of products and services is rapidly growing in importance because it provides benefits to both sellers and buyers, enabling sellers to increase their revenue by providing buyers with the items they are likely to purchase and allowing buyers to find immediately what they want without wasting their time in an age of information overload. As a result of this growing importance, many studies provided fundamental knowledge and techniques for developing recommendation systems, such as content-based technique (Belkin and Croft, 1992; Lang, 1995; Mooney and Roy, 1999; Pazzani and Billsus, 1997), collaborative filtering (CF) (Joaquin and Naohiro, 1999; Nakamura and Abe, 1998; Si and Jin, 2003; Yu et al., 2004; Yu et al., 2002), and association rule (Aggarwal et al., 2002; Huang and Huang, 2009; Wang et al., 2008). Among them, the CF technique has been most widely used and has proven to be useful in many practices.

So far, a number of studies have enhanced the effectiveness of CF-based recommendation systems by resolving typical problems of CF-based recommendation technique, including the new user problem (also called the cold start problem) (Kim et al., 2010; Park and Chang, 2009), the new item problem (also called the first rater problem) (Balabanovic and Shoham, 1998; Lee et al., 2008), and the sparsity problem (Jeong et al., 2009; Kim et al., 2010; Lee and Olafsson, 2009; Park and Chang, 2009). As a result, many CF-based recommendation systems have been developed. However, they still have considerable rooms for improvement in terms of recommendation quality and value. This paper proposes a novel approach to improving them. First, we suggest a new similarity function which reflects multiple aspects of similarity concept. Note that most recommendation systems compute similarity based on a single aspect of similarity concept (that is, correlation, angle and distance). Second, we suggest that temporal information should be taken into account when making recommendation, since it contains the changing preference of customers. Third, the way of measuring the effectiveness of recommendation system taken by the paper is different from those taken by other papers. Most recommendation systems make recommendation based on a part of past transaction data and then check the effectiveness of recommendation using the other part of the same past transaction data[1]. On the contrary to these systems, our system make recommendation of items which are likely to be purchased in the near future, based on the temporal information included in the past transaction data. It would be of great value for the sellers to know by when buyers will purchase the recommended items.

Based on these ideas, we hypothesize that when multiple similarities computed from different perspectives are integrated and temporal information is utilized, the resulting recommendation quality and value will be improved. Therefore, the objective of this paper is to test the hypothesis through experiments using MovieLens data. We first located items to recommend for individual users, considering three aspects of similarity when we sought for like-minded neighbors and then checked whether or not each user had watched the recommended movies by a specific point of time after the recommendation was made.

The rest of paper is organized as follows. Section 2 reviews the literature regarding recom-

1) This is not a recommendation in a strict sense, since we make recommendation of items to someone in order for him or her to buy them later.

mendation and recommendation systems. Section 3 explains our basic ideas on similarity integration and utilization of temporal information. Then, Section 4 addresses the overall framework for realizing our ideas and provides a detailed description of each step of the framework. In Section 5, we explain how our ideas make a difference using the results from four experiments and describe the implications of each experiment. The last section contains the concluding remarks, including a summary, implications, and limitations of this research.

2. Literature Review

Many recommendation systems have been developed so far, but the fundamental techniques used in most recommendation systems can be categorized into four types: 1) content-based filtering (CBF); 2) collaborative filtering (CF); 3) hybrid approach of CBF and CF techniques; 4) association-based approach.

CBF recommendation systems typically implement these steps: 1) build a content-based user profile from the features of the items that each user purchased, 2) construct an item profile by extracting a set of features for each and every item in the item set, 3) compute the similarity scores between the user profiles and the item profiles, and 4) recommend one or more items with high similarity scores. That is, they recommend items based on the similarity between items.

At the beginning, these systems were used to recommend documents such as net news (Lang, 1995), web pages (Pazzani and Billsus, 1997), and books (Mooney and Roy, 1999). User profiles and item profiles consist of weights given to a set of keywords extracted from documents using information retrieval techniques (Baeza-Yates and Ribeiro-Neto, 1999; Salton, 1988) or information filtering techniques (Belkin and Croft, 1992). Since both profiles are represented by weight vectors, a similarity score is computed using a heuristic function such as cosine similarity function or Pearson correlation (Balabanovic and Shoham, 1998; Lang, 1995). More recently, other techniques, such as classification models built from statistical approach (Mooney and Roy, 1999) or data mining approach (Pazzani and Billsus, 1997), have been used to classify whether a document item is relevant to a user. These systems have a few limitations: 1) it is not easy to get a sufficient number of features to build profiles (insufficient features problem) (Shardanand and Maes, 1995), 2) recommended items are limited to those that are similar to the items that have been purchased by a target user (over-specialization problem) (Adomavicius, 2005), and 3) new users who did not purchase items before or users with unusual preference cannot get a proper recommendation (new or unusual user problem) (Adomavicius, 2005; Billsus and Pazzani, 2002).

CF-based recommendation systems are implemented generally as follows: 1) select like-minded users using similarity function which represents the similarity between a target user and every other user based on the rating information on the common items that both users rated; 2) predict the ratings on items as an average, weighted sum or adjusted weighted sum of ratings on items given by the selected like-minded users; 3) recommend top-rated *k* items. That is, they recommend items based on the similarity between users.

These methods of rating prediction are called memory-based (Joaquin and Naohiro, 1999; Nakamura and Abe, 1998; Si and Jin, 2003; Yu et al., 2004; Yu et al., 2002). Another method of rating prediction, called model-based, is one in which a model such as a probabilistic model or a machine learning model is built from a large collection of ratings and is used to predict ratings of items (Billsus and Pazzani, 1998; Cheung et al., 2003; Getoor and Sahami, 1999; Goldberg et al., 2001; Hofmann, 2003; Hofmann, 2004; Kumar et al., 2001; Marlin, 2003; Pavlov and Pennock, 2002; Pennock et al., 1999; Shani et al., 2002).

After the term 'CF technique' was first coined, various recommendation systems were developed, including Tapestry for recommending news articles (Goldberg et al., 1992), GroupLens for net news (Resnick et al., 1994), and Ringo for music (Shardanand and Maes, 1995). CF-based recommendation systems also have some limitations: 1) it is difficult to recommend items for those users who have never rated items before (new user problem) (Kim et al., 2010; Park and Chang, 2009), 2) it is difficult to recommend items which have not been rated before (new item problem) (Balabanovic and Shoham, 1998; Lee et al., 2008), and 3) they make poor recommendations when rating information is insufficient (sparsity problem) (Jeong et al., 2009; Kim et al., 2010; Lee and Olafsson, 2009; Park and Chang, 2009).

