Available online at www.sciencedirect.com

ScienceDirect

Expert Systems with Applications 35 (2008) 350–360

Expert Systems with Applications

www.elsevier.com/locate/eswa

Product recommendation approaches: Collaborative filtering via customer lifetime value and customer demands

Ya-Yueh Shih [a,*], Duen-Ren Liu [b]

[a] Department of Information Management, Chung Hua University, Taiwan
[b] Institute of Information Management, National Chiao Tung University, Taiwan

Abstract

Recommender systems are techniques that allow companies to develop one-to-one marketing strategies and provide support in connecting with customers for e-commerce. There exist various recommendation techniques, including collaborative filtering (CF), content-based filtering, WRFM-based method, and hybrid methods. The CF method generally utilizes past purchasing preferences to determine recommendations to a target customer based on the opinions of other similar customers. The WRFM-based method makes recommendations based on weighted customer lifetime value – Recency, Frequency and Monetary. This work proposes to use customer demands derived from frequently purchased products in each industry as valuable information for making recommendations. Different from conventional CF techniques, this work uses extended preferences derived by combining customer demands and past purchasing preferences to identify similar customers. Accordingly, this work proposes several hybrid recommendation approaches that combine collaborative filtering, WRFM-based method, and extended preferences. The proposed approaches further utilize customer demands to adjust the ranking of recommended products to improve recommendation quality. The experimental results show that the proposed methods perform better than several other recommendation methods.

Keywords: Recommender systems; Collaborative filtering; Content-based filtering; WRFM-based CF method

1. Introduction

Recommender systems have emerged in e-commerce applications to support product recommendation (Kim, Lee, Shaw, Chang, & Nelson, 2001; Schafer, Konstan, & Riedl, 2001; Zeng, Xing, Zhou, & Zheng, 2004), which provide individual marketing decisions for each customer. They assist businesses in implementing one-to-one marketing strategies, relying on customer purchase history to reveal customer preferences and identify products that customers may purchase. One-to-one marketing introduces a fundamental new basis for competition in the marketplace by enabling organizations to differentiate based on customers rather than products (Peppers & Rogers, 1993). Schafer et al. (2001) presented a detailed taxonomy of e-commerce recommender systems, and elucidated how they can provide personalization to establish customer loyalty. Generally, such systems offer several advantages, including increasing the probability of cross-selling, establishing customer loyalty, and fulfilling customer needs by presenting products of possible interest to them.

Various recommendation methods have been proposed. The collaborative filtering (CF) method has been successfully used in various applications. It predicts user preferences for items in a word-of-mouth manner. User preferences are predicted by considering the opinions (in the form of preference ratings) of other "like-minded" users. The GroupLens system (Resnick, Iacovou, Suchak, Bergstrom, & Riedl, 1994) applied the CF method to recommend Usenet News and movies. Video recommender (Hill, Stead, Rosenstein, & Furnas, 1995) also used CF to generate recommendations on movies. Examples of music

* Corresponding author.
E-mail address: moon.shih@msa.hinet.net (Y.-Y. Shih).

0957-4174/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2007.07.055

recommender systems are Ringo (Shardanand & Maes, 1995) and MRS (Chen & Chen, 2001). Siteseer Rucker and Polanco (1997) provided Web page recommendations based on bookmarks of user's virtual neighbors, Amazon.com uses collaborative filtering to create books recommendations for customers (Linden, Smith, & York, 2003). Collaborative filtering requires a user to rate a reasonably large set of items, or the CF method has difficulty providing recommendations to novices (new users). Moreover, the CF method may suffer the sparsity problem, a situation in which transactional data is sparse and insufficient to identify similarities in user interests (Sarwar, Karypis, Konstan, & Riedl, 2000).

Firms increasingly recognize the importance of customer lifetime value (CLV) (Berger & Nasr, 1998). Generally, RFM (Recency, Frequency, and Monetary) method has been used to measure CLV (Kahan, 1998; Miglautsch, 2000). Identifying CLV or loyalty ranking of customer segments is important for helping decision-makers target markets more clearly in fiercely competitive environments. Additionally, the effect of CLV on recommendations should be investigated to make more effective marketing strategies. Recently, a weighted RFM-based CF method (WRFM-based CF method) (Liu & Shih, 2005b) has been proposed that integrates analytic hierarchy process (AHP) (Saaty, 1994) and data mining to recommend products based on customer lifetime value. This method employs association rule mining to identify recommendation rules from customer groups that are clustered according to weighted RFM values. Their experimental result demonstrated that the WRFM-based CF method can identify effective rules for making recommendations to customers with high lifetime value or loyalty. The WRFM-based CF method also suffers the sparsity problem.

The content-based filtering (CBF) offers a different approach to collaborative filtering and provides recommendations by matching customer profiles (e.g., interests) with content features (e.g., product attributes). Each customer profile is derived by analyzing the content features of products purchased by the customer. The simplest of these techniques is keyword matching (Claypool, Gokhale, & Miranda, 1999). Krakotoa Chronicle (Kamba, Bharat, & Albers, 1995) is an example of such system. However, the CBF method is limited in not being able to provide serendipitous recommendations, because the recommendation is based solely on the content features of products purchased by the customer. Some domains, such as music recommendations, have difficulty analyzing content features of products.

Several researchers are exploring hybrid methods of combining CF and CBF methods to smooth out the disadvantages of each (Basu, Hirsh, & Cohen, 1998; Claypool et al., 1999; Good et al., 1999). This work uses customer demands derived from the frequently purchased products in each industry as valuable information to integrate the CF method for making recommendations. Extended preferences derived by combining customer demands and past purchasing preferences are used to alleviate the sparsity problem of recommendation. Different from conventional CF techniques, this work uses extended preferences to identify similar customers. Accordingly, this work proposes several hybrid recommendation approaches that combine collaborative filtering, WRFM-based method, and extended preferences. Moreover, customer demands are considered in re-ranking recommended products to improve the quality of recommendation.

The remainder of this paper is organized as follows. Section 2 reviews related works on the typical KNN-based CF method, the WRFM-based method, hybrid works, and content-based filtering methods. Next, Section 3 outlines the proposed methods. Section 4 then describes the experimental setup and criteria to evaluate recommendation quality. Experimental results are also presented to confirm differences between methods. Finally, Section 5 draws conclusions, summarizing the contributions of this work and outlining areas for further research.

