Electronic Commerce Research and Applications 11 (2012) 309–317

Contents lists available at SciVerse ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

A hybrid online-product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis

Keunho Choi [a], Donghee Yoo [b], Gunwoo Kim [c,*], Yongmoo Suh [a]

[a] Business School, Korea University, Anam-Dong 5-Ga, Seongbuk-Gu, Seoul, Republic of Korea
[b] Department of Electronics Engineering & Information Science, Korea Military Academy, Gongneung 2-Dong, Nowon-Gu, Seoul, Republic of Korea
[c] College of Business & Economics, Hanbat National University, San 16-1, Duckmyoung-Dong, Yuseong-Gu, Daejeon, Republic of Korea

ARTICLE INFO

Article history:
Received 19 April 2011
Received in revised form 16 February 2012
Accepted 16 February 2012
Available online 24 February 2012

Keywords:
Collaborative filtering
Sequential pattern analysis
Implicit rating
Recommendation
Hybrid approach

ABSTRACT

Many online shopping malls in which explicit rating information is not available still have difficulty in providing recommendation services using *collaborative filtering* (CF) techniques for their users. Applying temporal purchase patterns derived from *sequential pattern analysis* (SPA) for recommendation services also often makes users unhappy with the inaccurate and biased results obtained by not considering individual preferences. The objective of this research is twofold. One is to derive implicit ratings so that CF can be applied to online transaction data even when no explicit rating information is available, and the other is to integrate CF and SPA for improving recommendation quality. Based on the results of several experiments that we conducted to compare the performance between ours and others, we contend that implicit rating can successfully replace explicit rating in CF and that the hybrid approach of CF and SPA is better than the individual ones.

1. Introduction

In an age of information overload, the importance of personalized recommendation systems for online products and services is rapidly growing. Such systems allow buyers to find what they want without wasting their time and also enable sellers to provide buyers with the items they are likely to purchase, thereby furnishing benefits to both parties. As a result of this growing importance, fundamental knowledge and techniques for developing recommendation systems have been studied, including content-based filtering (CBF) (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, 2002), association rule or sequential pattern analysis (Aggarwal et al. 2002, Huang and Huang 2009, Wang et al. 2008), and hybrid approaches (Balabanovic and Shoham 1998, Liu et al. 2010, Salter and Antonopoulos 2006, Wei et al. 2008).

A number of studies have attempted to resolve several typical problems of each recommendation technique such as the new user (or cold start) problem (Kim et al. 2010, Park and Chang 2009), the new item (or 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). However, there are still issues for how online shopping malls can make better recommendations for their users. This paper proposes a novel approach to improving recommendation quality and value, especially in the environment of e-commerce, focusing on the following two issues.

Online shopping malls rarely offer explicit rating information by users for items required by the CF technique (Jeong et al. 2009; Lee et al. 2010, 2008; Su et al. 2010). The technique has been widely used and has proved to be useful in practice, but it has a critical limitation – it cannot be adopted by recommendation systems for online shopping malls since explicit rating information on items is not available. So, our first research issue is:

- How can we derive implicit rating information from e-commerce transaction data that can be used for the CF technique, instead of explicit rating information?

In order for the CF technique to be used by more online shopping malls, it is necessary to derive implicit rating information from transaction data, which can be used as a proxy for explicit rating information. Based on the following two ideas, we defined a function which computes implicit ratings from the transaction data of users who purchased the same items many times: (1) a user who buys an item more than once implies the user likes it; and (2) a user who buys an item more frequently than another user implies that the former likes it more than the latter.

Many shopping malls have adopted sequential pattern analysis (SPA) to find temporal associations among items, but they may

* Corresponding author. Tel.: +82 42 821 1290; fax: +82 42 821 1288.
E-mail address: gkim@hanbat.ac.kr (G. Kim).

1567-4223/$ - see front matter
http://dx.doi.org/10.1016/j.elerap.2012.02.004

suffer from a higher probability of inaccurate and biased recommendations for items because they consider just purchasing information rather than rating information. Since purchasing an item does not always mean preferring it, the results of SPA may be useful but they will not always be useful for recommendation. This problem can be mitigated by integrating SPA with CF, which uses rating information. In contrast, CF uses only rating information of items without reflecting the changes in user preference over time. This limitation of CF can be reduced by integrating CF with SPA, which returns changes in user preferences over time in the form of sequential patterns. So, our second research issue is:

- How can we integrate CF and SPA to get better recommendations than either of the two alone?

In order to make CF and SPA supplement each other, we first calculated the predicted preference of a target user on an item to recommend from CF (*CFPP*) and from SPA (*SPAPP*). Then, we computed a linear summation of them by giving various weights to each to get a final predicted preference (*FPP*) of the target user on the item. Finally, the top *n* items with the highest *FPP* are recommended.

We implemented a Hybrid Online-Product rEcommendation (HOPE) system, which integrates CF-based recommendation using implicit ratings and SPA-based recommendations. Experiments to compare the performance of the HOPE system with those of other recommendation systems were conducted using a data set of a major online shopping mall in Korea.

The rest of this paper is organized as follows. Section 2 reviews previous works regarding recommendation systems. Section 3 describes the overall framework for realizing our approach and provides a detailed description of each step of the framework. Section 4 illustrates our approach with an example. In Section 5, we explain how much difference our approach makes based on the results from four experiments and describe the implications of each experiment. The last section contains concluding remarks, including a summary, implications, and a limitation of this research.

2. Previous works

The techniques used in most of recent recommendation systems can be generally categorized into four types: content-based filtering (CBF); collaborative filtering (CF); rule-based approaches; and hybrid approaches.

CBF recommendation systems typically: (1) construct an *item profile* by extracting a set of features from each item in the item set; (2) build a content-based *user profile* from a set of features of the items that each user purchased; (3) calculate the similarity between the user profiles and the item profiles using a specific similarity function; and (4) recommend top *n* items with high similarity scores. That is, they recommend items based on the similarity between items to recommend and items already purchased. Initially, 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). Both the user profiles and the item profiles have an associated weight 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). Other techniques, such as classification models built from a statistical approach (Mooney and Roy 1999) or a data mining approach (Pazzani and Billsus 1997), have been used to classify whether a document item is relevant to a user or not. CBF systems, however, have several limitations: (1) it is not easy to obtain 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 a target user purchased before (*over-specialization problem*) (Adomavicius and Tuzhilin 2005); and (3) new users who have not purchased items or users unusual in their preference cannot get a proper recommendation (*new or unusual user problem*) (Adomavicius and Tuzhilin 2005, Billsus et al. 2002).