Most of the limitations of CBF and CF-based recommendation systems have been overcome by a hybrid approach of combining these two kinds of recommendation systems (Balabanovic and Shoham, 1998; Liu et al., 2010; Salter and Antonopoulos, 2006; Wei et al., 2008). Fab system (Balabanovic and Shoham, 1998) combines the CF technique with the CBF technique to eliminate some weaknesses of CBF technique (e.g., insufficient features and over-specialization problems) and one of CF technique (e.g., new item problem). In this system, content-based user profiles are maintained to determine similar users for collaborative recommendation. Items are recommended to a target user when the following two conditions are simultaneously satisfied: 1) each item must have a high score against the target user's profile and 2) each item should be highly rated by users whose profiles are similar to the target user' profile. Liu et al. (2010) proposed another hybrid recommendation system to deal with the problems of the CBF and CF techniques such as sparsity and scalability. To resolve the sparsity problem, blanks in the user-item rating matrix are first filled with weighted rating average on items which a user rated, where the weight of the rated item is calculated by the similarity between an unrated item and the rated item in their feature values, and then CF technique is applied to the user-item rating matrix. To resolve the scalability problem, all users are classified into different groups with respect to the user personality features, and then neighbors of a target user are found within the group to which the user belongs instead of searching the entire user space.

Another simple but popular way of recommending items to a user is to make use of the association among the items. Association can be derived from a large amount of transaction database collected over time, using data mining techniques. It could be either an association rule among the items sold together (Aggarwal et al., 2002) or a sequential pattern among the items sold in sequence (Huang and Huang, 2009; Wang et al., 2008). Aggarwal et al. (2002) proposed a technique for discovering localized association rules among items which are helpful for target marketing. They first clustered market basket data using both the mushroom dataset and adult dataset in the UCI machine learning repository[2] and then derived association rules from each cluster. Huang and Huang (2009) proposed a sequential pattern-based recommendation system that predicts the customer's time-variant purchase behavior in a supermarket. They first clustered customers and derived sequential patterns among food items for each cluster in each time period. By taking into account the dynamic nature of a customer's purchase sequences, they improved the recommendation quality.

3. New Ideas for Improving the Performance of CF Technique

Although a number of studies have been conducted in an attempt to resolve the aforementioned problems of the CF technique, relatively few studies have addressed two critical issues of the CF technique (i.e., similarity integration and utilization of temporal information) which we think will contribute to the accuracy of CF-based recommendation systems. Thus, we would first look into these issues and describe our ideas on these issues in this section and then propose a new approach to resolving them in the next section.

As for the first issue of similarity integration, similarity in most existing CF-based recommendation systems is calculated using a similarity function based on Pearson correlation coefficient (Albadvi and Shahbazi, 2009; Kwon et al., 2009; Lee et al., 2008; Russell and Yoon, 2008; Salter and Antonopoulos, 2006), cosine similarity (Jeong et al., 2009; Lee et al., 2008; Symeonidis et al., 2008), or a distance measure (Adomavicius and Kwon, 2007; Kim et al., 2009; Kim et al., 2008; Park and Chang, 2009). Each of these functions measures the similarity between a target user and every other user to find like-minded users of the target user, but each from different perspective. The Pearson correlation coefficient calculates the similarity from the perspective of rating pattern (Albadvi and Shahbazi, 2009; Kwon et al., 2009; Lee et al., 2008; Russell and Yoon,

2) http://archive.ics.uci.edu/ml/.

2008; Salter and Antonopoulos, 2006), cosine similarity from the perspective of angle between two rating vectors (Albadvi and Shahbazi, 2009; Jeong et al., 2009; Lee et al., 2008; Symeonidis et al., 2008), and distance measure from the perspective of magnitude of rating (Kim et al., 2009; Kim et al., 2008; Park and Chang, 2009). Depending on the similarity function to be used, the set of like-minded users could be different, thus so are the items recommended. In our system, we integrated these three different aspects of similarity into a similarity function to find like-minded users, more similar to the target user.

As for the next issue of utilizing temporal information, users' tastes or preferences will change over time and users will have more interest in items that were launched more recently. Therefore, we need to pay attention to two kinds of temporal information such as purchasing time and item launch time. However, most CF-based recommendation systems have used rating information on items without taking such temporal information into account. Therefore, in our recommendation system, we divided both purchasing time and item launch time into several periods and put more weight on more recent periods, when making recommendations of items. There have been some researches related to this idea. Tang et al. (2005) made a similar attempt to make recommendations more efficient by considering only recently launched movies, but they ran into the problem of not being able to recommend old movies at all. Lee et al. (2008) proposed an implicit rating matrix using both item launch time and purchasing time, but their research encountered a problem that all users have the same rating on an item if they purchase the item during the same time period. Our system is different from this system in computing ratings for items. That is, the former first predicts ratings for items, assigns a weight for each category which is defined based on purchasing time and launch time, and then computes new ratings for items, reflecting the weight. On the other hand, the latter gives the same ratings to items belonging to the same category which is formed in the same manner as the former.

4. A Proposed Recommendation System

In this section, we first present an overall framework to realize our ideas on the two critical issues mentioned in previous section. Then we provide a detailed description of each step of the framework.

4.1 Overall Framework

As shown in <Figure 1>, the overall framework of our proposed recommendation system consists of six steps. Steps 1-4 are executed in each time period.

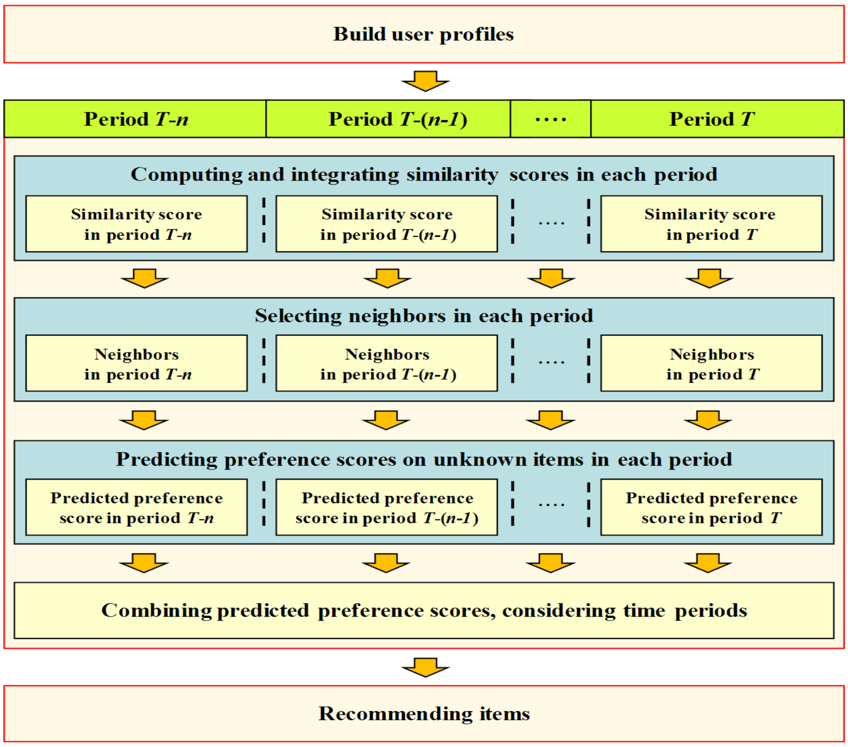
Step 1. ***Building user profiles***: user profiles are built from user-item rating data collected during the whole period and are divided into several pieces based on when items are purchased.