2. Recommendation methods and related works

2.1. Typical KNN-based collaborative filtering

Collaborative recommendation (or collaborative filtering) predicts user preferences on items in a word-of-mouth manner. Similarity measures between user preference ratings are derived to define the like-mindedness between users (Breese, Heckerman, & Kadie, 1998). Notably, preferences generally are defined in terms of customer purchasing behavior, namely, purchased/non-purchased (binary choice) of shopping basket data, or taste, namely, preference rating on product items. This work focused on product recommendation of retail transaction data that contains binary choice of shopping basket data.

A typical KNN-based collaborative filtering (CF) method (Resnick et al., 1994; Shardanand & Maes, 1995) employs nearest-neighbor algorithm to recommend products to a target customer u based on the preferences of *neighbors*, that is, those customers having similar preferences as customer u. The typical KNN-based CF method is detailed as follows (Sarwar et al., 2000). Customer preferences, namely, customer purchase history, are represented as a customer-item matrix $\boldsymbol{R}$ such that, r_{ij} is one if the ith customer had purchased the jth product, and is zero otherwise. The similarity among customers can be measured in various ways.

A common method is to compute the Pearson correlation coefficient defined as Eq. (1):

$$corr_P(c_i, c_j) = \frac{\sum_{s \in I}(r_{c_i,s} - \bar{r}_{c_i})(r_{c_j,s} - \bar{r}_{c_j})}{\sqrt{\sum_{s \in I}(r_{c_i,s} - \bar{r}_{c_i})^2 \sum_{s \in I}(r_{c_j,s} - \bar{r}_{c_j})^2}} \tag{1}$$

The notations $\bar{r}_{c_i}$ and $\bar{r}_{c_j}$ denote the average number of products purchased by customers c_i and c_j, respectively. Moreover, the variable I denotes the set of products.

Additionally, the $r_{c_i,s}$ and $r_{c_j,s}$ indicate whether customers c_i and c_j purchased product item s.

Customers are ranked by their similarity measures in relation to the target customer u, as determined using the Pearson correlation coefficient. The k most similar (highest ranked) customers are selected as the k-nearest neighbors of customer u. Finally, the top-N recommended products are determined from the k-nearest neighbors of u, as follows. The frequency count of products is calculated by scanning the purchase data of the k-nearest neighbors. The products then are sorted based on frequency count. The N most frequent products that have not yet been purchased by target customer u are selected as the top-N recommendations.

2.2. Weighted RFM-based CF method

This method (Liu & Shih, 2005b) primarily integrated AHP, clustering, and association rule mining techniques for product recommendation. It employs the AHP to evaluate the weighting (relative importance) of each RFM variable, and specifically asks decision-makers to make intuitive judgments about ranking ordering to make pairwise comparisons. The RFM values of each customer are normalized as follows. The profit form, $x' = (x - x^S)/(x^L - x^S)$,was used to normalize the F (frequency) and M (monetary) values, since F and M positively influenced CLV or loyalty. The cost form, $x' = (x^L - x)/(x^L - x^S)$, was used to normalize the R value, since it negatively impacted CLV. x' represented the normalized and original R (F, M) values, while x^L and x^S represented the largest and smallest R (F, M) value of all customers. The normalized RFM values of each customer are then multiplied by the relative importance of RFM variables, w_R, w_F and w_M, which are determined by the AHP. The similarity among customers is measured by computing the Pearson correlation coefficient based on the weighted RFM values of customers, defined as Eq. (2):

$$corr_{\mathrm{WRFM}}(c_i, c_j) = \frac{\sum_{s\in V}(WRFM_{c_i,s} - \overline{WRFM}_{c_i})(WRFM_{c_j,s} - \overline{WRFM}_{c_j})}{\sqrt{\sum_{s\in V}(WRFM_{c_i,s} - \overline{WRFM}_{c_i})^2 \sum_{s\in V}(WRFM_{c_j,s} - \overline{WRFM}_{c_j})^2}} \tag{2}$$

In Eq. (2), $\overline{WRFM}_{c_i}$ and $\overline{WRFM}_{c_j}$ are the average weighted RFM (WRFM) value of customer c_i and c_j, respectively. The variable V denotes the set of RFM variables. The variables $WRFM_{c_i,s}$ and $WRFM_{c_j,s}$ indicate the weighted value R (F or M) of customer c_i and c_j, respectively.

K-means clustering is employed to group customers with similar lifetime value or loyalty based on weighted RFM. Association rule based recommendation (Sarwar et al., 2000) is then employed to recommend product items for each customer group as follows. Association rule mining is used to extract the recommendation rule set RS_j from transactions associated with each cluster, rather than from all customer transactions. The cluster C_j to which a customer, u, belongs is first identified. Then, RS_j, the recommendation rule set extracted from C_j is used to select the top-N candidate products to be recommended to customer u. Let X_u represent the set of products previously purchased by customer u. For each recommendation rule $X \Rightarrow Y$ in RS_j, if $X \subseteq X_u$, then all products in $Y - X_u$ are candidate products for recommendation to customer u. Each candidate product is associated with the confidence value of the corresponding recommendation rule. The candidate products are sorted by associated confidence value, where the N highest ranked candidate products are selected as the recommendation set.

2.3. Preference-based CF method

Preference-based CF method is similar to the WRFM-based CF method, except that clustering is performed based on purchase preferences. The method first constructs a customer-item matrix $\boldsymbol{R}$ such that, r_{ij} is one if the ith customer purchased the jth product, and is zero otherwise. The Pearson correlation coefficient (Eq. (2)) is used to compute the similarity of preference among customers. K-means clustering is then employed to group customers with similar purchase preferences. Finally, association rule mining is used to extract recommendation rules from each customer group. Candidate products for recommendation are selected and ranked based on the recommendation rules, as described in Section 2.2.

2.4. Combining WRFM-based method and preference-based CF method

The core concept of the WRFM-based CF method is to group customers based on weighted RFM (CLV), and then extract recommendation rules from each customer group. Furthermore, the preference-based CF is similar to the WRFM-based method except that the preference-based method groups customers by purchase preferences. The two methods consider either CLV or purchase preferences separately. A hybrid method that groups customers by considering both the WRFM values and purchase preferences can improve the quality of recommendations (Liu & Shih, 2005a). The hybrid method is named WRFMCP method in this paper. The WRFMCP method conducts association rule based recommendations by extracting recommendation rules from customer groups clustered according to the weighted combination of the WRFM values and purchase preferences.

