Mobile commerce product recommendations based on hybrid multiple channels

Duen-Ren Liu, Chuen-He Liou *

Institute of Information Management, National Chiao Tung University, 1001 University Road, Hsinchu 300, Taiwan

ARTICLE INFO

ABSTRACT

The number of third generation (3G) subscribers conducting mobile commerce has increased as mobile data communications have evolved. Multi-channel companies that wish to develop mobile commerce face difficulties due to the lack of knowledge about users' consumption behavior on new mobile channels. Typical collaborative filtering (CF) recommendations may be affected by the so-called *sparsity problem* because relatively few products are browsed or purchased on the mobile Web. In this study, we propose a hybrid multiple channel method to address the lack of knowledge about users' consumption behavior on a new channel and the difficulty of finding similar users due to the sparsity problem of typical CF recommender systems. Products are recommended to users based on their browsing behavior on the new mobile channel as well as the consumption behavior of heavy users of existing channels, such as television, catalogs, and the Web. Our experiment results show that the proposed method performs well compared to the other recommendation methods.

1. Introduction

In the last decade, mobile communications have evolved from 2G/2.5G to 3G/3.5G. As a result, the data transfer rate has been progressively upgraded from 64 Kbps (2.5G/GPRS) to 384 Kbps (3G/WCDMA) and 3.5 Mbps (3.5G/HSDPA), which is comparable to that of the wired Internet. The evolution has triggered an increase in the use of mobile devices, such as mobile phones, to conduct mobile commerce (m-commerce) on the mobile Web (Venkatesh et al. 2003, Ngai and Gunasekaran 2007). M-commerce covers a large number of services, one of which is mobile shopping (m-shopping) (Wu and Wang 2006). Retailers have increased their investment in mobile shopping channels to deliver content, products, and promotions to customers. However, it is hard to determine consumption patterns since there are very few purchase orders in the developmental stage of an m-shopping channel. The number of product recommendations is also low because of the small number of consumption patterns that have been identified.

In the mobile commerce environment, the screens of mobile devices are small and have limited resolution, and the input mechanisms are poor (Ho and Kwok 2003, Venkatesh et al. 2003). Moreover, few products are browsed on the mobile Web because Internet fees for mobile communications are still high; hence, one-to-one product recommendations are important (Brunato and Battiti 2003). Recommender systems are widely used to recommend various items, such as movies and music, to customers

according to their interests (Hill et al. 1995, Shardanand and Maes 1995). Generally, recommender systems are based on either collaborative or content-based filtering techniques. Collaborative filtering (CF), which has been used successfully in various applications, utilizes preference ratings given by customers with similar interests to make recommendations to a target customer (Resnick et al. 1994, Sarwar et al. 2000). In contrast, content-based filtering (CBF) derives recommendations by matching customer profiles with content features (Lang 1995, Pazzani and Billsus 1997). Some studies have combined collaborative filtering and content-based filtering techniques as a hybrid recommendation method (Balabanović and Shoham 1997, Claypool et al. 1999).

The typical CF method relies on finding users with similar interests to make recommendations. However, it suffers from the *sparsity problem*, which arises because users rate very few items and the user–item rating matrix is very sparse; thus, the recommendation quality is poor due to the difficulty of finding users with similar interests (Sarwar et al. 2000). In mobile shopping environments, active users may only browse and purchase a few items on the mobile Web. As a result, it is difficult to find users with similar interests on the mobile Web based on product preferences, which are derived from users' browsing and purchasing histories.

In addition, multi-channel companies may experience difficulties when they develop a new channel due to the lack of knowledge about users' consumption behavior on the new channel. Most companies use feedback from advertisements and marketing campaigns to gain an understanding of users' consumption behavior on a new channel. Information about customers' consumption behavior can also be derived by analyzing customer transaction data. Customers may exhibit migratory behavior between channels

* Corresponding author.
   E-mail addresses: dliu@iim.nctu.edu.tw (D.-R. Liu), chlioua@ms22.hinet.net (C.-H. Liou).

(Thomas and Sullivan 2005, Ansari et al. 2008). Hence, some new channel users may be old customers who have migrated from existing channels. In addition, customers may purchase products across channels. Thus, the consumption behavior of new mobile channel users may partially correlate with the behavior of the users of other channels (e.g., television, catalogs, and the Web) with different degrees of overlap.