Step 2. ***Computing and integrating similarity scores***: three similarity scores between a target user and every other user in the search space are calculated using three similarity functions (i.e., Pearson correlation coefficient, cosine similarity, and distance measure), and then they are integrated into one score.

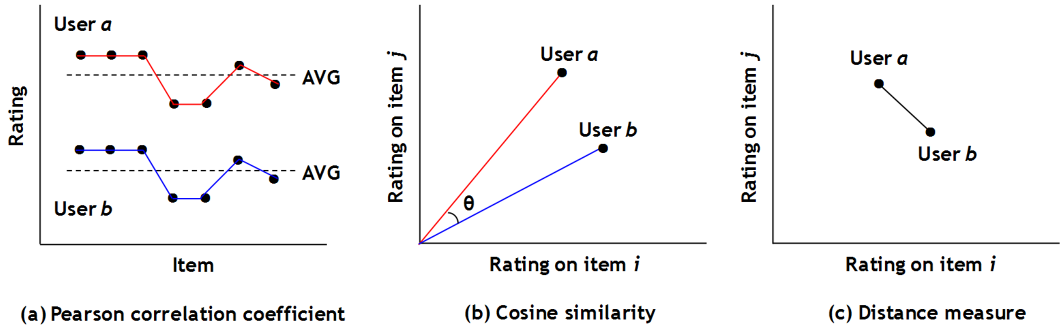
Step 3. ***Selecting neighbors***: based on the integrated similarity scores, top k neighbors with the most similar preferences to the target user are selected.

Step 4. ***Predicting preference scores on unknown items***: preference scores (that is, ratings) on unknown items by a target user are predicted based on the ratings by the top m neighbors.

Step 5. ***Combining predicted preference scores, considering time periods***: the predicted preference scores computed in each time period are merged with different weights put to categories of different purchasing time and different launch time, to get final predicted preference score.



<Figure 1> Overall Framework of the Proposed Recommendation System



<Figure 2> Different Perspective of Three Similarity Functions

Step 6. ***Recommending items***: top n items with the highest final predicted preference scores are recommended to the target user.

4.2 Building User Profiles

User profiles are initially built from user-item rating information collected during the whole period. Then purchasing time is divided into several periods. The number and the duration of a time period can be set differently according to the policy of the company and the characteristics of items sold. To make the explanation simple, we set the number of time period to three. Once the number of time period is set, the user profile is built in each period and updated whenever the time period moves forward. The user profile is represented as follows.

$$User\ Profile = \{(R_i, \cdots, R_n), (R_{i'}, \cdots, R_{n'}), (R_{i''}, \cdots, R_{n''})\},$$

where R_i, $R_{i'}$, and $R_{i''}$ in the first, the second, and the third parentheses denote the ratings of a user on items i, i', and i'' in periods T, T-1, and T-2, respectively.

4.3 Computing and Integrating Similarity Scores

Most CF-based recommendation systems have used only one of the similarity functions (i.e., Pearson correlation coefficient, cosine similarity, or distance measure) to compute the similarity between users, although some systems used two or more similarity functions to compare their performance (Adomavicius and Kwon, 2007).

Pearson correlation coefficient estimates the similarity between a target user and every other user based only on the rating pattern between the two, as shown in <Figure 2>(a).

Pearson correlation coefficient between two users is defined as

$$P(a,b) = \frac{\Sigma_{i=1}^{m}(R_{a,i}-\overline{R_a})(R_{b,i}-\overline{R_b})}{\sqrt{\Sigma_{i=1}^{m}(R_{a,i}-\overline{R_a})^2}\sqrt{\Sigma_{i=1}^{m}(R_{b,i}-\overline{R_b})^2}}, \quad (1)$$

where $R_{a,i}$, $R_{b,i}$, $\overline{R_a}$, $\overline{R_b}$ denote the ratings of users a and b on item i, the averages of all $R_{a,i}$

and all $R_{b,i}$, respectively and m denotes the number of items purchased by both users a and b. The coefficient can be any value between +1.0 and -1.0, inclusive. The coefficient, +1.0 implies that the rating pattern (or tendency) of two users is perfectly identical, while the coefficient, -1.0 means that the rating pattern of the two differs diametrically.

Cosine-based approach treats two users as two vectors in an *m*-dimensional rating vector space. In this approach, the similarity between a target user and every other user is measured by calculating the cosine value of the angle between the two vectors, as defined in Equation (2). The angle between the two vectors approaches zero as the preferences of two users becomes more similar, as shown in <Figure 2>(b).

$$C(a, b) = \frac{\Sigma_{i=1}^{m}(R_{a,i})(R_{b,i})}{\sqrt{\Sigma_{i=1}^{m}(R_{a,i})^2}\sqrt{\Sigma_{i=1}^{m}(R_{b,i})^2}} \quad (2)$$

Finally, the distance measure is used to calculate the absolute magnitude of the similarity between a target user and every other user in the *m*-dimensional space of rating on an item, as shown in <Figure 2>(c), where two users, *a* and *b*, are thought to be more similar to each other when the distance is shorter. Therefore, the distance-based similarity can be defined as an inverse of the distance, as is shown in Equation (3). The similarity may take any value between 0 and 1.

$$D(a, b) = \frac{1}{1+\sqrt{\Sigma_{i=1}^{m}(R_{a,i}-R_{b,i})^2}} \quad (3)$$

Since these three functions measure the similarity between two users from different perspectives, depending on the functions used, the set of neighbors whose information is used to predict the rating on items unknown to the target user could be different and thus so are the items recommended. Suppose, for example, that the ratings of a target user *A* and other users *B*, *C* and *D* for five items are given as shown in <Table 1>. Then <Table 2> shows that the similarity scores between any two users computed using different similarity functions are different and so are the ranks among the scores.

<Table 1> Example of User-Item Rating Matrix

User	Item1	Item2	Item3	Item4	Item5
A	5	4	5	5	4
B	5	4	5	4	4
C	5	3	5	3	4
D	2	1	2	2	1

<Table 2> Example of Similarity Measure and Rank, Depending on Similarity Function

	Pearson corr.		Cosine		Dis.-based	
	Sim.	Rank	Sim.	Rank	Sim.	Rank
A,B	0.667	2	0.996	1	0.500	1
A,C	0.456	3	0.981	3	0.309	2
A,D	1.000	1	0.982	2	0.130	3

Therefore, to lessen the absurdity resulting from using a single similarity function, we integrated the three similarity scores computed from three different perspectives (i.e., pattern, angle and magnitude) into one. Another thing that we took into account while integrating them is that a user's rating information may be more reliable if the user rated more items that a target user rated than other users did. Therefore, we defined a new term, degree of match, as n_b/N, where n_b is the number of commonly rated items between the user b and a target user, and N is the maximum of all such n_u for every user u. The similarity scores are multiplied by the degree of match prior to integrating the three similarity scores, each computed using Equations (1) through (3).