2.5. Hybrid works

The success of collaborative filtering relies on the availability of a sufficiently large set of quality preference ratings provided by users. In practice, users may be reluctant to provide preference ratings; at least it is hard when they first register onto the system. Accordingly, providing accurate recommendations under sparse data conditions (few preference ratings) is a primary challenge for collaborative filtering (Konstan et al., 1997). Finding users with similar preferences is difficult, if the user-rating matrix is very sparse, causing the sparsity problem for the CF method.

Hybrid approaches have been proposed to overcome drawbacks of the CF and CBF methods. They combine content-based filtering and collaborative filtering to improve recommendation accuracy. Two such approaches, the weighted model and the meta-level model, had been proposed that used different strategies to combine content-based and collaborative filtering (Burke, 2002; Li et al., 2003). The weighted model used linear combinations of prediction results of collaborative and content-based filtering. For example, the method was applied to recommend news in an online newspaper (Chen & Chen, 2001), using an adaptive weighted average to combine content-based and collaborative filtering predictions.

The meta-level model employed a sequential combination of collaborative and content-based filtering, where the output generated by content-based filtering is used as the input of collaborative filtering (Burke, 2002). The sequential combination is to combine two matrices concurrently via different weightings based on the spirit of hybrid methods to improve the recommendation accuracy, respectively. The user profile contains user preferences of each product features (describing the user's interests). Similarity measures between user profiles and product profiles (features of product items) are then derived to predict users' preference ratings on unrated product items.

This process aims to convert a sparse user-rating matrix into a dense user-rating matrix. Collaborative filtering then uses the dense matrix to provide recommendations. For instance, Melville, Moony, and Nagarajan (2002) presented a content-boosted collaborative filtering (CBCF) approach for movie recommendations, where pseudo user-ratings are derived by combining users' actual ratings and content-based predictions on unrated items. Then, the method performs collaborative filtering based on this dense matrix.

3. Proposed recommendation methods

Section 3.1 describes the derivation of customer demands in each industry. Furthermore, the derivation of extended preferences by combining purchase preferences and customer demands are introduced in Section 3.2. The proposed hybrid methods that combine CF, WRFM method and valuable customer demands/extended preferences to make recommendations are detailed in Sections 3.3 and 3.4, respectively.

3.1. Customers demands

CBF methods make recommendations by analyzing the descriptions (features) of the items that users rated to derive user profiles. Sometimes the descriptions of items are not easy to obtain. For example, it may be difficult to extract features of music. Customer demands for each industry are used as valuable information in this work. Most customers belonging to the same industry tend to have similar demands for some products. These demands are determined by simple statistics that calculate the frequency count of each product item purchased by customers in each industry. If the frequency count of a product item purchased in an industry is greater than a given threshold θ, then customers in such industry tend to have a demand for the item. Content-based recommendations generally deal with new items unseen by others (Balabanovic & Shoham, 1997). However, because this work was limited to extracting product features from a specific data set, it could not address new items. But the problem of new customers can still be solved. The system will know the industry of a new customer ahead of time, so when he/she wants to buy products, it can recommend products to him/her. The element r_{ij} of a customer-demand matrix $\boldsymbol{CD}$ represents whether the ith customer tends to have a demand for the jth product. If the ith customer tends to have a demand for the jth product, i.e., the frequency count of the jth product purchased in ith customer's industry is greater than θ, r_{ij} is 1; otherwise it is 0. The similarity of customer demands ($Corr_{cd}$) among customers can also be measured by computing the Pearson correlation coefficient.

3.2. Customers' extended preferences

In a real domain, customers may purchase very few product items and thus the customer-item matrix $\boldsymbol{R}$ is generally sparse. But the fact that a customer has not bought a product does not imply that the customer does not need or is not interested in that product. Therefore, this work proposed a denser matrix based on the spirit of hybrid methods that use sequential combination to improve the recommendations accuracy. Limited to available content information, in which product features are not provided in the data set, customer demands are integrated with purchase preferences to reduce the sparsity of customer-item matrix $\boldsymbol{R}$. Herein, the denser matrix is named the extended-preference matrix, $\boldsymbol{EP}$. The element r_{ij} of the extended-preferences matrix $\boldsymbol{EP}$ represents whether the ith customer had purchased or tend to have the demand for the jth product. If the ith customer tends to have a demand for or had purchased the jth product, r_{ij} is 1; otherwise it is 0. The similarity of extended preferences ($Corr_{ep}$) among customers can also be measured by computing the Pearson correlation coefficient.

3.3. Combining WRFM and customer demands

The WRFM-method (Liu & Shih, 2005b) effectively identifies recommendation rules for customers with high lifetime value or loyalty. However, as described previously, the method belongs to collaborative filtering and suffers the drawbacks of CF methods. Accordingly, this work proposes a hybrid method combining WRFM-based CF and

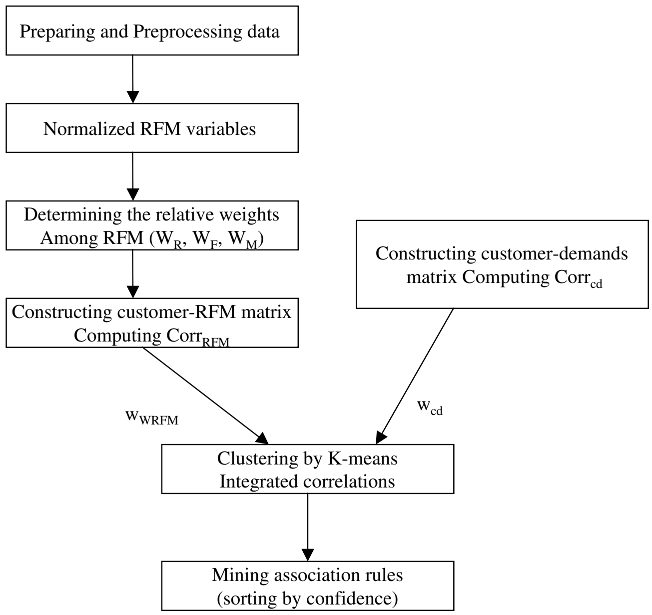

Fig. 1. WRFMCD method for product recommendation.

customer demands, termed WRFMCD. Fig. 1 illustrates the WRFMCD method.