CF-based recommendation systems typically: (1) build a user profile from rating information of each user on items; (2) identify like-minded users who rate items similarly to a target user using a similarity function such as cosine similarity, Pearson correlation coefficient, or distance-based similarity; and (3) recommend top *n* items that the like-minded users preferred after their ratings are predicted as an average, weighted sum or adjusted weighted sum of ratings given on items by the identified like-minded users. 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, 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, 2004; Kumar et al. 2001, Marlin 2003, Pavlov and Pennock 2002, Pennock et al. 1999, Shani et al. 2002). Many CF-based recommendation systems have been 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 recommendation systems, however, also have some limitations: (1) it is difficult to recommend items for 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 never 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).

Another simple but popular way of recommending items to a user is the rule-based approach. Rules are derived from a large transaction database collected over time, using data mining techniques. It could be either an *association rule* among items that are purchased together (Aggarwal et al. 2002) or a *sequential pattern* among items that are purchased in sequence over time (Huang and Huang 2009, Wang et al. 2008). The rule-based approach to recommending items, however, has a limitation in that it is difficult to recommend items that do not appear in association rules or sequential patterns. Aggarwal et al. (2002) proposed a technique for discovering localized association rules that 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 (archive.ics.uci.edu/ml/) 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.

Hybrid recommendation systems have been developed to overcome, or at least to mitigate, the limitations of CBF, CF, and rule-based recommendation systems (Balabanovic and Shoham 1998, Liu et al. 2009, 2010; Salter and Antonopoulos 2006, Wei et al.

2008). The Fab system (Balabanovic and Shoham 1998) combines the CF and the CBF techniques to eliminate the *insufficient features and over-specialization problems* of CBF technique and *new item problem* of CF technique. In this system, content-based user profiles are maintained to determine similar users for collaborative recommendation, and items are recommended to a target user when 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 that of the target user.

Liu et al. (2009) selected the top k neighbors from the cluster to which a target user belongs using binary choice (purchased/not-purchased) analysis of shopping basket data and derived the prediction scores of items (not yet purchased by the target user) based on the frequency count of them by scanning the purchase data of the k neighbors. Meanwhile, we selected neighbors from entire user space based on the newly derived implicit rating information of users and we derived the prediction scores of items (not yet purchased or already purchased by the target user) based on the adjusted weighted sum of ratings given on them by the k neighbors. In addition, they divided the entire time period into three periods and clustered transactions of users in each time period, and then, they derived sequential patterns represented by a sequence of transaction clusters over the three periods, while we derived sequential patterns represented by a sequence of items over the entire period. That is, they tried to find cluster sequence, while we tried to find item sequence. As such, our approach is believed to be better than theirs in that we can make a more personalized recommendation.

Depending on the techniques to be applied to recommendation system, the types of information required for making recommendations are fairly different from one another. CBF recommendation systems have used *content information of items* to build user profile and find similar items to the items that a target user purchased based on the content similarity (Albadvi and Shahbazi 2009, Lee and Kwon 2008; Liang et al. 2008, Park and Chang 2009, Ricci and Nguyen 2007, Salter and Antonopoulos 2006). On the other hand, CF recommendation systems have used *rating information of users on items* to represent the user's preference on corresponding items, and predict ratings of a target user on items based on the user similarity in ratings (Adomavicius and Kwon 2007, Goldberg et al. 2001, Jeong et al. 2009, Kwon et al. 2009, Russell and Yoon 2008). Rule-based approach has used *purchase behavior information* of users to derive meaningful association rules and sequential patterns and make recommendations based on them (Cho et al. 2002, Huang and Huang 2009, Jiang et al. 2009, Liu et al. 2009, Shih and Liu 2008).

3. Hybrid Online Product rEcommendation (HOPE) system

We developed a recommendation system, called HOPE, which integrates CF-based recommendation using implicit rating and SPA-based recommendation. This section presents the overview of the system, followed by the detailed description of each step of the framework.

3.1. System overview

Fig. 1 shows an overall framework of our recommendation system, HOPE system, which consists of two main processes: CF process and SPA process. The CF process, depicted in the upper left part of the figure, is the same as the traditional CF process, except that an implicit rating derived from transaction data of users is used instead of explicit rating. Thus, it calculates the similarity between a target user and other users using the implicit rating and selects the top k users based on the similarity score as neighbors of a target user. Finally, the predicted preferences of a target user on items purchased by the top k neighbors (*CFPP*) are calculated based on the ratings of the neighbors. The SPA process, depicted in the upper right part of Fig. 1, derives sequential patterns from transaction data of other users, and predicted preferences on items (*SPAPP*) are calculated by matching all subsequences of a target user's purchase sequence data with each derived sequential pattern. Finally, the weighted sum of normalized *CFPP* and *SPAPP* is calculated as a final predicted preference (*FPP*) on each candidate item to recommend, and then the top n items with the highest *FPP* are recommended.

3.2. Collaborative filtering-based recommendation

3.2.1. Deriving implicit ratings of users on items

It is usually difficult to obtain explicit rating information on items. In order to use the CF technique in such circumstance, this paper suggests a method of deriving implicit ratings of users on items from transaction data as an alternative to explicit ratings. The absolute preference of user u on item i, $AP(u,i)$, is computed from the following equation.

$$AP(u,i) = \ln\left(\frac{\text{The number of transactions of user } u \text{ including item } i}{\text{The number of transactions of user } u} + 1\right) \quad (1)$$

It is computed solely based on the purchase data of user u. This value, however, is far from representing the exact preference of user u on item i because it only takes into account the frequency of purchase and because the frequency is quite different depending on the item price, item lifetime, and the like. For example, since expensive items or items with long lifespan, such as jewelry or electric home appliances, are usually purchased infrequently. So the preferences of users for them cannot be higher than cheap items or those with a short lifespan, such as hand creams or tissues. Also, when a user u purchased item i four times out of ten transactions (i.e., $AP(u,i)$ is 1.4), we may think that he does not prefer item i if other users purchased the same item eight times out of ten transactions, while we may think that he prefers item i if other users purchased the same item only once. It is therefore necessary to define relative preference so it is comparable among users. The relative preference of user u on item i, $RP(u,i)$ is thus defined as in the following equation.

$$RP(u,i) = \frac{AP(u,i)}{\operatorname{Max}_{c \in U}(AP(c,i))} \quad (2)$$

where U denotes every user who purchased item i. Note that RP represents a user's preference to some extent, while AP does not. In Eq. (2), we used a maximization function in the denominator. The reason for using maximization is to make $RP(u,i)$ range from 0.0 to 1.0 (i.e., normalization).