In this paper, we propose a hybrid multiple channel method to address the lack of knowledge about the consumption behavior of new channel users and the difficulty of finding similar users due to the *sparsity problem* inherent in typical CF systems. The method recommends products to new channel users. The recommendations are based on the browsing behavior of mobile users, as well as the consumption behavior of heavy users of existing channels, such as television, catalogs, and the Web. We define heavy users as customers who have purchased products in the recent past, purchased products frequently, and spent large amounts of money. Other studies offer similar definitions of heavy users. For example, Lim et al. (2005) and Gensler et al. (2007) define heavy users as customers whose average monthly purchase amount is above the median monthly purchase amount. Heavy Internet users spend more time surfing the WWW than other users (Van den Poel and Leunis 1999). Moreover, Miglautsch (2000) showed that a relatively small percentage of customers frequently contribute most of the revenue based on the Pareto principle (80/20 rule). In this work, we use the purchase behavior of heavy users of existing channels (e.g., television, catalogs, and the Web) to represent users' consumption behavior of each channel based on the 80/20 rule. Analyzing the purchase behavior of heavy users of existing channels could provide sufficient transaction instances that could be used to find more similar users for a new channel. This could solve the sparsity problem of the new channel and improve the recommendation quality.

The remainder of this paper is organized as follows. In Section 2, we discuss the background of our research. In Section 3, we explain how we select the heavy users of a channel. We also describe the proposed recommendation scheme and the recommendation engine. In Section 4, we present the experiment results. In Section 5, we discuss the results and their implications. In Section 6, we summarize our findings, state the limitations of the present study, and consider future research avenues.

2. Background and related work

2.1. Multiple sales channels

Multiple sales channels can be divided into physical channels (e.g., department stores) and virtual channels (e.g., the Web, catalogs, and television) (Tiernan 2001). In the past, most companies only provided single sales channels for customers to purchase products. However, because of advances in information technology and increased demand, companies now use the physical and virtual channels to provide customers with seamless services. In this way, companies create more value for their customers, including greater choice and convenience. The channels can also be designed to allow customers to move from one channel to another seamlessly by reducing transaction costs during the purchase process (Tiernan 2001, Chircu and Mahajan 2006, Schröder and Zaharia 2008). Existing studies do not consider how to determine the product preferences of new channel users based on the consumption behavior of users in other sales channels.

2.2. Market clustering

Clustering techniques, which are usually used to segment markets (Punj and Stewart 1983, Chen et al. 1996), seek to maximize the variance among groups while minimizing the variance within groups. A number of clustering algorithms have been developed, such as K-means, hierarchical, and fuzzy c-means algorithms (Omran et al. 2007). K-means clustering (MacQueen 1967) is a widely used similarity grouping method that partitions a dataset into k groups. The K-means algorithm assigns instances to clusters based on the minimum distance principle. An instance is assigned to a cluster based on the minimum distance to the center of the cluster over all k clusters.

The recency, frequency and monetary (RFM) framework is used to analyze customer behavior and define market segments. It is widely used in direct marketing and database marketing (Kahan 1998, Miglautsch 2000). Bult and Wansbeek (1995) defined the framework's terms as follows: (1) Recency is the period since the last purchase. A lower value corresponds to a higher probability that the customer will make a purchase in the near future. (2) Frequency is the number of purchases made within a certain period; higher frequency indicates greater loyalty. (3) Monetary represents the amount of money spent during a certain period; the higher the amount, the more the company should focus on that customer. Most direct marketing firms target market segments that have lower recency, higher frequency, and higher monetary values (Kahan 1998). Miglautsch (2000) suggested using the RFM model as a market segmentation tool to quantify customer behavior. His findings showed that a relatively small percentage of customers frequently contribute most of the revenue based on the Pareto principle (80/20 rule).

2.3. Association rule-based recommendation method

Association rule mining tries to find the associations between two sets of products in a transaction database. Agrawal et al. (1993) formalized the problem of finding association rules that satisfy the minimum support and the minimum confidence requirements. For example, assume that a set of purchase transactions includes a set of product items I. An association rule is an implication of the form $X \Rightarrow Y$, where $X \subset I, Y \subset I$, and $X \cap Y = \Phi$. X is the antecedent (body) and Y is the consequent (head) of the rule. Two measures, support and confidence, are used to indicate the quality of an association rule. The support of a rule is the percentage of transactions that contain both X and Y, whereas the confidence of a rule is, among all transactions that contain X, the fraction that also contains Y.