The WRFMCD method first establishes a customer-WRFM and a customer-demand matrix. Then, the WRFM-based and the customer-demand correlation coefficients are computed using the Pearson correlation coefficient, respectively. Subsequently, *K*-means clustering is used to group customers with similar CLV and customer demands based on weighted correlation coefficients. Finally, the association rule mining approach is applied to extract recommendation rules from each group derived from *K*-means clustering. Candidate products are sorted by their associated confidence values, and the top-*N* highest ranked candidate products are recommended to users. The following subsections detail the WRFMCD method.

3.3.1. Data preparing

A data set is used to elucidate the proposed methodology. The case concerns a hardware retailing company that manufactures wheels, casters, platforms and hand trucks for industrial, medical, hospital and institutional use. Its decision-makers must target customer groups and develop market strategies to satisfy customer needs and thereby increase the market share of the company. Two years of data on purchase transactions and approximately 70,000 records have been collected. For each customer, a customer transaction is created to record all products previously purchased. Customers purchased an average of 34 product items. The data set is preprocessed to extract customer transactions. Unreasonable records such as those of customers who have a non-zero amount of purchases but have never made any transactions are also removed. In this study, 895 customer transactions are extracted from the database. RFM values of customer transactions are extracted to measure the customers' CLV. Table 1 shows some CLV expresses in terms of RFM.

Table 1
RFM values for each customer

Customer no.	Recency (days)	Frequency	Monetary (NT. Dollars)
1260003	159	87	313,763
1260006	135	44	146,444
1260009	111	379	1,426,665
...	...	...	...
1300050	256	1	7700

3.3.2. Grouping customers with similar CLV and customer's demands by K-means

This method clusters customers by integrating their CLV and customer demands (refer to Section 3.1). Customer demands are determined by simple statistics that calculate the frequency count of each product item purchased by customers in each industry. Customers who belong to the same industry tend to have similar demands for products. This work begins by establishing the customer-WRFM matrix and customer-demand matrix. The RFM value is normalized and then multiplied by the relative importance of the RFM variables. Eq. (2) is used to compute the WRFM-based Pearson correlation coefficient, $corr_{\mathrm{WRFM}}$, while Eq. (1) is used to compute the customer-demand based correlation coefficient, $corr_{cd}$. The integrated correlation coefficient is then derived according to Eq. (3):

$$Corr_{\mathrm{WRFMCD}}(c_i, c_j) = w_{\mathrm{WRFM}} \times corr_{\mathrm{WRFM}}(c_i, c_j) + w_{cd} \times corr_{cd}(c_i, c_j) \tag{3}$$

w_{WRFM} and w_{cd} represent the relative importance (weights) of the CLV elements and customer demands, respectively. If $w_{\mathrm{WRFM}} = 0$, customer demands are used for recommendations; if $w_{\mathrm{WRFM}} = 1$, the method becomes a WRFM-based CF method.

K-means technique is employed to cluster customers according to the integrated correlation coefficients. In general, a coefficient between the centroid c_j of a cluster and a customer c_i is measured using Eq. (3). The centroid is represented by both the average WRFM values and the average customer demands of the customers in the cluster. Customers are assigned to a cluster with a maximum integrated correlation coefficient. The weights of parameters w_{WRFM} and w_{cd} are used to yield an integrated correlation coefficient. The proper weighting values of w_{WRFM} and w_{cd} can be determined by performing analytical experiments to evaluate the quality of recommendations under different weight combination (for example, w_{WRFM} equals 0.8 and w_{cd} equals 0.2).

3.3.3. Recommendation phase

Association rule mining is used to extract a set of recommendation rules from the transactions associated with each cluster, which is generated by grouping customers according to the weighted correlation coefficients of CLV and customer demands. The set of recommendation rules extracted from cluster C_j is then used to select the top-N candidate products to be recommended to customer u. Notably, the candidate products are ranked according to the associated confidence values, as described in Section 2.2. Here, customer demands can also be used as a factor to design an adjusted recommendation method, termed the *adjusted* WRFMCD (A-WRFMCD) method. Candidate products are re-ranked according to their recommendation values derived by multiplying their associated confidence values with the degree of customer demands on products. The recommendation value of a candidate product b, $R(b)$, is equal to $P_r(b|a)$ multiplied by d_b, the degree of customer demand on product b. $P_r(b|a)$ is the confidence value, i.e., the probability of buying b given a. The resulting formula is

$$R(b) = P_r(b \mid a) * \ d_b$$

Preliminary experimental results show that the recommendation quality improves by setting $d_b = 1$, if the customer tends to have the demand for product b; $d_b = 0.5$, otherwise. The recommended product item matching customer demands thus gains high priority in ranking among candidate products.

3.4. Combining WRFM and extended preferences

The WRFMCD method combines customer lifetime value and customer demands to recommend products to customers, but does not consider customer purchase preferences. This work proposes another hybrid method, namely the WRFMEP method, by which customers are clustered by integrating the dimensions of customer lifetime value and extended preferences. This method adopts relative weighting to adjust the importance of customer lifetime value and extended preferences in clustering. The association rule mining approach is then applied to extract recommendation rules from each group derived by K-means clustering. Fig. 2 illustrates the WRFMEP method.

3.4.1. Grouping customers with similar CLV and extended preferences by K-means

A customer-WRFM matrix and the extended-preferences matrix are first established in this method. Customer RFM values are normalized and then multiplied by the relative importance of the RFM variables. The integrated correlation coefficient is then derived according to Eq. (4):

$$Corr_{\mathrm{WRFMEP}}(c_i, c_j) = w_{\mathrm{WRFM}} \times corr_{\mathrm{WRFM}}(c_i, c_j) + w_{ep} \times corr_{ep}(c_i, c_j) \quad (4)$$

The K-means technique uses integrated correlation coefficients to cluster customers. Generally, such a coefficient between the centroid c_j of a cluster and the customer c_i is measured by Eq. (4). The centroid here is represented by both the average WRFM values and the average extended preferences of customers in the cluster. Customers are assigned to a cluster with a maximum integrated correlation coefficient. The weights of parameters w_{WRFM} and w_{ep} are used to yield an integrated correlation coefficient. The proper weighting values of w_{WRFM} and w_{ep} can be determined by performing analytical experiments to evaluate the quality of recommendations under different weight combinations (for example, w_{WRFM} equals 0.8 and w_{ep} equals 0.2). If $w_{\mathrm{WRFM}} = 0$, the method is termed the CFEP method, which combines the preference-based CF method and extended preferences for making recommendations. If $w_{\mathrm{WRFM}} = 1$, the method becomes the WRFM-based CF method.