Finally, we multiplied $RP(u,i)$ by 5 and rounded up so that implicit rating ranges from 1 to 5, as is mostly used in current recommendation systems, which is explained by the following equation.

$$\text{Implicit rating}(u,i) = \text{Round up } (5 \times RP(u,i)) \quad (3)$$

3.2.2. Calculating similarity score based on implicit ratings

With the implicit ratings of users on items, similarity between a target user and every other user is calculated, as is done in traditional CF technique, using such similarity function as 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 distance measures (Adomavicius and

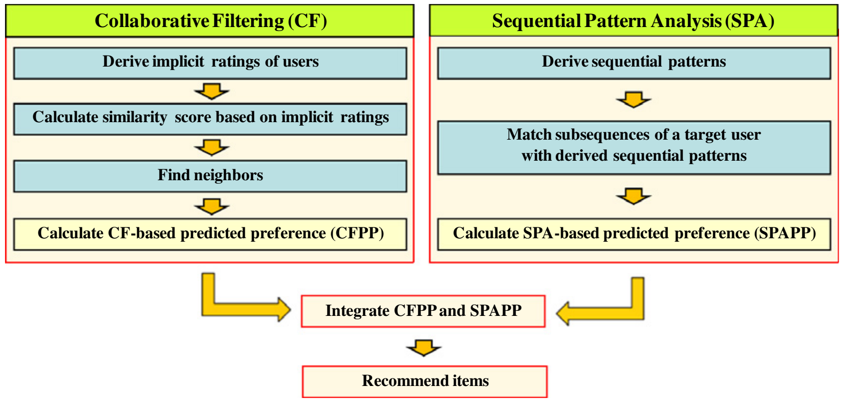

Fig. 1. Overall framework of HOPE system.

Kwon 2007, Kim et al. 2009, 2008; Park and Chang 2009). The Pearson correlation coefficient estimates the similarity based on the *rating pattern* between two users. Cosine similarity treats two users as two vectors in the m-dimensional rating vector space, where m denotes the set of all items rated by both users, and estimates the similarity by calculating the *cosine value of the angle* between the two vectors. Finally, distance measure estimates the similarity between a target user and other user by calculating the *absolute magnitude of the similarity* between two users in the m-dimensional rating vector space, so that distance-based similarity is defined as an inverse of the distance. The three similarity functions are defined in Eqs. (4)–(6) as follows:

$$\text{Pearson correlation coefficient}(a,b) = \frac{\sum_{i=1}^{m}(R_{ai}-\overline{R}_a)(R_{b,i}-\overline{R}_b)}{\sqrt{\sum_{i=1}^{m}(R_{a,i}-\overline{R}_a)^2}\sqrt{\sum_{i=1}^{m}(R_{b,i}-\overline{R}_b)^2}} \tag{4}$$

$$\text{Cosine similarity}(a,b) = \frac{\sum_{i=1}^{m}(R_{a,i})(R_{b,i})}{\sqrt{\sum_{i=1}^{m}(R_{a,i})^2}\sqrt{\sum_{i=1}^{m}(R_{b,i})^2}} \tag{5}$$

$$\text{Distance-based similarity}(a,b) = \frac{1}{1+\sqrt{\sum_{i=1}^{m}(R_{a,i}-R_{b,i})^2}} \tag{6}$$

where $R_{a,i}, R_{b,i}, \overline{R}_a$, and $\overline{R}_b$ denote the ratings of users a and b on item i, average of all $R_{a,i}$ and average of all $R_{b,i}$, respectively.

Since the above three similarity functions estimate the similarity between two users from different perspectives, depending on the similarity functions to be used, the set of neighbors whose rating information is used to predict the preference of a target user on candidate items to recommend could be different, and thus, so are the items finally recommended. It means that the choice of similarity function should be made properly based on the data set at hand. Therefore, to find a similarity function that is more appropriate for our data set, we attempted to use all these three similarity functions and compared their accuracy. Fig. 2 shows different perspectives of the three similarity functions conceptually.

3.2.3. Finding neighbors

Having calculated the similarity between a target user and every other user using each similarity function, users are sorted by similarity in descending order and then the top k users are selected as neighbors of target user a. We also changed the number of neighbors from 1 to 2 to 3 to 4 to 5 to find the appropriate number of like-minded neighbors.

3.2.4. Calculating the CF-based predicted preference (CFPP)

The rating information of the top k neighbors is then used to predict CF-based predicted preference of user a on item i, $CFPP(a,i)$, as is shown in Eq. (7).

$$CFPP(a,i) = \overline{R}_a + \frac{1}{\sum_{b=1}^{k}|sim(a,b)|} \times \sum_{b=1}^{k} sim(a,b) \times (R_{b,i}-\overline{R}_b), \tag{7}$$

where k denotes the number of user a's neighbors and $sim(a,b)$ denotes the similarity between users a and b, which is computed using Pearson correlation coefficient, cosine similarity or distance measure.

3.3. Sequential pattern analysis-based recommendation

3.3.1. Deriving sequential patterns

In order to calculating predicted preferences of items based on SPA method, sequence data of each user is generated firstly by sorting transaction data for the person according to the transaction date. Sequence data is a series of item sets, ordered by their purchase time stamp. And then, sequential patterns are derived from sequence data of users except a target user using the SPA method.

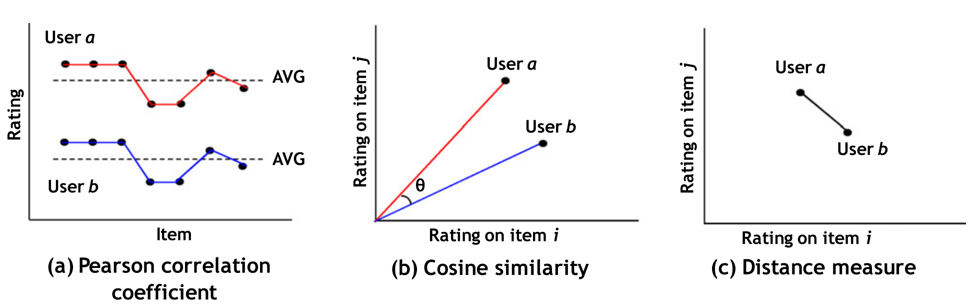

Fig. 2. Different perspectives of three similarity functions.

Table 1
An example of implicit rating data derived from an original transaction data.

	Item1		Item2		Item3		Item4		Item5	
	Date	Rating	Date	Rating	Date	Rating	Date	Rating	Date	Rating
User1	01/01	3	–	–	01/02	1	01/03	5	–	–
User2	01/01	4	–	–	01/02	3	01/03	1	01/04	2
User3	–	–	01/01	1	01/02	2	–	–	01/03	4
User4	01/01	5	01/02	4	01/03	3	–	–	–	–
User*T*	–	–	01/01	4	01/02	3	01/03	2	–	–

3.3.2. Matching subsequences of a target user with derived sequential patterns

After deriving sequential patterns, all the subsequences of a target user's sequence data are enumerated. For example, when the sequence data of target user *T* is ⟨Item1⟩⟨Item3⟩⟨Item2⟩, the possible subsequences can be ⟨Item1⟩, ⟨Item3⟩, ⟨Item2⟩, ⟨Item1⟩⟨Item3⟩, ⟨Item1⟩⟨Item2⟩, ⟨Item3⟩⟨Item2⟩, and ⟨Item1⟩⟨Item3⟩⟨Item2⟩. And then, each of these subsequences is matched with each of the derived sequential patterns to find the item(s) to recommend.