The support of an association rule indicates how frequently the rule applies to the target data. A high level of support corresponds to a strong correlation between the product items. The confidence score is a measure of the reliability of an association rule. The higher the level of confidence, the more significant will be the correlation between the product items. The Apriori algorithm (Agrawal and Srikant 1994) is normally used to find association rules by discovering frequent item sets of product items. An item set is considered to be frequent if its support exceeds a user-specified minimum support threshold. Association rules that meet a user-specified minimum confidence threshold can be generated from the frequent item sets.

Sarwar et al. (2000) described the association rule-based recommendation method as follows. For each customer, a customer transaction is created to record all the products that he or she purchased previously. An association rule mining algorithm is then applied to find all the recommendation rules that satisfy the given minimum support and minimum confidence. The top N products to be recommended to a customer, u, are then determined as follows. Let X_u be the set of products purchased by u previously. The method first finds all the recommendation rules $X \Rightarrow Y$ in the rule set. If $X \subseteq X_u$ then all products in $Y–X_u$ are deemed to be candidate products for recommendation to the customer u. The

candidate products are then sorted and ranked according to the associated confidence of the recommendation rules, and the top N candidate products are selected as the top N recommended products.

2.4. Most frequent item-based recommendation method

The most frequent item-based recommendation method (Sarwar et al. 2000) counts the purchase frequency of each product by scanning the products purchased by the users in a cluster. Next, all the products are sorted by the purchase frequency in descending order. Finally, the method recommends the top N products that have not been purchased by the target customer.

2.5. Collaborative filtering

Collaborative filtering (CF) (Resnick et al. 1994, Shardanand and Maes 1995) utilizes the nearest-neighbor principle to recommend products to a target audience. The neighbors are identified by computing the similarity between customers' purchase behavior patterns or tastes. The similarity is measured by Pearson's correlation coefficient, which is defined as follows:

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

where $\bar{r}_{c_i}$ and $\bar{r}_{c_j}$ denote the average number of products purchased by customers C_i and C_j, respectively; the variable I denotes the mix of the set of products; and $r_{ci,s}$ and $r_{cj,s}$ indicate, respectively, that customers C_i and C_j purchased product item S.

The KNN-based CF method utilizes k-nearest neighbors to recommend N products to a target user (Sarwar et al. 2000). The k-nearest neighbors are identified by computing the similarity between customers' purchase behavior or tastes. The similarity is measured by Pearson's coefficient, as shown in Eq. (1). After the neighborhood has been formed, the N recommended products are determined by the k-nearest neighbors as follows. The frequency count of products is calculated by scanning the data about the products purchased or browsed by the k-nearest neighbors. The products are then sorted based on the frequency count, and the N most frequently occurring products that have not been purchased by the target customers are selected as the top-N recommendations.

2.6. Collaborative filtering for e-commerce and m-commerce

Several electronic commerce applications use the collaborative filtering technique to make recommendations. GroupLens is a netnews recommendation system based on the collaborative filtering technique that helps users find news articles of interest. The system predicts scores based on the opinions of other readers who have already rated articles (Resnick et al. 1994). Sarwar et al. (2000) proposed a recommendation method that combines collaborative filtering and association rule mining techniques to recommend movies in the MovieLens database. Amazon.com uses an item-based collaborative filtering technique to recommend products that are similar to products purchased by the customer previously (Linden et al. 2003). Cho et al. (2005) designed a methodology that combines collaborative filtering and clustering techniques to recommend products to a department store's customers based on their sequential purchase patterns.

The drawback of many CF methods is that they may suffer from the data sparsity problem because customers only purchase a few products. Several studies have attempted to solve the problem in various applications. For example, Zeng et al. (2004) suggested using a class-based collaborative filtering technique to recommend

movies in the EachMovie database. Huang and Huang (2009) developed a food recommendation system based on a two-stage technique that discovers sequential patterns in product categories to mitigate the problem. And Kim et al. (2010) proposed a collaborative filtering method based on collaborative tagging to recommend webpages on a social bookmark website.