Finally, we defined the integrated similarity between users a and b from the three similarity scores to be their harmonic average, as shown in Equation (4), where $P(a, b)$, $C(a, b)$ and $D(a, b)$ represent similarity scores computed using Equations (1), (2) and (3), respectively.

$$Isim(a,b) = \frac{3 \times P(a,b) \times C(a,b) \times D(a,b)}{P(a,b) \times C(a,b) + C(a,b) \times D(a,b) + D(a,b) \times P(a,b)} \times \frac{n_b}{N} \quad (4)$$

4.4 Selecting Neighborhoods and Predicting Preferences Scores on Unknown Items

Once the integrated-similarity scores are calculated, candidate neighbors are sorted in descending order of the score and the top k users are selected as neighbors of a target user in each period. To predict preferences on items unknown to a target user, we use an existing aggregation function, shown in Equation (5), which CF-based recommendation systems have generally used.

$$P_{a,i}^{predicted} = \overline{R}_a + \frac{1}{\Sigma_{b=1}^{N_c} |Isim(a,b)|} \times \Sigma_{b=1}^{N_c} Isim(a,b) \times (R_{b,i} - \overline{R}_b) \quad (5)$$

where $P_{a,i}^{predicted}$ denotes the predicted preference score of user a on item i, N_c the number of user a's neighbors, $\overline{R_a}$ the average rating of user a, $Isim(a, b)$ the integrated similarity between users a and b defined in Equation (4).

1.1. The aggregation function in Equation (5) was introduced to lessen the rating scale problem by considering the increase or decrease $(R_{b,i} - \overline{R_b})$ of neighbors' ratings on the item from their average rating. However, the rating scale problem still remains in that the average rating of a target user and that of a neighbor may be different $(\overline{R_a} \neq \overline{R}_b)$. Thus, by considering the rate of increase or decrease $(R_{b,i} - \overline{R_b})/\overline{R}$ of neighbors' ratings on the item from their average rating, we propose Equation (6) as a new aggregation function.

$$P_{a,i}^{predicted} = \overline{R}_a + \frac{1}{\Sigma_{b=1}^{N_c} |Isim(a,b)|} \times \Sigma_{b=1}^{N_c} Isim(a,b) \times \overline{R_a} \times \frac{(R_{b,i} - \overline{R_b})}{\overline{R}_b} \quad (6)$$

4.5 Combining Predicted Preference Scores, Considering Time Periods

Once the predicted preference scores (that is, ratings) on items unknown to a target user in each period is computed, the final predicted preference score on each item is calculated in two steps: (1) predicted ratings in each period are multiplied by temporal weights. More weight is given to more recent periods; (2) the weighted preference scores of all periods are added together.

For example, if the predicted preference scores of a target user on old launch item i is 3, 5, and 4 in periods T-2, T-1, and T, respectively, then the final predicted preference score rating using the temporal weight matrix in <Table 3> is 3.7 (= 3 × 0.2 + 5 × 0.3 + 4 × 0.4). Similarly, we can compute the final predicted preference scores for middle-launch items and recent-launch items.

<Table 3> Temporal Weight Matrix

	Time period		
	T-2	T-1	T
Old launch	0.2	0.3	0.4
Middle launch	0.5	0.6	0.7
Recent launch	0.8	0.9	1

We determined a set of appropriate temporal weights that should be given to each purchase time period and to each launch time period, after we had attempted to compute ratings with different sets of adjusted weights. The weights given in <Table 3> are those that are proven to return the best final predicted preference scores. Note that the resulting rating can be out of the range of item rating scale, but it does not matter since our concern is to find the top n items with the highest final combined ratings.

4.6 Recommending Items

After the final combined ratings on items unknown to a target user are calculated, the top n items with the highest final predicted preference scores are recommended.

5. Experiments

5.1 Experimental Design

To evaluate the effectiveness of our proposed recommendation system, we made two different data sets from MovieLens[3)] data, collected from April 25, 2000 to February 24, 2003 (34 months), comprising 515,088 rating records by 1,000 users on 3,646 movies, where the rating scales are from one to five. One (Dataset 1) is prepared to evaluate the effect of integrating three similarity measures without taking temporal information into account, as was done by most researches on recommendation system, and the other (Dataset 2) to evaluate the effect of both utilizing temporal information and integrating different similarity perspectives. We implemented all benchmark systems and our proposed recommendation systems using Transact-SQL in Microsoft SQL Server 2008.

Dataset 1 Among the entire users, 20 users who had watched movies more than 20 times were randomly selected as target users. 70 percent of each target user's data were used as training data, and the remaining data were used as test data.

In order to ensure that integrating different similarity perspectives actually improves the accuracy of recommendation system, we imple-

3) http://movielens.umn.edu/.

mented quite a few recommendation systems. We classified them into the following two groups (Note: all systems are implemented based on the aggregation function in Equation (5)):

1-A: Three conventional CF-based recommendation systems, CF_P, CF_C and CF_D (all represented in one as CF_P/C/D): these systems compute similarity using Pearson correlation coefficient, cosine similarity, and distance-based similarity, respectively.

1-B: A system integrating similarities computed from three different perspectives, CF_PCD: This system implements our idea of similarity integration.

Dataset 2 The whole time period of 34 months were divided into three six-month time periods and one final sixteen-month time period. The data from the first three time periods (i.e., the first 18 months) were used for training, and the remaining data (i.e., the last 16 months) were used for testing. 20 target users were randomly selected based on whether they had watched movies more than 20 times in the training data, and had watched a movie within the first six months in the test data. For the evaluation of our recommendation system, we compared the movies the target users had watched during the entire test period with the movies our system recommended.

We divided item launch times into three periods (i.e., old, middle, and recent), where old launch time period contains 1,087 movies released before 1985 (163,996 records), middle contains 1,227 movies released from 1986 to 1995 (195,274 records), and recent contains 1,332 movies released from 1996 to 2000 (155,818 records).

Then, in order to ensure that integrating different similarity perspectives and utilizing temporal information actually improve the accuracy of recommendation system, we implemented quite a few recommendation systems. We classified them into the following three groups (Note: all benchmark systems are implemented based on the aggregation function in Equation (5)):

2-A: Three conventional CF-based recommendation systems, CF_P', CF_C' and CF_D' (all represented in one as CF_P'/C'/D'): these systems compute similarity using Pearson correlation coefficient, cosine similarity, and distance-based similarity, respectively.

2-B: Three CF-based recommendation systems, CF_P'_T, CF_C'_T and CF_D'_T (all represented in one as CF_P'/C'/D'_T): CF_P'_T is a new version of CF_P' that implements the ideas of temporal information. Similarly, CF_C'_T and CF_D'_T are new versions of CF_C' and CF_D', respectively.

2-C: Our proposed system, CF_P'C'D'_T: This system implements both ideas of temporal information and similarity integration.

We conducted our experiments by setting

the number of neighbors to 10[4]) and the number of recommended movies to 10, 20, 30, 40 or 50.