3.3.3. Calculating the SPA-based predicted preference (SPAPP)

Last, from the support of matched sequential patterns, the SPA-based predicted preference of user *a* for candidate item *i* to recommend, *SPAPP*(*a*,*i*), is calculated using the following equation.

$$SPAPP(a,i) = \sum_{s \in SUB} Support_s^i \qquad (8)$$

where *SUB* denotes the set of all subsequences of user *a*, and $Support_s^i$ denotes the support of item *i* from a subsequence *s*.

3.4. Integrating CFPP and SPAPP

CFPP and *SPAPP* are normalized to get *N_CFPP* and *N_SPAPP*, respectively, since they are different in the range of values. User *a*'s final predicted preference on item *i*, *FPP*(*a*,*i*), is calculated using the following equation:

$$FPP(a,i) = \alpha \times N\_CFPP(a,i) + (1-\alpha) \times N\_SPAPP(a,i) \qquad (9)$$

where α and $1-\alpha$ are weights given to CF technique and SPA method, respectively, and α ranges from 0.0 to 1.0.

3.5. Recommending items

After obtaining *FPP* values of items purchased by neighbors of the target user, the top *n* items are recommended. In this step, unlike usual recommendation systems, items purchased by the target user may be included in recommendation list because users may purchase the same items again.

4. An illustration with an example

This section illustrates HOPE system with a simple example. Table 1 represents an example transaction data, where it is assumed that the rating of each item by each user is calculated using Eqs. (1)–(3) from the original transaction data.

HOPE system first calculates *CFPP* values of all items by the target user *T* as follows. Neighbors of the user *T* can be identified by CF technique since the implicit rating information became available. To select the best similarity measure for this data, we have to calculate the similarity between a target user *T* and every other user using each of the three similarity measures. Cosine similarity is, however, assumed to be selected as the best similarity measure, which is thought to be sufficient for the purpose of illustration. From Table 1, the similarities between target user *T* and each one of user1, user2, user3, and user4 are 0.7071, 0.9648, 0.8944, and 1, respectively. Thus, if the number of neighbors is set to 2, the neighbors of target user *T* are users 2 and 4. Now, *CFPP*(*T*,1), *CFPP*(*T*,2), *CFPP*(*T*,3), *CFPP*(*T*,4) and *CFPP*(*T*,5) calculated using Eq. (7) are 4.7455, 3.5, 3.2365, 2, and 3. See Table 2.

HOPE system then calculates *SPAPP* values of all items by the target user *T*. The sequence data are constructed from Table 1 for the users except the target user *T*[1]:

- User1: ⟨Item1⟩⟨Item3⟩⟨Item4⟩
- User2: ⟨Item1⟩⟨Item3⟩⟨Item4⟩⟨Item5⟩
- User3: ⟨Item2⟩⟨Item3⟩⟨Item5⟩
- User4: ⟨Item1⟩⟨Item2⟩⟨Item3⟩

Suppose the minimum support 0.5 is given for sequential pattern mining. Then, the following list of sequential patterns with supports in parentheses is generated from the above sequence database:

- ⟨Item1⟩⟨Item3⟩(0.75)
- ⟨Item2⟩⟨ Item3⟩(0.5)
- ⟨Item3⟩⟨Item4⟩(0.5)
- ⟨Item3⟩⟨Item5⟩(0.5)
- ⟨Item1⟩⟨Item4⟩(0.5)
- ⟨Item1⟩⟨Item3⟩⟨Item4⟩(0.5)

Assume that the target user *T* has the sequence data ⟨Item1⟩⟨Item3⟩⟨Item2⟩ in test data. The test data belong to B ∗ D in Fig. 3. Its subsequences are ⟨Item1⟩, ⟨Item3⟩, ⟨Item2⟩, ⟨Item1⟩⟨Item3⟩, ⟨Item1⟩⟨Item2⟩, ⟨Item3⟩⟨Item2⟩ and ⟨Item1⟩⟨Item3⟩⟨Item2⟩. By matching each subsequence with the starting part of the sequential patterns, we can decide candidate items to recommend and their support. For example, since the first subsequence ⟨Item1⟩ appears in the starting part of the first and the fifth sequential patterns, Item3 and Item4 can be decided as candidate items to recommend with supports 0.75 and 0.5. Similarly from the second subsequence ⟨Item3⟩, both Item4 and Item5 are identified as candidate items to recommend with the same support 0.5 from the third and the fourth sequential patterns, respectively. After all candidate items to recommend are identified in this way, *SPAPP*(*T*,1), *SPAPP*(*T*,2), *SPAPP*(*T*,3), *SPAPP*(*T*,4), and *SPAPP*(*T*,5) are calculated using Eq. (8). They are 0, 0, 1.25(=0.75 + 0.5), 1.5(=0.5 + 0.5 + 0.5), and 0.5. See Table 2.

When the weights given to CF technique and SPA method are set to 0.1 and 0.9, after conducting Experiment 3 as is explained in Section 5, respectively, *FPP*(*T*,1), *FPP*(*T*,2), *FPP*(*T*,3), *FPP*(*T*,4) and *FPP*(*T*,5) calculated using Eq. (9) are 0.5, 0.2732, 0.6419, 0.5, and 0.3488, as shown in Table 2. Thus, if the number of items to recommend, *n*, is 3, then Item3, Item1, and Item4, which have the highest *FPP* values are selected as items to recommend by the HOPE system.

[1] Since the derived sequential patterns are to be applied to the target user, the sequence data of the target user should be excluded in process of deriving sequential patterns.

Table 2
The integration of results from CF-based and SPA-based recommendation system.

Predicted Items	CFPP	SPAPP	N_CFPP	N_SPAPP	FPP	Rank
Item1	4.7455	0	1	0	0.5	2
Item2	3.5	0	0.5463	0	0.2732	5
Item3	3.2365	1.25	0.4504	0.8333	0.6419	1
Item4	2	1.5	0	1	0.5	2
Item5	3	0.5	0.3642	0.3333	0.3488	4

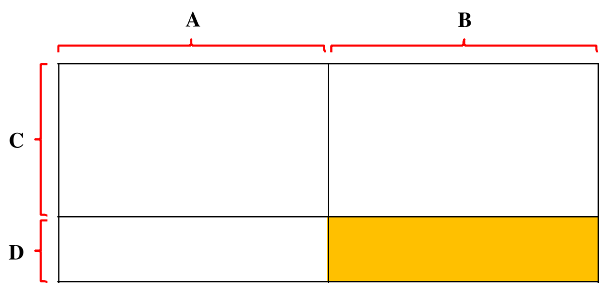

Fig. 3. Dataset partitioning.