In mobile commerce, several applications use the collaborative filtering technique to provide recommendations. For example, Li et al. (2009) proposed a two-stage collaborative filtering method that incorporates clustering and sequential pattern techniques for mobile services (product advertising and ring tone downloads) based on users' profiles, preferences and locations. Liou and Liu (2009) combined mobile phone features (MPF) and product preference methods based on collaborative filtering techniques to provide product recommendations on the mobile Web. VISCOR is a mobile wallpaper recommender system that combines collaborative and content-based filtering to reduce users' search costs and provide better wallpaper recommendations (Kim et al. 2004). MONERS, a hybrid news recommender system, determines users' preferences for news articles based on the importance of the news event, the recency of the news, changes in users' preferences, user segments, and article preferences (Lee and Park 2007). In addition, mobility information about users' locations obtained from global positioning systems (GPS) is normally combined with recommendation methods in m-commerce. For example, PILGRIM is a location-based collaborative filtering system that recommends webpages to users who are in an ellipsoid area (Brunato and Battiti 2003). These methods use a single channel (mobile phone) to collect users' preferences and make recommendations. Relatively few studies have investigated the use of multiple channels (e.g., television, catalogs, and the Web) to recommend products. Table 1 summarizes the applications based on collaborative filtering (CF) recommendation techniques.

In this paper, we propose a multiple channel-based collaborative filtering technique to recommend products on a mobile shopping website. The proposed method successfully integrates the heterogeneous databases of a CRM system and the mobile website.

Table 1
Applications based on collaborative filtering recommendation techniques.

CF	Application	
Channel	e-commerce	m-commerce
Single	Techniques	Techniques
	• Item-based (Linden et al. 2003)	• Location-based (Brunato and Battiti 2003)
	• User-based (Resnick et al. 1994)	• User-based (Kim et al. 2004)
	• Clustering (Cho et al. 2005)	• Clustering (Lee and Park 2007)
	• Association rules (Sarwar et al. 2000)	• Association rules (Liou and Liu 2009)
	• Sequential pattern (Huang and Huang 2009)	• Sequential pattern (Li et al. 2009)
	Applications	Applications
	• News (Resnick et al. 1994)	• Webpages (Brunato and Battiti 2003)
	• Movies (Sarwar et al. 2000, Zeng et al. 2004)	• Wallpaper images (Kim et al. 2004)
	• Products (Linden et al. 2003, Cho et al. 2005, Huang and Huang 2009)	• Products (Liou and Liu 2009)
	• Bookmarks (Kim et al. 2010)	• News (Lee and Park 2007)
		• Ringtone (Li et al. 2009)
Multiple		Products: The approach proposed in this paper

3. Methodology

In this section, we describe the proposed hybrid recommendation scheme based on multiple channels: mobile, television, catalog, and Web channels. Fig. 1 provides an overview of the scheme. A channel's current users are divided into RFM groups to identify heavy users of the channel based on their recency, frequency and monetary values. The heavy users are then further divided into preference groups based on their product category preferences to provide recommendations for users of a new mobile channel.

First, we use the K-means clustering algorithm to cluster a channel's current users into RFM groups based on the Euclidean distance of the R, F, and M values. Then, we compare the average R, F, and M values of the groups to the average R, F, and M values of all users of the channel. Finally, the groups of heavy users are selected based on lower recency values, but higher frequency and monetary values. The heavy users of other existing channels could be used to provide more transaction instances for finding more similar users for the new channel.

Second, we use the K-means clustering method to cluster the heavy users of each channel into preference groups based on the users' similarity, which is measured by Pearson's correlation coefficient of users' product category preferences. Heavy users in the preference group could be used to find more similar users for the new channel. This would solve the sparsity problem of the new

channel and derive more association rules to improve the quality of recommendations. For each target mobile channel user, similar users are selected from the clusters of mobile, television, catalog, and Web channel users based on their product category preferences. The system then finds the association rules of products and product categories as well as the most frequent items of similar users of each channel. The association rules and most frequent items of the hybrid multiple channels are determined, respectively, from the rules and items of multiple channels using the weighted sum of the associated confidence scores and frequent counts with different hybrid weights of the mobile (w_M), television (w_T), catalog (w_C), and Web (w_W) channels. The hybrid weights indicate the relative importance of the multiple channels to the mobile channel, and are determined according to the best recommendation quality derived from the preliminary analytical data. The method uses the hybrid weights to recommend products based on the association-rules and most frequent-items approaches.