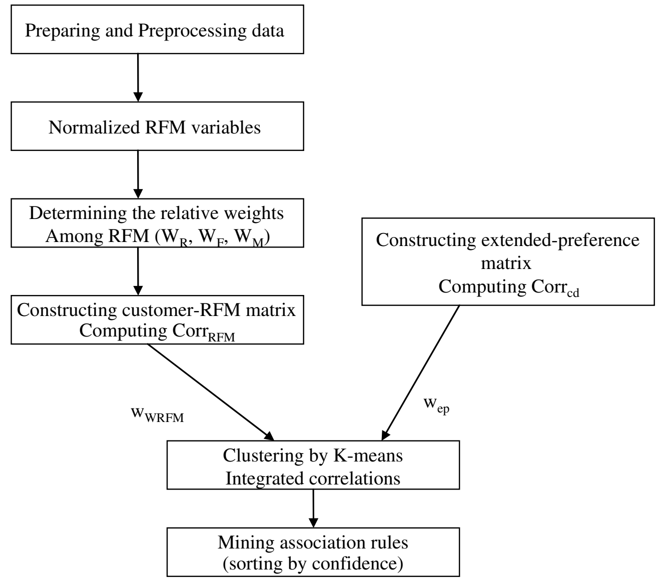

Fig. 2. WRFMEP method for product recommendation.

3.4.2. Recommendation phase

Association rule mining extracts a set of recommendation rules from transactions associated with each cluster. A cluster is generated by grouping customers according to the weighted correlation coefficient of CLV and extended preferences. We begin by identifying cluster C_j to which a customer, u belongs. Next, the set of recommendation rules extracted from C_j is used to select the top-N candidate products to be recommended to customer u. Similar to the A-WRFMCD method, the *adjusted* WRFMEP (A-WRFMEP) method uses customer demands to adjust and re-rank candidate recommended products (as described in Section 3.3.3).

3.5. Summary

A figure of proposed methods combining various sources is shown in Fig. 3. A WRFMCD method considered different weights of RFM and customer demands for product recommendation is introduced in Section 3.3. A CFEP method combines the customer preferences and customer demands for making recommendation is introduced in Section 4.1. The WRFMEP integrated the dimensions of WRFM and extended preferences is detailed in Section 3.4. Finally, customer demands is used as a factor to design adjusted recommendation methods, termed the *adjusted* WRFMCD (A-WRFMCD), the adjusted CFEP (A-CFEP) and adjusted WRFMEP (A-WRFMEP), respectively.

4. Experimental evaluation

In order to clarify all methods used in this paper, we summarized all acronyms in Table 2.

4.1. Experimental setup

The proposed methods, WRFMCD (see Section 3.3) and WRFMEP (see Section 3.4) are compared with several other methods, including WRFM-based, preference-based CF methods. Shih et al. (2005) implemented a CFEP (collaborative filtering based on extended preferences) method, which combines the preference-based CF method and customer demands for making recommendations. The CFEP method first establishes an extended-preference matrix (see Section 3.2). Second, customers are grouped according to similarity measures derived from extended preferences. The association rule mining technique is then employed to extract recommendation rules from each group. Additionally, this work implements an EP-based k-NN method to make recommendations. The EP-based k-NN method employs nearest neighbor to recommend products to a target customer based on those customers having similar extended preferences. Accordingly, the CFEP and EP-based k-NN are also compared with the proposed methods.

WRFMCD, WRFMEP, and CFEP are compared with A-WRFMCD (Section 3.3.3), A-WRFMEP (Section 3.4.2) and A-CFEP to confirm the usefulness of re-ranking candidate products based on customer demands. Similar to the A-WRFMCD and A-WRFMEP methods, the adjusted CFEP (A-CFEP) method also uses customer demands to adjust and re-rank candidate recommended products. Most hybrid works use CBF to reduce the sparsity problem so as to enhance the CF method. The proposed hybrid methods integrate customer demands and purchase preferences to derive extended preferences and further improve recommendation accuracy. To verify that methods based on extended preferences not only improve the overall quality of recommendation, but are also useful in alleviating sparsity problems for customers making few purchases, this work conducts experiments to evaluate whether methods that consider extended preferences are better than those that do not consider extended preferences.

The hardware retailing data set was divided into a 75% training set and a 25% testing set. The training set included product items purchased by customers in a specified period and was used to extract recommendation rules by association rule mining. Moreover, a preliminary analytical experiment was conducted to determine the proper weighting of w_{cd} and w_{ep} in WRFMCD and WRFMEP methods, respectively. The training set was used as the analytical data set in the preliminary analytical experiment, where 65% of the data set was used for deriving recommendation rules and 10% for analyzing recommendation quality. The minimum confidence level was set to 0.8 and the minimum support to 0.1. Identifying all frequent *itemsets* was difficult, since the average number of product items purchased by customers is 34. Hence, association rule mining explored only frequent *itemsets* with sizes less than or equal to three.

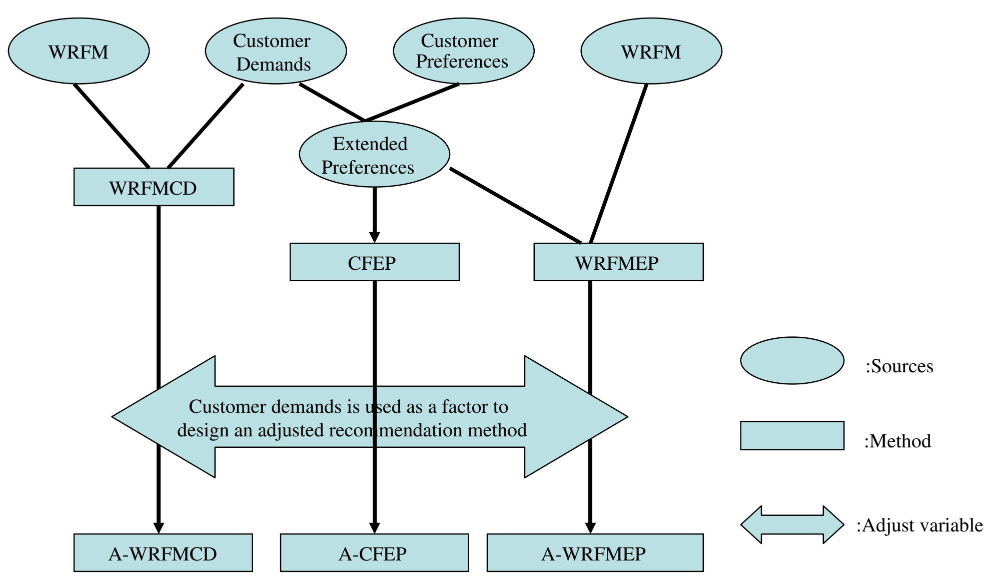

Fig. 3. Proposed methods combining various sources.