5. Experiments

5.1. Experimental design

The data used in our experiment were provided by one of the three largest online shopping malls in Korea, after removing confidential information. The entire data set was collected from August 16, 2008 through August 15, 2009 (12 months) in four tables, for customers, sellers, products, and purchases. Since most recommendation systems have difficulty in recommending items to users who are involved in a small number of transactions, we focused on the users who have purchased more than 30 times among total 1000 users. As a result, 16,486 transactions of 247 users on 1911 items were used in our experiments.

Before conducting our experiments, we partitioned our data set into four parts, as shown in Fig. 3. It was partitioned first by time into Part A and B, and second by random sampling of users into Parts C and D. Part A consists of transaction data collected during the first 6 months and Part B during the second 6 months. Part C consists of transaction data from 70% of the users, randomly chosen, and Part D the transaction data of the remaining users.

When making recommendations using CF technique, we used Part $A * (C + D)$ in order to calculate the similarity between each target user and every other user, and recommended items for the target users in Part $B*D$. On the other hand, when making recommendations using the SPA method, we used Part $A * C$ in order to derive sequential patterns, and recommended items for the users in Part $B * D$.

In our experiment, as a means to measure the quality of recommendation, we also used precision, recall and F1, as used elsewhere to evaluate and compare the accuracy of recommendations (Albadvi and Shahbazi 2009, Chen et al. 2008, Huang and Huang 2009, Kim et al. 2009, Kwon et al. 2009, Lee et al. 2010, 2009; Park and Chang 2009, Symeonidis et al. 2008). F1 is the harmonic average of precision and recall.

In order to ensure that our proposed technique actually improves the accuracy of recommendation system, we implemented several other recommendation systems using Transact-SQL on Microsoft SQL Server 2008 and SAS 9.1 Enterprise Miner. We classified them into the following three groups:

(1) Three CF-based recommendation systems, CF_P, CF_C and CF_D: these systems compute similarity using Pearson correlation coefficient, cosine similarity, and distance-based similarity.
(2) Three SPA-based recommendation systems, SPA1, SPA2, and SPA3: these systems recommend items from derived sequential patterns when the minimum support is 1%, 2%, and 3%.
(3) Our proposed system, HOPE: this system integrates CF technique and SPA method.

5.2. Experimental results and implications

In this section, we explain how the ideas proposed in Section 3 affect the accuracy of recommendations by analyzing the results from four experiments, and describe the implications of each idea.

5.2.1. Experiment 1

The objective of this experiment is to find a similarity function appropriate to our dataset. As mentioned in Section 3.2.2, according to the similarity functions (i.e., Pearson correlation coefficient, cosine similarity, and distance-based similarity) to be used, the set of neighbors could be different and so are the items finally recommended. Therefore, the choice of similarity function should be made appropriately based on the characteristics of the online shopping mall dataset at hand. To do this, we used all the three similarity functions and compared the accuracy of three CF systems that belong to group A.

Each CF system was executed with the number of neighbors set to 1, 2, 3, 4, or 5 and with the number of recommended items set to 10, 20, 30, 40 or 50. In all cases, CF_P outperformed both CF_C and CF_D. From the results, we can conclude that Pearson correlation coefficient is more appropriate as a similarity function for our online shopping mall data set. The accuracy of CF_P increases when the number of neighbors increases from 1 to 3, while it decreases when the number of neighbors is 4 or 5. The best accuracy of CF_P was obtained when the number of recommended items is 50, as is shown in Fig. 4. Increasing the size of recommendation leads to an increase in recall but a decrease in precision, because items with lower predicted ratings are likely to be included in a recommendation list as the number of recommended items increases.

However, on the contrary, the precision of CF_C and CF_D increases when the number of recommendations increases from 10 to 30 and from 10 to 40, while it decreases when the number of recommendations increases from 30 to 50 and from 40 to 50. The precision of CF_P decreases when the number of recommendations increases from 10 to 20, while it increases when the number of recommendations increases from 20 to 50 in our experiment. As a reason for this, we presume that items recommended to users may be somewhat different from those accepted by them in practice. This phenomenon also was reported to take place in other studies (Chen et al. 2008, Zenebe and Norcio 2009, Zhen et al. 2010).

As such, Experiments 2, 3, and 4 were carried out with the same setting (i.e., the number of neighbors is 3 and the number of recommended items is 50). Note that F1 values of CF_P, CF_C, and CF_D increase as the number of recommended items increases.

5.2.2. Experiment 2

The objective of this experiment is to identify the best minimum support for sequential pattern mining. Depending on the minimum support, the derived sequential patterns and the number of them could be different and so are the items finally recommended. Thus, it is important to determine appropriate minimum support depending on the data set at hand. So, we attempted to use