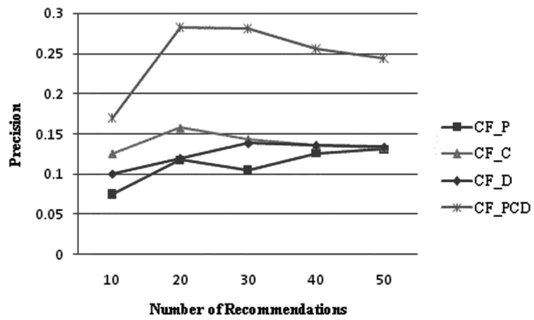
5.2 Experimental Results and Analysis

In this section, we explain how our proposed ideas affect the recommendation quality by analyzing the results from four experiments and describe the implication of each idea.

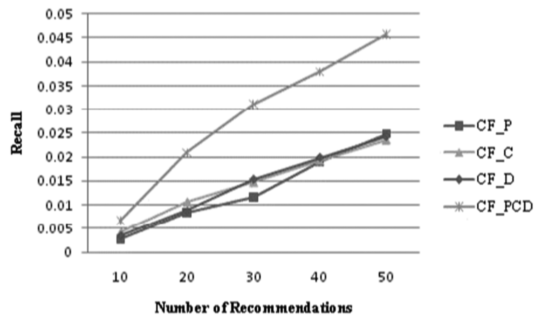
Experiment 1 To evaluate the effect of integrating similarities computed from three different perspectives using dataset 1, we compared three benchmark systems in group 1-A (i.e., CF_ P/C/D) with a system in group 1-B (i.e., CF_ PCD). We used precision, recall, F1 (the harmonic average of precision and recall), and mean absolute error (MAE) to measure and compare the recommendation quality, as have been used in many other studies to evaluate the recommendation quality (Albadvi and Shahbazi, 2009; Chen et al., 2008; Huang and Huang, 2009; Kim et al., 2009; Kwon et al., 2009; Liu et al., 2009; Park and Chang, 2009; Russell and Yoon, 2008).

A system adopting the integrated similarity showed considerably better recommendation quality than the three benchmark systems in group 1-A in terms of precision, recall, F1 (see <Figure 3>(a), <Figure 3>(b), and <Figure 3>(c))[5]), and MAE (see <Figure 4>). Especially, CF_PCD system showed relatively lower MAE than was shown by other CF-based recommendation systems using the same MovieLens data (Ahn, 2008; Bobadilla et al., 2010).

This result from Experiment 1 confirmed our idea that the integration of similarity scores computed from three different perspectives (i.e., pattern, angle and magnitude) will improve the recommendation quality.



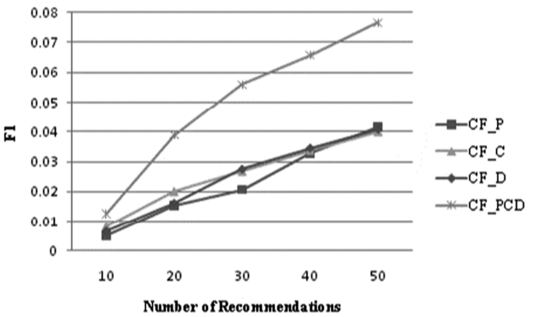
(a) Precision



(b) Recall

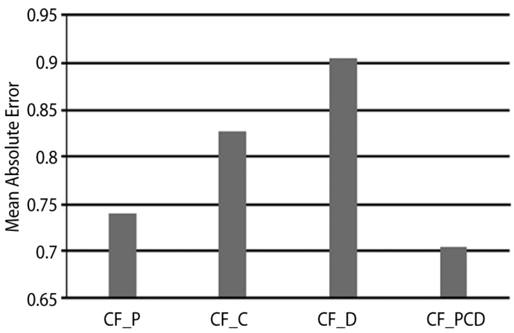
4) We assumed that the number of neighbors will make little or no difference when comparing various CF-based recommendation systems in terms of recommendation quality.

5) As the number of recommendations increases over 30, CF_PCD system shows a slightly decreased precision. So do CF_C and CF_D systems. We think that this is because movies with low predicted rating are likely to be included in recommendation list as the number of recommended movies increases.



(c) F1

<Figure 3> Comparison Between Systems that Use a Single Similarity and a System that Uses Integrated Similarities in Dataset 1 in terms of Precision, Recall, and F1

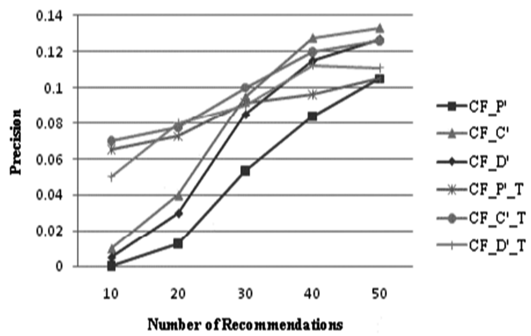


<Figure 4> Comparison Between Systems that Use a Single Similarity and a System that Uses Integrated Similarities in Dataset 1 in terms of MAE

Experiment 2 To evaluate the effect of our ideas including both utilization of temporal information and integration of different similarity perspectives, we further conducted three experiments using dataset 2. We used precision, recall, and F1 to measure and compare the recommendation quality.

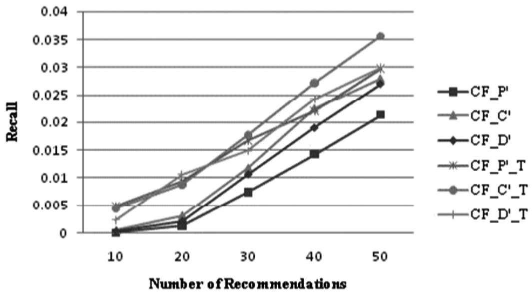
Experiment 2-1 To evaluate the effect of using temporal information, we compared three conventional CF-based recommendation systems in group 2-A (i.e., CF_P'/C'/D') with three other benchmark systems in group 2-B (i.e., CF_P'/C'/D'_T).

Experimental results are given in <Figure 5>, where CF_P'/C'/D'_T systems in group 2-B showed better precision, recall, and F1 than the corresponding CF_P'/C'/D' systems in group 2-A, in most cases[6]. This experiment confirmed our idea that users' tastes and preferences change over time and users have more interest in newly purchased items and/or newly launched items. Therefore, we can come to the conclusion that recommendation quality can be improved by considering temporal information that reflects the changing preferences of customers over time.

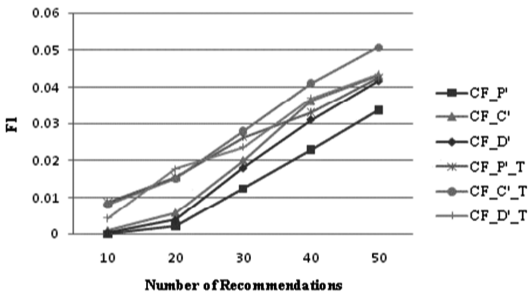


(a) Precision

6) Note the exception that CF_C' and CF_D' show slightly better precision than CF_C'_T and CF_D'_T, respectively when the number of recommendations are 40 and 50.