Table 2
Listing of terms and acronyms

Terms	Acronyms
Content-based filtering	CBF
Customer demands	CD
Collaborative filtering	CF
Extended preferences	EP
Customer lifetime value	CLV
Weighted RFM-based CF method	WRFM-based CF method
A hybrid method combining WRFM-based CF and customer demands	WRFMCD
Adjusted WRFMCD	A-WRFMCD
Collaborative filtering based on extended preferences	CFEP
Adjusted CFEP	A-CFEP
A hybrid method combining WRFM-based CF and extended preferences	WRFMEP
Adjusted WRFMEP	A-WRFMEP
A hybrid method combining WRFM-based CF and customer preferences	WRFMCP
EP-based by *K*-means method	EP-based KNN

Testing data were used to verify recommendation quality of the various methods.

4.2. Experimental metrics

Two metrics, precision and recall, are commonly used to measure the quality of a recommendation. These are also used measures in information retrieval (Salton & McGill, 1983; van Rijsbergen, 1979). Product items can be classified into products that customers are interested in purchasing, and those that they are not interested in purchasing. A recommendation method may recommend interesting or uninteresting products. The recall-metric indicated the effectiveness of a method for locating interesting products. The precision-metric represented the extent to which the product items recommended by a method really are interesting to customers.

Recall is the fraction of interesting product items that can be located:

$$\mathtt{Recall} = \frac{\mathtt{number\ of\ correctly\ recommended\ items}}{\mathtt{number\ of\ interesting\ items}}$$

Precision is the fraction of recommended products (predicted to be interesting) that are really found to be interesting:

$$\mathtt{Precision} = \frac{\mathtt{number\ of\ correctly\ recommended\ items}}{\mathtt{number\ of\ recommended\ items}}$$

Items interesting to customer u were those products purchased by u in the test set. Correctly recommended items were those that match interesting items. However, increasing the number of recommended items tended to reduce the precision and increase the recall. An F1-metric (Sarwar et al., 2000) could be used to balance the trade-off between precision and recall. F1 metric assigned equal weight to precision and recall and was given by

$$\mathtt{F1} = \frac{2 \times \mathtt{recall} \times \mathtt{precision}}{\mathtt{recall} + \mathtt{precision}}$$

Each metric was computed for each customer, and the average value computed for each cluster, as well as the overall average (over all customers) as measures of the quality of the recommendation.

4.3. Experimental results

4.3.1. Evaluation of WRFMCD method

The WRFMCD method considered different weightings on the dimensions of CLV and customer demands. The analytical experiment used the training set as the analytical data set (65%) to derive recommendation rules and 10% to determine the proper weightings, w_{WRFM} and w_{cd} ($w_{cd} = 1 - w_{\mathrm{WRFM}}$). Candidate products were sorted by associated confidence value, where the top-N highest ranked candidate products were selected as the recommendation set. Accordingly, the WRFMCD method achieved the best recommendation quality when $w_{\mathrm{WRFM}} = 0.3$ and $w_{cd} = 0.7$. Based on the analytical results, Table 3 summarizes the experimental results of the WRFMCD and A-WRFMCD methods on the testing set (25% data set) by setting $w_{\mathrm{WRFM}} = 0.3$ and $w_{cd} = 0.7$ to derive top-N recommendations. Overall, the analytical results suggest that the A-WRFMCD method is better than the WRFMCD method.

4.3.2. Evaluation of WRFMEP method

The WRFMEP method considered different weightings on the dimensions of CLV and extended preferences. The analytical experiment used the training set as the analytical data set (65%) to derive recommendation rules and 10% to determine the proper weightings, w_{WRFM} and w_{ep} ($w_{ep} = 1 - w_{\mathrm{WRFM}}$). If the $w_{\mathrm{WRFM}} = 0$, the method is CFEP; otherwise, the method is the WRFM-based CF method. Candidate products were sorted by associated confidence value, where the top-*20* highest ranked candidate products were selected as the recommendation set. The analytical result is shown in the left-hand side of Table 4. The WRFMEP method achieved the best recommendation quality when $w_{\mathrm{WRFM}} = 0.3$ and $w_{ep} = 0.7$. The right-hand side of Table 4 shows the analytical result of A-WRFMEP. Overall, when $w_{ep} > w_{\mathrm{WRFM}}$, the F1 metric of WRFMEP method exceeds that obtained using the WRFM-based CF ($w_{ep} = 0$; $w_{\mathrm{WRFM}} = 1$) and CFEP methods ($w_{ep} = 1$; $w_{\mathrm{WRFM}} = 0$). The A-WRFMEP method also outperformed the WRFMEP method. Based on the analytical results, further experiments (Sections 4.3.3–4.3.5) were conducted to evaluate the WRFMEP method by setting $w_{\mathrm{WRFM}} = 0.3$ and $w_{ep} = 0.7$.

4.3.3. Verifying the importance of extended preferences

Experiments were conducted to compare the EP-based k-NN, WRFMEP, and CFEP with the KNN-based, WRFM-based CF and preference-based CF methods to

Table 3
Analytical results of WRFMCD method under different N (top-N)

Top-N	WRFMCD method			A-WRFMCD method		
	Precision	Recall	F1-metric	Precision	Recall	F1-metric
Top-4	0.323	0.327	0.312	0.324	0.331	0.314
Top-10	0.447	0.639	0.516	0.453	0.658	0.528
Top-20	0.453	0.647	0.522	0.467	0.672	0.539
Top-30	0.451	0.648	0.521	0.459	0.664	0.528
Top-40	0.415	0.626	0.488	0.418	0.638	0.494
Top-50	0.412	0.624	0.483	0.417	0.628	0.489

verify the importance of extended preferences. Moreover, the WRFMEP method was also compared with the WRFM-based CF, WRFMCP, and WRFMCD methods. The training set (75%) included product items purchased by customers during a specified period and was used to extract recommendation rules by association rule mining. The analytical data set (25% data set) was used verify the quality of the recommendations. Methods were compared by varying N, the number of recommended items.