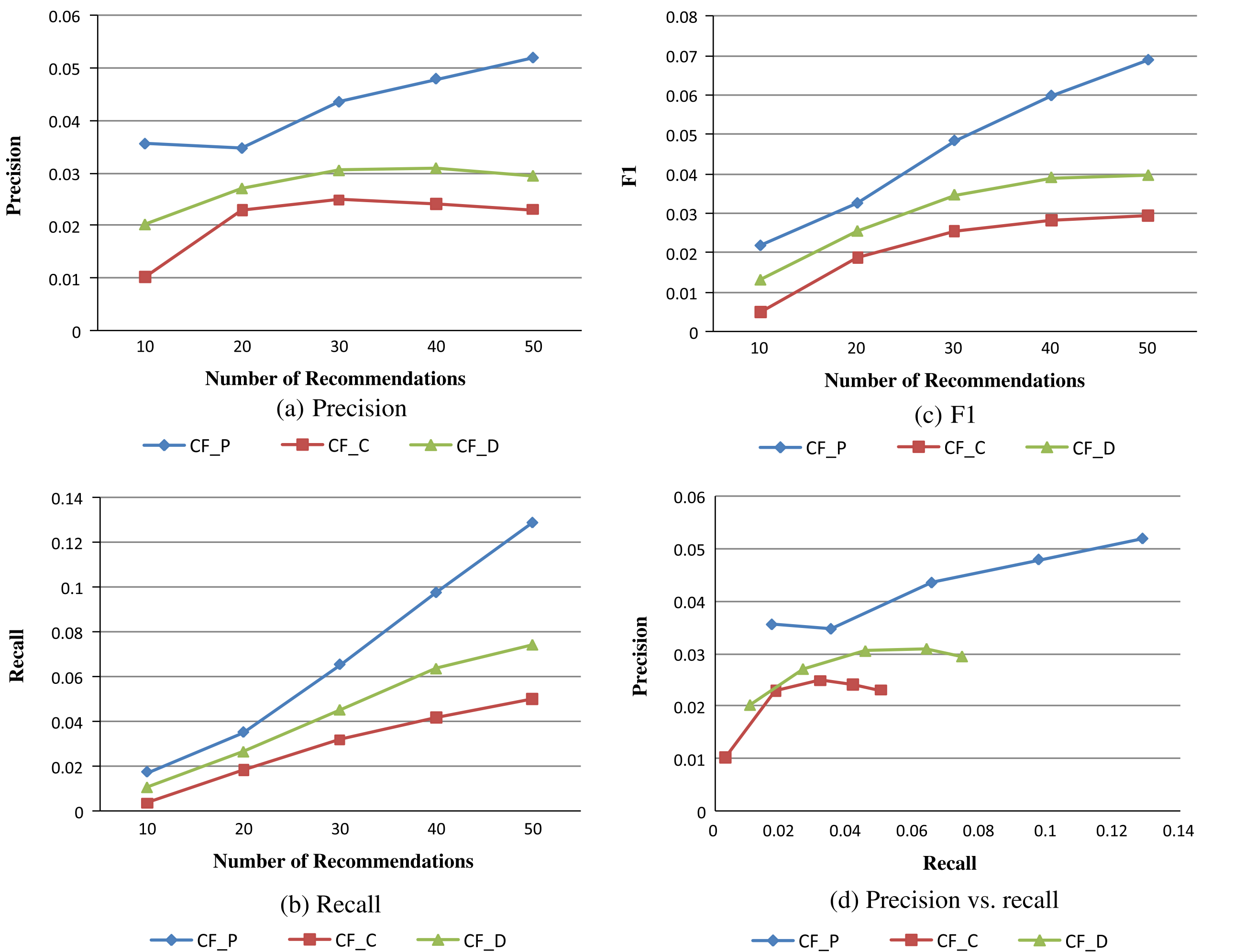

Fig. 4. Comparison of precision, recall, and F1 among CF-based recommendation systems using different similarity function, when the number of neighbors is set to 3.

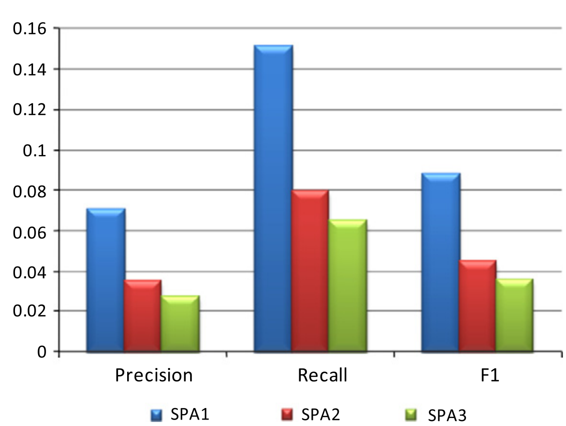

Fig. 5. Comparison of precision, recall, and F1 among SPA-based recommendation systems, each of which is implemented with different minimum support.

different minimum support.[2] When the minimum support was 1% (SPA1), 2% (SPA2), and 3% (SPA3), the number of derived sequential patterns was 497, 146, and 64. As shown in Fig. 5, SPA1 outperformed SPA2 and SPA3, and SP2 outperformed SP3, in all measures of precision, recall, and F1. This experimental result indicates that recommending items based on more sequential patterns gives better performance. Therefore, this minimum support (i.e., 1%) was used for Experiments 3 and 4.

5.2.3. Experiment 3

The objective of this experiment is to find a pair of weights that are to be given to the results of CF-based and SPA-based recommendations. Since the accuracy of the final recommendation can be affected by the way of integrating the results of CF-based and SPA-based recommendations, we should determine the suitable weights to be given to the results of the two different recommendations.

To evaluate the effect of integration weight on the accuracy of hybrid recommendation system, we attempted to measure the performance of HOPE system, while changing the weight α given to CF technique from 0.05 to 0.9. Fig. 6 shows that precision, recall, and F1 of the system increase as α increases from 0.05 to 0.1, but they decrease as α increases from 0.1 to 0.9. That is, when $\alpha = 0.1$, the system shows the best performance. Therefore, the weights given to CF-based and to SPA-based recommendation were set to 0.1 and 0.9, respectively. This result indicates that SPA contributes more than CF when making final recommendation.

5.2.4. Experiment 4

This is the main experiment in which we want to compare the performance of our hybrid recommendation system, HOPE system,

[2] When the minimum support is smaller than 1%, all sequence of purchase are regarded as sequential patterns even though they appear only once. Therefore, we did not use minimum support smaller than 1%.

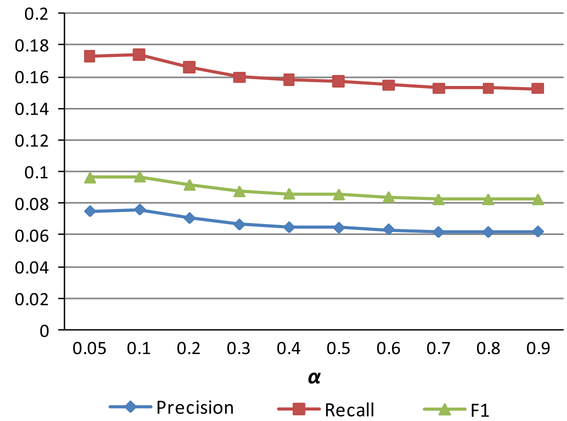

Fig. 6. Adjustment of integration weight.

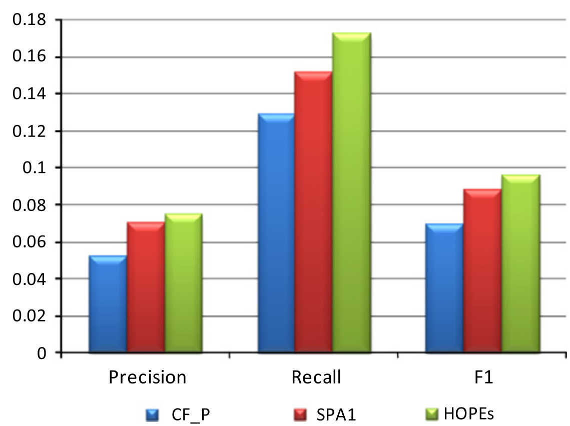

Fig. 7. Comparison of precision, recall, and F1 among three approaches.

with those of CF_P and SPA1 with regard to precision, recall, and F1. As shown in Fig. 7, HOPE system outperformed CF_P and SPA1 in all measures of precision, recall, and F1, while SPA1 outperformed CF_P recommendation system in all measures of precision, recall, and F1. This indicates that considering both users' preference on items and sequential relationships between items is effective on improving the accuracy of recommendation.