(b) Recall

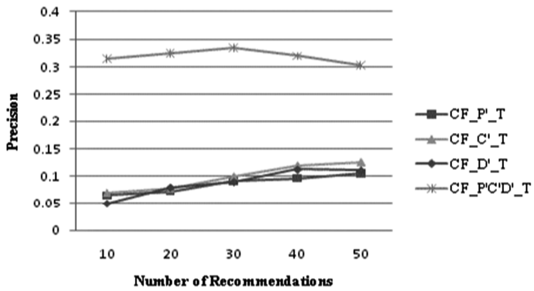


(c) F1

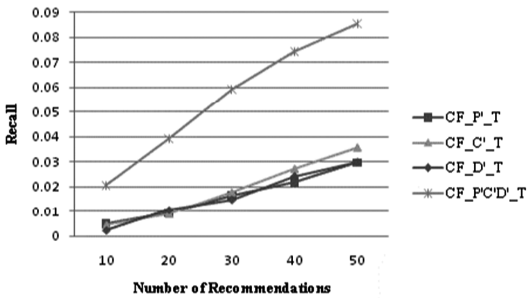
<Figure 5> Comparison Between Systems that Implemented the Idea of Using Temporal Information and Systems that did not in Dataset 2 in terms of Precision, Recall, and F1

Experiment 2-2 To evaluate the effect of integrating different similarity perspectives as well as taking temporal information into account, we compared three benchmark systems in group 2-B (i.e., CF_P'/C'/D'_T) with our recommendation system in group 2-C (i.e., CF_P' C'D'_T).

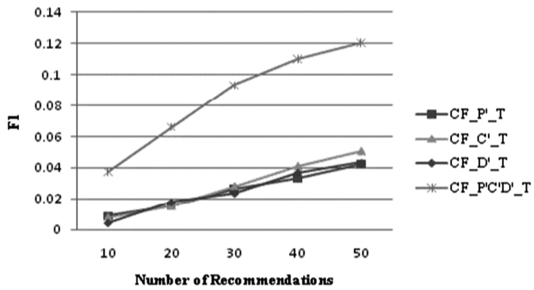
<Figure 6>(a), <Figure 6>(b) and <Figure 6>(c) show that our proposed system showed considerably better precision, recall, and F1 than the three benchmark systems. This result suggests that the integration of different similarity perspectives is likely to be an important factor in improving the recommendation accuracy, as was shown also by the result from Experiment 1.



(a) Precision



(b) Recall

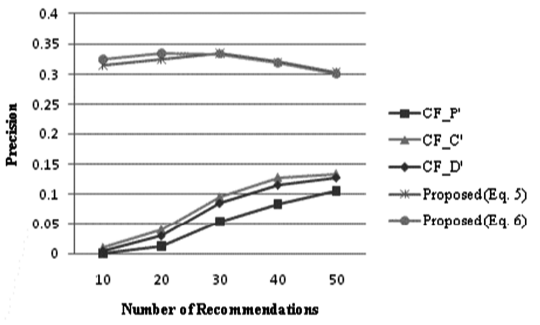


(c) F1

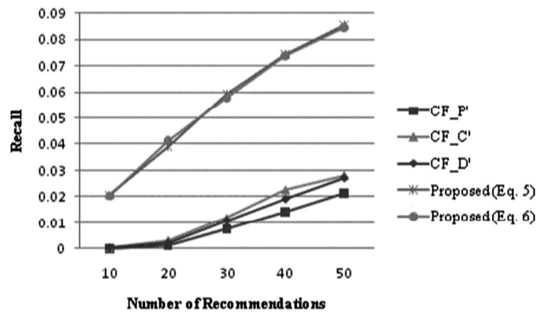
<Figure 6> Comparison Between Systems that Use a Single Similarity and a System that Uses Integrated Similarities in Dataset 2 in terms of Precision, Recall, and F1

Experiment 2-3 To evaluate the overall improvement offered by our proposed recommendation system, we compared our proposed system in group 2-C (i.e., CF_PCD_T) with three conventional benchmark systems in group 2-A (i.e., CF_P/C/D). We also tested whether two different aggregation functions, represented as Equation (5) and Equation (6), make any difference in recommendation quality. Our proposed CF systems implemented based on Equation (5) and Equation (6) are represented as Proposed (Eq. 5) and Proposed (Eq. 6) in <Figure 7>, respectively.

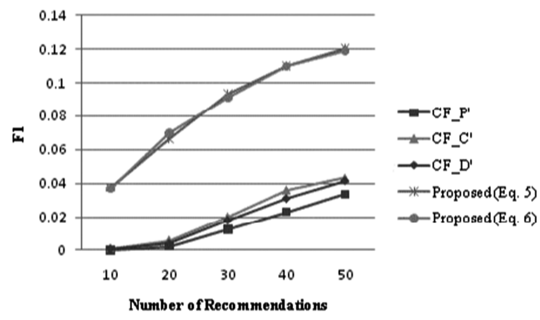
<Figure 7>(a), <Figure 7>(b) and <Figure 7>(c) show that our proposed CF systems are considerably better than the three conventional benchmark systems in terms of all three measures. This result suggests that our approach contributes to the accuracy of CF-based recommendation systems. On the other hand, the difference in three measures between our systems implemented based on Equation (5) and based on Equation (6) is negligible.



(a) Precision



(b) Recall



(c) F1

<Figure 7> Comparison Between Our Proposed Recommendation Systems Using Two Kinds of Aggregation Functions (Represented as Proposed (Eq. 5) and Proposed (Eq. 6) and Three Conventional CF Systems in Dataset 2 in terms of Precision, Recall, and F1.

6. Conclusions

This research proposed a new novel approach to improving the accuracy of current CF-based recommendation system, based on two ideas of integrating different similarity perspectives and of integrating temporal information.

As for the first idea of integrating different similarity perspectives, we integrated the similarity scores computed from three different perspectives so that similarity is not biased toward a specific perspective. As for the second idea of integrating temporal information, we used two kinds of temporal information, purchasing time and item launch time, in the recommendation process of CF technique. We hypothesize that when multiple similarity perspectives are integrated and temporal information is utilized, the resulting recommendation quality will be improved.

To validate our hypothesis, we conducted a series of experiments, from which we obtained some meaningful results as follows:

- CF-based recommendation systems which implement the idea of integrating different similarity perspectives perform considerably better than other CF-based recommendation systems. It is because the resulting integrated similarity is helpful in finding more appropriate neighbors of a target user, which in turn leads to the higher accuracy of recommendation.
- CF-based recommendation systems which implement the idea of utilizing temporal information perform better than other CF-based recommendation systems in general. It is because user preference on items changes over time and thus temporal information should be taken into account when making recommendations for more accurate recommendation.
- Our proposed recommendation system, which implements the above two ideas, highly outperforms the conventional CF-based recommendation systems in all measures, regardless of the number of recommendations.

Through a series of experiments, it is shown that integrating three different similarity perspectives and utilizing temporal information contribute to improving the accuracy of CF-based recommendation systems. This study, however, has some limitations. If the dataset we have used in our experiments were collected for much longer time span, it would be possible to conduct the same experiments with more time periods for the training data and with longer time period for the test data. Then, we may derive a knowledge such as "The recommended items will be purchased within the next X years with the expected accuracy of Y %.", which would provide more value to the sellers of the items. Also, our experiments have been conducted with a fixed number of neighbors and it may be worthwhile to do the experiments while changing the number of neighbors, although the same results are expected.