Table 5 summarizes the recommendation quality obtained using these various methods. From the experimental results, the F1-metrics of CFEP exceed those of the preference-based CF method. Moreover, the F1-metrics of WRFMEP exceed those of the WRFM-based CF method, as well as the WRFMCP and WRFMCD methods. EP-based k-NN also provides better recommendations than the KNN-based method. Generally, the performance ranking of these methods with extended preferences is WRFMEP $\succ$ CFEP $\succ$ EP-based k-NN method; while the ranking of these methods without extended preferences is WRFMCP method $\succ$ WRFM-based CF method $\succ$ preference-based CF method $\succ$ KNN-based method. This ranking implies that extended preferences, which are derived by combining customer demands and purchase preferences, are useful for improving the quality of recommendation.

4.3.4. Verifying the importance of re-ranking candidate products

Experiments were conducted to compare various methods using the 75% training set and the 25% testing set to verify the proposed adjusting approaches via varying the N, the number of recommendation items. Table 6 shows the F1-metrics of various methods under different top-N recommendations. Generally, the F1-metrics of both adjusting methods exceed those methods without re-ranking candidate products. Re-ranking candidate products according to customer demands offers a promising approach for improving recommendation accuracy.

4.3.5. Experiments on customers who purchase few product items

Previous experiments focused on confirming the overall recommendation accuracy, but did not consider sparse problems. Accordingly, experiments were conducted to compare various methods for those users who purchased product items not exceeding 5, 10 and 15 items from the 75% training set. The numbers of such customers were 73, 161 and 260, respectively.

Table 7 lists the experimental results displaying a trend similar to those of experiments involving all customers. The F1 metrics of methods with extended preferences exceed those of methods without considering extended preferences. CFEP outperformed the typical CF method. EP-based k-NN performed better than the KNN-based method. Furthermore, WRFMEP outperformed the WRFMCP, WRFMCD and WRFM-based methods. The result implies that the proposed hybrid method improves the overall quality of recommendation. Additionally, making recommendations for customers who purchased few product items based on extended preferences is better than

Table 4
Analytical results of WRFMEP method (top-20)

W_{WRFM}	WRFMEP method			A-WRFMEP method		
	Precision	Recall	F1-metric	Precision	Recall	F1-metric
0	0.413	0.621	0.507	0.432	0.652	0.514
0.1	0.448	0.662	0.528	0.474	0.675	0.542
0.2	0.451	0.665	0.532	0.473	0.675	0.542
0.3	0.457	0.669	0.533	0.474	0.677	0.543
0.4	0.457	0.667	0.533	0.474	0.677	0.543
0.5	0.457	0.664	0.531	0.474	0.675	0.541
0.6	0.453	0.664	0.528	0.469	0.674	0.540
0.7	0.452	0.664	0.527	0.465	0.675	0.539
0.8	0.448	0.658	0.522	0.460	0.673	0.535
0.9	0.442	0.654	0.518	0.459	0.669	0.533
1	0.436	0.647	0.511	0.452	0.663	0.531

Table 5
Analytical results of various methods to verify the importance of extended preferences

Top-N	Preference-based CF method	CFEP	WRFM-based CF method	WRFMCD	WRFMCP	WRFMEP	KNN-based ($k = 100$)	EP-based k-NN ($k = 100$)
Top-4	0.335	0.294	0.333	0.312	0.342	0.323	0.286	0.298
Top-10	0.476	0.497	0.499	0.516	0.486	0.497	0.487	0.490
Top-20	0.518	0.518	0.524	0.522	0.533	0.535	0.515	0.515
Top-30	0.502	0.525	0.504	0.521	0.525	0.533	0.498	0.518
Top-40	0.496	0.500	0.484	0.488	0.496	0.513	0.467	0.482
Top-50	0.473	0.495	0.477	0.483	0.489	0.505	0.422	0.467

Table 6
F1-metrics of various methods under different N (top-N)

Methods	WRFMCD	A-WRFMCD	WRFMEP	A-WRFMEP	CFEP	A-CFEP
Top-4	0.312	0.314	0.323	0.294	0.294	0.319
Top-10	0.516	0.528	0.497	0.508	0.497	0.506
Top-20	0.522	0.539	0.535	0.543	0.518	0.524
Top-30	0.521	0.528	0.533	0.542	0.525	0.527
Top-40	0.488	0.494	0.513	0.522	0.500	0.507
Top-50	0.483	0.489	0.505	0.512	0.495	0.496

Table 7
Experimental results of various methods for customers with few purchases

	Purchased items $\leqslant$ 5 (73)			Purchased items $\leqslant$ 10 (161)			Purchased items $\leqslant$ 15 (260)		
	Top-10	Top-20	Top-30	Top-10	Top-20	Top-30	Top-10	Top-20	Top-30
Preference-based CF	0.3524	0.3858	0.3852	0.3655	0.3632	0.3667	0.3643	0.3637	0.3612
CFEP	0.3718	0.3861	0.3714	0.3824	0.3856	0.3807	0.4151	0.3935	0.3912
KNN-based	0.3438	0.3506	0.3201	0.3502	0.3721	0.3286	0.3523	0.3608	0.3200
EP-based KNN	0.3615	0.3688	0.3239	0.3859	0.3817	0.3306	0.4293	0.3688	0.3239
WRFM-based	0.3712	0.3618	0.3540	0.3856	0.3746	0.3729	0.4214	0.3884	0.3805
WRFMCD	0.3945	0.3914	0.3857	0.4112	0.3969	0.3835	0.4243	0.4064	0.3893
WRFMCP	0.3715	0.3662	0.3547	0.3913	0.3841	0.3803	0.4221	0.4022	0.3814
WRFMEP	0.4042	0.3935	0.3882	0.4197	0.4025	0.3881	0.4273	0.4086	0.3930

those methods without considering extended preferences. Generally, the quality of recommendation improves as the number of purchased items increases.

5. Conclusions

The collaborative filtering method has been successfully used in a number of applications, but suffers several limitations. This work uses customer demands derived from frequently purchased products in each industry to integrate with the CF method to make recommendations. This work also combines customer demands and past purchase preferences to reduce the sparsity of customer-item matrix and further improves recommendation accuracy. Customer demands are included as a factor in re-ranking candidate products to provide recommendations. Several experiments were conducted to compare the effectiveness between various methods.