Finally, in order to test whether the differences in precision, recall, and F1 between HOPE system and the other two recommendation systems are statistically significant, we conducted the paired T-test. The result showed that the differences between HOPE system and CF_P were statistically significant ($p < 0.001$) in all measures of precision, recall, and F1. However, the difference of precision between HOPE system and SPA1 was not statistically significant, while the difference of recall between the two systems was statistically significant ($p < 0.05$) and so was that of F1 ($p < 0.1$).

In summary, by using implicit ratings of users on items derived from transaction data, we could successfully apply CF technique to our data set, and by integrating implicit rating-based CF technique and SPA method, we were able to increase the accuracy of recommendation.

6. Conclusion

CF techniques have been used successfully to recommend various items such as movie and document. But it requires that there should be many users who rated many items, so that items that are rated high by like-minded users of a user can be recommended to the user. Therefore, it has inherent problems such as the new user problem, the new item problem and the sparsity problem. As mentioned earlier, these problems can be removed by the hybrid of content-based technique, collaborative filtering technique, etc.

Since it is not easy to get rating information on items more often than not, it is important to examine a new approach to applying the CF technique when there is no such information. Therefore, rather than suggesting a new approach to resolving these inherent problems we proposed a new CF-based approach which makes use of implicit rating information instead of explicit rating information that is required by the original CF technique, and we explained how implicit rating information can be computed from the transaction dataset. We also suggested an approach to obtaining better recommendation quality by integrating CF and SPA each of which considers the rating information of users on items and the associations among items, and proved it through our experiments.

We conducted four experiments, which led to the following conclusions. First, the implicit rating derived from purchase data works well as an alternative of explicit rating, especially when Pearson correlation coefficient is used as a similarity function for CF-based recommendation. But it should be noted that the best similarity function to be used in CF-based recommendation depends on the dataset to be analyzed. Second, proper minimum support should be input when conducting sequential pattern mining. In our dataset, SPA-based recommendation shows the better performance when the minimum support is set to 1%, than the other cases in which the minimum support is set to 2% or 3%. Third, it is important to use appropriate weights when integrating two things by linear sum of them. After several attempts to integrate CF-based recommendation and SPA-based recommendation, the weight given to the CF-based recommendation is determined to 0.1, Finally, our proposed recommendation system, HOPE system, outperforms the CF-based and SPA-based recommendation systems in all measures of precision, recall and F1, and the differences in precision, recall and F1 are statistically significant in most cases but the difference in precision between HOPE system and SPA-based recommendation is not. The mining approach assumes that the volume of the data set to be analyzed is huge. But our dataset is not the case, which is a limitation.

Acknowledgment

This research is supported by 2011 academic research funding from the Hanbat National University.

References

Adomavicius, G., and Kwon, Y. New recommendation techniques for multicriteria rating systems. *IEEE Intelligent Systems*, 22, 3, 2007, 48–55.

Adomavicius, G., and Tuzhilin, A. 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*, 17, 6, 2005, 734–749.

Aggarwal, C. C., Procopiuc, C., and Yu, P. S. Finding localized associations in market basket data. *IEEE Transactions on Knowledge and Data Engineering*, 14, 1, 2002, 51–62.

Albadvi, A., and Shahbazi, M. A hybrid recommendation technique based on product category attributes. *Expert Systems with Applications*, 36, 9, 2009, 11480–11488.

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

Balabanovic, M., and Shoham, Y. Content-based, collaborative recommendation. *Communications of the ACM*, 40, 3, 1998, 66–72.

Belkin, N. J., and Croft, W. B. Information filtering and information retrieval: two sides of the same coin. *Communications of the ACM*, 35, 12, 1992, 29–38.

Billsus, D., and Pazzani, M. J. Learning collaborative information filters. In *Proceedings of the Fifteenth International Conference on Machine Learning*, 1998.

Billsus, D., Brunk, C. A., Evans, C., Gladish, B., and Pazzani, M. Adatpvie interfaces for ubiquitous web access. *Communications of the ACM*, 45, 5, 2002, 34–38.

Chen, L. S., Hsu, F. H., Chen, M. C., and Hsu, Y. C. Developing recommender systems with the consideration of product profitability for sellers. *Information Sciences*, 178, 4, 2008, 1032–1048.

Cheung, K. W., Kwok, J. T., Law, M. H., and Tsui, K. C. Mining customer product ratings for personalized marketing. *Decision Support Systems*, 35, 2, 2003, 231–243.

Cho, Y. H., Kim, J. K., and Kim, S. H. A personalized recommender system based on web usage mining and decision tree induction. *Expert systems with Applications*, 23, 3, 2002, 329–342.

Getoor, L., and Sahami, M. Using probabilistic relational models for collaborative filtering. In *Workshop on Web Usage Analysis and User Profiling* (WEBKDD'99), 1999.

Goldberg, D., Nichols, D., Oki, B. M., and Terry, D. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35, 12, 1992, 61–70.

Goldberg, K., Roeder, T., Gupta, D., and Perkins, C. Eigentaste: a constant time collaborative filtering algorithm. *Information Retrieval*, 4, 2, 2001, 133–151.

Hofmann, T. Collaborative filtering via gaussian probabilistic latent semantic analysis. In *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2003, 259–266.

Hofmann, T. Latent semantic models for collaborative filtering. *ACM Transactions on Information Systems*, 22, 1, 2004, 89–115.

Huang, C. L., and Huang, W. L. Handling sequential pattern decay: developing a two-stage collaborative recommendation system. *Electronic Commerce Research and Applications*, 8, 3, 2009, 117–129.

Jeong, B., Lee, J., and Cho, H. An iterative semi-explicit rating method for building collaborative recommender systems. *Expert Systems with Applications*, 36, 3, 2009, 6181–6186.

Jiang, Y., Shang, J., and Liu, Y. Maximizing customer satisfaction through an online recommendation system: a novel associative classification model. *Decision Support Systems*, 48, 3, 2009, 470–479.

Joaquin, D., and Naohiro, I. Memory-based weighted-majority prediction. In *ACM SIGIR' 99 Workshop on Recommender Systems: Algorithms and, Evaluation*, 1999.