References

Adomavicius, G., "Towards the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions", *IEEE Transactions on Knowledge and Data Engineering*, Vol.17, No.6(2005), 734~749.

Adomavicius, G. and Y. Kwon, "New Recommendation Techniques for Multicriteria Rating Systems", *IEEE Intelligent Systems*, Vol.22, No.3(2007), 48~55.

Aggarwal, C. C., C. Procopiuc and P. S. Yu, "Finding Localized Associations in Market Basket Data", *IEEE Transactions on Knowledge and Data Engineering*, Vol.14, No.1 (2002), 51~62.

Ahn, H. J., "A New Similarity Measure for Collaborative Filtering to Alleviate the New User Cold-Starting Problem", *Information Sciences*, Vol.178, No.1(2008), 37~51.

Albadvi, A. and M. Shahbazi, "A Hybrid Recommendation Technique based on Product Category Attributes", *Expert Systems with Applications*, Vol.36, No.9(2009), 11480~11488.

Baeza-Yates, R. A. and B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley Longman Publishing Co., Inc., Boston, 1999.

Balabanovic, M. and Y. Shoham, "Content-Based, Collaborative Recommendation", *Communications of the ACM*, Vol.40, No.3(1998), 66~72.

Belkin, N. J. and W. B. Croft, "Information Filtering and Information Retrieval: Two Sides of the Same Coin", *Communications of the ACM*, Vol.35, No.12(1992), 29~38.

Billsus, D. and M. J. Pazzani, "Learning Collaborative Information Filters", *15th International Conference on Machine Learning*, (1998), 48~56

Billsus, D. and M. J. Pazzani, "User Modeling for Adaptive News Access", *User Modeling and User-Adapted Interaction*, Vol.10, No.2 (2000), 147~180.

Bobadilla, J., F. Serradilla and J. Bernal, "A New Collaborative Filtering Metric that Improves the Behavior of Recommender Systems", *Knowledge-Based Systems*, Vol.23, No.6(2010), 520~528.

Chen, L. S., F. H. Hsu, M. C. Chen and Y. C. Hsu, "Developing Recommender Systems with the Consideration of Product Profitability for Sellers", *Information Sciences*, Vol.178, No.4(2008), 1032~1048.

Cheung, K. W., J. T. Kwok, M. H. Law, and K. C. Tsui, "Mining Customer Product Ratings for Personalized Marketing", *Decision Support Systems*, Vol.35, No.2(2003), 231~243.

Getoor, L. and M. Sahami, "Using Probabilistic Relational Models for Collaborative Filtering", *Workshop on Web Usage Analysis and User Profiling*, (1999), 1~6.

Goldberg, D., D. Nichols, B. M. Oki and D. Terry, "Using Collabortive Filtering to Weave an Information Tapestry", *Communications of the ACM*, Vol.35, No.12(1992), 61~70.

Goldberg, K., T. Roeder, D. Gupta and C. Perkins, "Eigentaste: A Constant Time Collaborative Filtering Algorithm", *Information Retrieval*, Vol.4, No.2(2001), 133~151.

Hofmann, T., "Collaborative Filtering via Gaussian Probabilistic Latent Semantic Analysis", *26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, (2002), 259~266.

Hofmann, T., "Latent Semantic Models for Collaborative Filtering", *ACM Transactions on Information Systems*, Vol.22, No.1(2004), 89~115.

Huang, C. L. and W. L. Huang, "Handling Sequential Pattern Decay: Developing a Two Stage Collaborative Recommendation System", *Electronic Commerce Research and Applications*, Vol.8, No.3(2009), 117~129.

Jeong, B., J. Lee and H. Cho, "An Iterative Semi-Explicit Rating Method for Building Collaborative Recommender Systems", *Expert Systems with Applications*, Vol.36, No.3

(2009), 6181~6186.

Joaquin, D. and I. Naohiro, "Memory-based Weighted-Majority Prediction", *ACM SIGIR '99 Workshop on Recommender Systems: Algorithms and Evaluation*, (1999), 1~5.

Kim, H. N., A. T. Ji, I. Ha and G. S. Jo, "Collaborative Filtering based on Collaborative Tagging for Enhancing the Quality of Recommendation", *Electronic Commerce Research and Applications*, Vol.9, No.1(2010), 73~83.

Kim, H. K., J. K. Kim and Y. U. Ryu, "Personalized Recommendation over a Customer Network for Ubiquitous Shopping", *IEEE Transactions on Services Computing*, Vol.2, No.2(2009), 140~151.

Kim, J. K., H. K. Kim and Y. H. Cho, "A User-Oriented Contents Recommendation System in Peer-to-Peer Architecture", *Expert Systems with Applications*, Vol.34, No.1(2008), 300~312.

Kumar, R., P. Raghavan, S. Rajagopalan and A. Tomkins, "Recommender Systems: A Probabilistic Analysis", *Journal of Computer and System Science*, Vol.63, No.1(2001), 42~61.

Kwon, K., J. Cho and Y. Park, "Multidimensional Credibility Model for Neighbor Selection in Collaborative Recommendation", *Expert Systems with Applications*, Vol.36, No.3 (2009), 7114~7122.

Lang, K., "NewsWeeder: Learning to Filter Netnews", *12th International Conference on Machine Learning*, (1995), 1~9.

Lee, J. S. and S. Olafsson, "Two-way Cooperative Prediction for Collaborative Filtering Recommendations", *Expert Systems with Applications*, Vol.36, No.3(2009), 5353~5361.

Lee, T. Q., Y. Park and Y. T. Park, "A Time-based Approach to Effective Recommender Systems Using Implicit Feedback", *Expert Systems with Applications*, Vol.34, No.4 (2008), 3055~3062.

Liu, D. R., C. H. Lai and W. J. Lee, "A Hybrid of Sequential Rules and Collaborative Filtering for Product Recommendation", *Information Sciences*, Vol.179, No.20(2009), 3505~3519.

Liu Z., W. Qu, H. Li and C. Xie, "A Hybrid Collaborative Filtering Recommendation Mechanism for P2P Networks", *Future Generation Computer Systems*, Vol.26, No.8(2010), 1409~1417.

Marlin, B., "Modeling User Rating Profiles for Collaborative Filtering", *Advances in Neural Information Processing Systems*, Vol.16(2003), 627~634.

Mooney, R. J. and L. Roy, "Content-Based Book Recommending Using Learning for Text Categorization", *ACM SIGIR '99 Workshop on Recommender Systems: Algorithms and Evaluation*, (1999), 1~8.

Nakamura, A. and N. Abe, "Collaborative Filtering using Weighted Majority Prediction Algorithm", *15th International Conference on Machine Learning*, (1998), 395~403.

Park, Y. J. and K. N. Chang, "Individual and Group Behavior-based Customer Profile Model for Personalized Product Recommendation", *Expert Systems with Applications*, Vol.36, No.2(2009), 1932~1939.