According to the experimental results, generally, the performance ranking of those methods with extended preferences is WRFMEP $\succ$ CFEP $\succ$ EP-based k-NN method; while the ranking of those methods without considering extended preferences is WRFMCP method $\succ$ WRFM-based CF method $\succ$ preference-based CF method $\succ$ KNN-based method. This ranking implies that extended preferences, derived by combining customer demands and purchase preferences, are useful for improved recommendation quality. Furthermore, re-ranking candidate products according to customer demands offers a promising approach to improve recommendation accuracy. Finally, the experimental results show that the proposed hybrid methods not only improve the overall quality of recommendation, but also can be extended to recommend product items to customers who purchased few product items based on extended preferences. In general, the quality of recommendation improves as the number of purchased items increases.

Future works will address three themes. First, the proposed approach was evaluated experimentally using a data set obtained from a hardware retailer. Further studies are needed to evaluate the application of the proposed approach to other application domains. Second, the present work focused on product recommendation of retail transaction data which contains binary choice of shopping basket data; the customer preference is represented as one, if the customer purchased the product; and zero, otherwise.

Further investigation is needed to evaluate the effectiveness of the proposed methods for data sets with non-binary preference ratings. Finally, owing to the limitations of available content information of the data set concerned, this work could not address new and unseen items. Further studies are required to verify the proposed methods on other real cases that can support more content information.

Acknowledgements

The authors gratefully acknowledge the Editor and anonymous reviewers for their valuable comments and constructive suggestions. This research was supported in part by the National Science Council of the Republic of China under the Grant NSC 94-2416-H-159-004 and NSC 95-2416-H-216-009.

References

Balabanovic, M., & Shoham, Y. (1997). Fab: Content-based, collaborative recommendation. *Communications of the ACM, 40*(3), 66–72.

Basu, C., Hirsh, H., & Cohen, W. (1998). Recommendation as classification: Using social and content-based information in recommendation. In *Proceedings of the 15th national conference on artificial intelligence* (pp. 714–720).

Berger, P., & Nasr, N. (1998). Customer lifetime value: Marketing models and applications. *Journal of Interactive Marketing, 12*(1), 17–30.

Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithm for collaborative filtering. In *Proceedings of the 14th annual conference on uncertainty in artificial intelligence* (pp. 43–52).

Burke, R. D. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction, 12*(4), 331–370.

Chen, H. C., & Chen, A. L. P. (2001). A music recommendation system based on music data grouping and user interests. In *Proceedings of the ACM conference on information and knowledge management* (pp. 231–238).

Claypool, M., Gokhale, A., & Miranda, T. (1999). Combining content-based and collaborative filters in an online newspaper. In *Proceedings of the ACM SIGIR workshop on recommender system*.

Good, N., Shafer, J. B., Konstan, J. A., Borchers, A., Sarwar, B. M., Herlocker, J. L., et al. (1999). Combining collaborative filtering with personal agents for better recommendations. In *Proceedings of the 6th national conference on artificial intelligence* (pp. 439–446).

Hill, W., Stead, L., Rosenstein, M., & Furnas, G. (1995). Recommending and evaluating choices in a virtual community of use. In *Proceedings of the ACM (CHI'95)*.

Kahan, R. (1998). Using database marketing techniques to enhance your one-to-one marketing initiatives. *Journal of Consumer Marketing, 15*(5), 491–493.

Kamba, T., Bharat, K., & Albers, M. C. (1995). The Krakatoa Chronicle—an interactive personalized newspaper on the Web. In *Proceedings of the 4th international World Wide Web conference, December* (pp. 11–14).

Kim, J. W., Lee, B. H., Shaw, M. J., Chang, H. L., & Nelson, M. (2001). Application of decision-tree induction techniques to personalized advertisements on Internet storefronts. *International Journal of Electronic Commerce, 5*(3), 45–62.

Konstan, J. A., Miller, B. N., Maltz, D. M., Herlocker, J. L., Gordon, L. R., & Rieddl, J. (1997). GroupLens: Applying collaborative filtering to UseNet News. *Communications of the ACM, 40*(3), 77–87.

Li, Q., & Kim, B. M. (2003). An approach for combining content-based and collaborative filters. In *Proceedings of the 6th international workshop on information retrieval with asian languages* (pp. 17–24).

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing, 7*(1), 76–80.

Liu, D. R., & Shih, Y. Y. (2005a). Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences. *Journal of Systems and Software, 77*(2), 181–191.

Liu, D. R., & Shih, Y. Y. (2005b). Integrating AHP and data mining for product recommendation based on customer lifetime value. *Information & Management, 42*(3), 387–400.

Melville, P., Moony, R. J., & Nagarajan, R. (2002). Content-based collaborative filtering for improved recommendations. In *Proceedings of the eighteenth national conference on artificial intelligence* (pp. 187–192).

Miglautsch, J. (2000). Thoughts on RFM scoring. *Journal of Database Marketing, 8*(1), 67–72.

Peppers, D., & Rogers, M. (1993). *The one-to-one future: Building relationships one customer at a time*. New York: Currency and Doubleday.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of Netnews. In *Proceedings of the ACM conference on computer supported cooperative work* (pp. 175–186).

Rucker, J., & Polanco, M. J. (1997). Personalized navigation for the Web. *Communications of the ACM, 40*(3), 73–75.

Saaty, T. L. (1994). *Fundamentals of decision making and priority theory with the analytic hierarchy process*. Pittsburgh, PA: RWS Publications.

Salton, G., & McGill, M. J. (1983). *Introduction to modern information retrieval*. New York: McGraw-Hill.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Analysis of recommendation algorithms for e-commerce. In *Proceedings of the 2nd ACM conference on electronic commerce, Minneapolis* (pp. 158–167).

Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. *Journal of Data Mining and Knowledge Discovery, 5*(1/2), 115–152.

Shardanand, U., & Maes, P. (1995). Social information filtering: Algorithms for automating 'world of mouth'. In *Proceedings of the ACM (CHI'95)*.

Shih, Y. Y., & Liu, D. R. (2005). Hybrid recommendation approaches: Collaborative filtering via valuable content information. In *Proceedings of the thirty-eight Hawaii international conference on system sciences (HICSS-38), Hawaii*.

van Rijsbergen, C. J. (1979). *Information retrieval* (2nd ed.). London: Butterworth.

Zeng, C., Xing, C. X., Zhou, L. Z., & Zheng, X. H. (2004). Similarity measure and instance selection for collaborative filtering. *International Journal of Electronic Commerce, 8*(4), 115–129.