Kim, H. N., Ji, A. T., Ha, I., and Jo, G. S. Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. *Electronic Commerce Research and Applications*, 9, 1, 2010, 73–83.

Kim, H. K., Kim, J. K., and Ryu, Y. U. Personalized recommendation over a customer network for ubiquitous shopping. *IEEE Transactions on Services Computing*, 2, 2, 2009, 140–151.

Kim, J. K., Kim, H. K., and Cho, Y. H. A user-oriented contents recommendation system in peer-to-peer architecture. *Expert Systems with Applications*, 34, 1, 2008, 300–312.

Kumar, R., Raghavan, P., Rajagopalan, S., and Tomkins, A. Recommender systems: a probabilistic analysis. *Journal of Computer and System Science*, 63, 1, 2001, 42–61.

Kwon, K., Cho, J., and Park, Y. Multidimensional credibility model for neighbor selection in collaborative recommendation. *Expert Systems with Applications*, 36, 3, 2009, 7114–7122.

Lang, K. NewsWeeder: learning to filter netnews. In *Proceedings of the Twelfth International Conference on, Machine Learning*, 1995.

Lee, J. S., and Olafsson, S. Two-way cooperative prediction for collaborative filtering recommendations. *Expert Systems with Applications*, 36, 3, 2009, 5353–5361.

Lee, K. C., and Kwon, S. Online shopping recommendation mechanism and its influence on consumer decisions and behavior: a causal map approach. *Expert Systems with Applications*, 35, 4, 2008, 1567–1574.

Lee, S. K., Cho, Y. H., and Kim, S. H. Collaborative filtering with ordinal scale-based implicit ratings for mobile music recommendations. *Information Sciences*, 180, 11, 2010, 2142–2155.

Lee, T. Q., Park, Y., and Park, Y. T. A time-based approach to effective recommender systems using implicit feedback. *Expert Systems with Applications*, 34, 4, 2008, 3055–3062.

Liang, T. P., Yang, Y. F., Chen, D. N., and Ku, Y. C. A semantic-expansion approach to personalized knowledge recommendation. *Decision Support Systems*, 45, 3, 2008, 401–412.

Liu, D. R., Lai, C. H., and Lee, W. J. A hybrid of sequential rules and collaborative filtering for product recommendation. *Information Sciences*, 179, 20, 2009, 3505–3519.

Liu, Z., Qu, W., Li, H., and Xie, C. A hybrid collaborative filtering recommendation mechanism for P2P networks. *Future Generation Computer Systems*, 26, 8, 2010, 1409–1417.

Marlin, B. Modeling user rating profiles for collaborative filtering. *Advances in Neural Information Processing Systems*, 16, 2003, 627–634.

Mooney, R. J., and Roy, L. Content-based book recommending using learning for text categorization. In *ACM SIGIR '99 Workshop on Recommender Systems: Algorithms and Evaluation, Berkeley, CA*, 1999.

Nakamura, A., and Abe, N. Collaborative filtering using weighted majority prediction algorithm. In 15th *International Conference of Machine Learning (ICML '98), Madison, Wisconsin, USA*, 1998, 395–403.

Park, Y. J., and Chang, K. N. Individual and group behavior-based customer profile model for personalized product recommendation. *Expert Systems with Applications*, 36, 2, 2009, 1932–1939.

Pavlov, D. Y., and Pennock, D. M. A maximum entropy approach to collaborative filtering in dynamic, sparse, high-dimensional domains. *Advances in Neural Information Processing Systems*, 15, 2002, 1441–1448.

Pazzani, M., and Billsus, D. Learning and revising user profile: the identification of interesting web sites. *Machine Learning*, 27, 3, 1997, 313–331.

Pennock, D. M., Horvitz, E., Lawrence, S., and Giles, C. L. Collaborative filtering by personality diagnosis: a hybrid memory- and model-based approach. In *Proceedings of the 16th Conference on Uncertainty in, Artificial Intelligence*, 1999, 473–480.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. GroupLens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, 1994, 175–186.

Ricci, F., and Nguyen, Q. N. Acquiring an revising preferences in a critique-based mobile recommender system. *IEEE Intelligent Systems*, 22, 3, 2007, 22–29.

Russell, S., and Yoon, V. Applications of wavelet data reduction in a recommender system. *Expert Systems with Applications*, 34, 4, 2008, 2316–2325.

Salter, J., and Antonopoulos, N. Cinema screen recommender agent: combining collaborative and content-based filtering. *IEEE Intelligent Systems*, 21, 1, 2006, 35–41.

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

Shani, G., Heckerman, D., and Brafman, R. I. An MDP-based recommender system. *Journal of Machine Learning Research*, 6, 2, 2002, 1265–1295.

Shardanand, U., and Maes, P. Social information filtering algorithms for automating "word of mouth". In *Proceedings of the SIGCHI Conference on Human Factors in, Computing Systems*, 1995, 210–217.

Shih, Y. Y., and Liu, D. R. Product recommendation approaches: collaborative filtering via customer lifetime value and customer demands. *Expert Systems with Applications*, 35, 1–2, 2008, 350–360.

Si, L., and Jin, R. Flexible mixture model for collaborative filtering. In *Proceedings of 20th International Conference on, Machine Learning*, 2003.

Su, J. H., Wang, B. W., Hsiao, C. Y., and Tseng, V. S. Personalized rough-set-based recommendation by integrating multiple contents and collaborative information. *Information Sciences*, 180, 1, 2010, 113–131.

Symeonidis, P., Nanopoulos, A., and Manolopoulos, Y. Providing justifications in recommender systems. *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems an Humans*, 38, 6, 2008, 1262–1272.

Wang, Y., Dai, W., and Yuan, Y. Website browsing aid: a navigation graph-based recommendation system. *Decision Support Systems*, 45, 3, 2008, 387–400.

Wei, C. P., Yang, C. S., and Hsiao, H. W. A collaborative filtering-based approach to personalized document clustering. *Decision Support Systems*, 45, 3, 2008, 413–428.

Yu, K., Schwaighofer, A., Tresp, V., Xu, X., and Kriegel, H. P. Probabilistic memory-based collaborative filtering. *IEEE Transactions on Knowledge and Data Engineering*, 16, 1, 2004, 56–69.

Yu, K., Xu, X., Tao, J., Ester, M., and Kriegel, H. P. Instance selection techniques for memory-based collaborative filtering. In *Proceeding of the Second SIAM International Conference on Data Mining (SDM' 02)*, 2002.

Zenebe, A., and Norcio, A. F. Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems. *Fuzzy Sets and Systems*, 160, 1, 2009, 76–94.

Zhen, L., Jiang, Z., and Song, H. Distributed recommender for peer-to-peer knowledge sharing. *Information Sciences*, 180, 18, 2010, 3546–3561.