Pavlov, D. Y. and D. M. Pennock, "A Maximum Entropy Approach to Collaborative Filtering in Dynamic, Sparse, High-Dimensional Domains", *Advances in Neural Information Processing Systems*, Vol.15(2002), 1441~1448.

Pazzani, M. and D. Billsus, "Learning and Revising User Profile: The Identification of

Interesting Web Sites", *Machine Learning*, Vol.27, No.3(1997), 313~331.

Pennock, D. M., E. Horvitz, S. Lawrence and C. L. Giles, "Collaborative Filtering by Personality Diagnosis: A Hybrid Memory-and Model- Based Approach", *16th Conference on Uncertainty in Artificial Intelligence*, (1999), 473~480.

Resnick, P., 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*, (1994), 175~186.

Russell, S. and V. Yoon, "Applications of Wavelet Data Reduction in a Recommender System", *Expert Systems with Applications*, Vol.34, No.4(2008), 2316~2325.

Salter, J. and N. Antonopoulos, "Cinema Screen Recommender Agent: Combining Collaborative and Content-Based Filtering", *IEEE Intelligent Systems*, Vol.21, No.1(2006), 35~41.

Salton, G., *Automatic Text Processing*, Addison-Wesley Longman Publishing Co., Inc., Boston, 1988.

Shani, G., D. Heckerman, and R. I. Brafman, "An MDP-based Recommender System", *Journal of Machine Learning Research*, Vol.6, No.2 (2002), 1265~1295.

Shardanand, U. and P. Maes, "Social Information Filtering Algorithms for Automating 'Word of Mouth'", *SIGCHI Conference on Human Factors in Computing Systems*, (1995), 210~217.

Si, L. and R. Jin, "Flexible Mixture Model for Collaborative Filtering", *20th International Conference on Machine Learning*, (2003), 1~8.

Symeonidis, P., A. Nanopoulos and Y. Manolopoulos, "Providing Justifications in Recommender Systems", *IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems an Humans*, Vol.38, No.6(2008), 1262~1272.

Tang, T. Y., P. Winoto and K. C. C. Chan, "Scaling Down Candidate Sets based on the Temporal Feature of Items for Improved Hybrid Recommendations", *Intelligent Techniques for Web Personalization*, (2005), 169~186.

Wang, Y., W. Dai and Y. Yuan, "Website Browsing Aid: A Navigation Graph-based Recommendation System", *Decision Support Systems*, Vol.45, No.3(2008), 387~400.

Wei, C. P., C. S. Yang and H. W. Hsiao, "A Collaborative Filtering-based Approach to Personalized Document Clustering", *Decision Support Systems*, Vol.45, No.3(2008), 413~428.

Yu, K., A. Schwaighofer, V. Tresp, X. Xu and H. P. Kriegel, "Probabilistic Memory-based Collaborative Filtering", *IEEE Transactions on Knowledge and Data Engineering*, Vol.16, No.1(2004), 56~69.

Yu, K., X. Xu, J. Tao, M. Ester, and H. P. Kriegel, "Instance Selection Techniques for Memory-based Collaborative Filtering", *2nd SIAM International Conference on Data Mining*, (2002), 1~16.

Abstract

통합유사도 함수의 이용과 시간정보를 고려한 협업필터링 기반의 추천시스템

최근호* · 김건우** · 유동희*** · 서용무*

상품 및 서비스에 대한 개인화된 추천 서비스가 중요해짐에 따라, 많은 연구자들은 추천시스템 개발을 위한 다양한 지식과 기법들을 제공해왔다. 이러한 기법들 중에서 협업 필터링(Collaborative Fitering) 기법은 여러 분야에서 널리 사용되고 있으며, 그 유용성이 입증되었다. 하지만, 추천시스템의 성능을 더욱 높이기 위해서 현재의 협업 필터링 기법은 다음과 같은 점들을 고려해야 한다. 첫째, 대부분의 추천시스템과 관련한 연구에서 특정 고객과 성향이 유사한 다른 고객들을 찾기 위해 사용되는 유사도 함수들(Similarity Functions)은 대부분 특정한 관점에 초점을 두고 있기 때문에 다양한 관점에서 성향이 유사한 다른 고객들을 찾는데 한계를 가진다. 따라서, 특정 관점에 치우치지 않는 통합된 유사도 함수를 사용해야 할 필요가 있다. 둘째, 고객들의 성향은 시간이 지남에 따라 변화하기 때문에, 이를 추천결과에 반영하기 위해서는 시간에 따른 고객들의 구매 성향의 변화를 고려해야 한다. 본 연구는 여러 실험들을 통해 다음의 가설을 검정하는 것을 목적으로 하였다-다양한 관점이 동시에 반영된 통합 유사도 함수의 이용과 시간정보를 이용한 사용자의 구매 성향의 변화를 반영할 경우 추천의 정확도가 향상될 것이다. 다양한 실험을 통해, 본 연구에서 제시한 추천시스템은 전통적인 협업 필터링 기반의 추천시스템들에 비해 일반적으로 상당히 높은 정확도를 보였으며 이를 통해, 본 연구에서 제시한 가설이 채택될 수 있음을 확인하였다.

Keywords : 추천시스템, 협업필터링, 시간정보, 유사도함수

* 고려대학교 경영대학 경영학과
** 한밭대학교 경상대학 경영학과
*** 육군사관학교 전자정보학과

저 자 소 개

Keunho Choi

Is a Ph.D. candidate of Korea University Business School (KUBS), South Korea. He received his M.S. in Management Information Systems from Inha University in 2008, and BSs in Material Engineering and Techno Business from Inha University in 2006. His current research interests include recommendation systems, data mining, and semantic web.

Gunwoo Kim

Is an assistant professor in department of Business Administration at Hanbat National University, South Korea. He received his Ph.D. degree in MIS from the Korea University in 2010, his Master degree and Bachelor of Science degree in MIS and computer science from Korea University and Yonsei University, Korea, in 2005 and 2002, respectively. He worked as Information System consultants in several companies from 1996 to 2003. He has published papers on the application of semantic web technology to business domain such as business process management, knowledge management, tourism. His current research interests include use of ontology in several industry areas, collaboration, recommendation system, and social network.

Donghee Yoo

Is an assistant professor of Department of Electronics Engineering & Information Science at Korea Military Academy. He received his Ph.D. degree in MS/IS from the Korea University Business School in 2009. His current research interests include the semantic web, ontology, linked data, enterprise architecture, decision support system, and recommendation system.

Yongmoo Suh

Is a professor of Korea University Business School, South Korea (KUBS). He obtained his Ph.D. from the University of Texas at Austin in 1992, MS in Computer Science from Korea Advanced Institute of Science and Technology in 1980 and BS from Seoul National University in 1978. Before going abroad for Ph.D. study, he worked as a researcher of the Korea Institute of Science and Technology. And prior to joining KUBS, he taught at Sejong University and Konkuk University in Korea. He has published papers on various areas, such as object-oriented systems, collaboration, and data mining. His current research interests include use of ontology in business domain, recommendation, and business process management.