3.1. User selection and clustering of the existing channels

Recall that heavy users are valuable customers who have purchased products in the recent past, purchased products frequently, and spent large amounts of money. First, we calculate the R, F, and M values of each user in a channel. Second, we cluster users into groups by the K-means clustering method based on the Euclidean distance of the R, F, and M values. Then, we compare the average R,
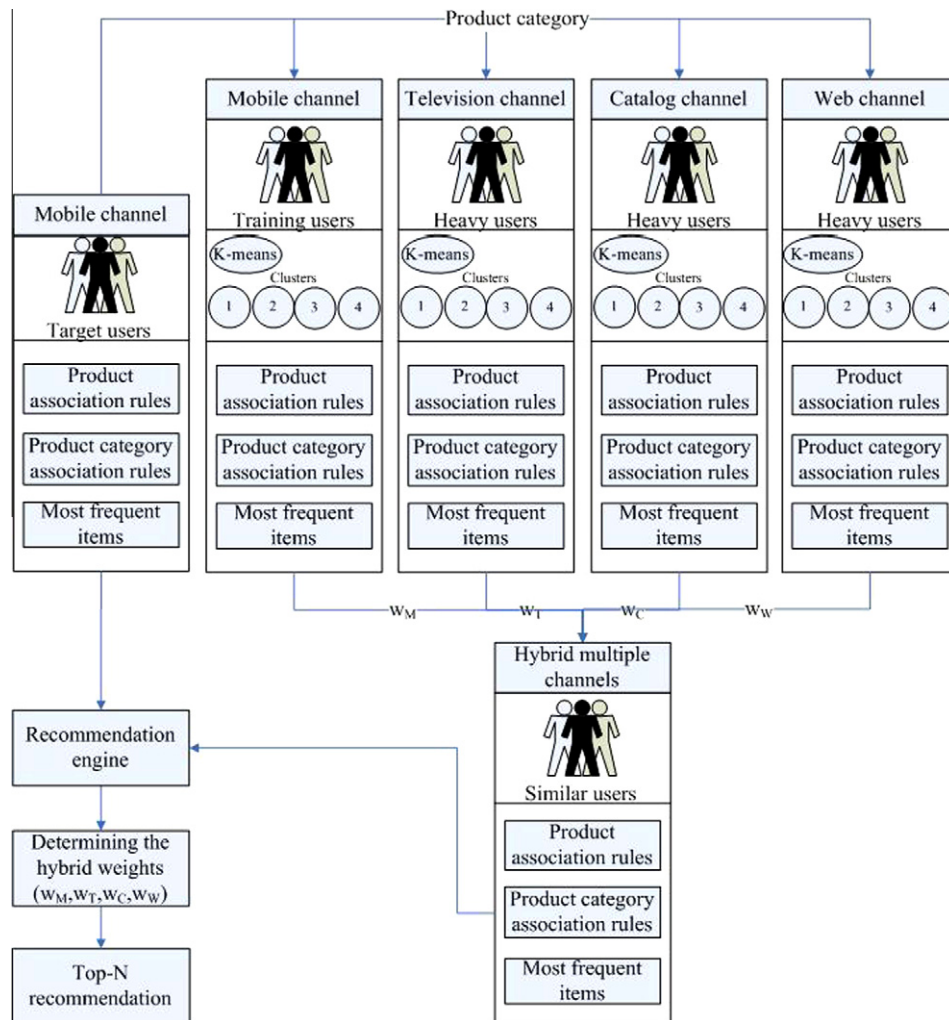

Fig. 1. An overview of the proposed recommendation scheme.

Table 2
User–product category rating matrix.

User ID	Cosmetics	Perfumes	Skincare	Pants	Shoes	Toys	Shirts	Notebooks	...
1	1	0	1	1	1	0	0	1	...
2	0	1	1	0	1	1	0	0	...
3	1	0	0	1	0	0	1	1	...
4	0	1	1	1	0	1	1	0	...
...	...	...	...	...	...	...	...	...	...

F, and M values of each group (the group average) to the average R, F, and M values of all users in a channel (the channel average). Heavy users are selected from groups with lower R values, but higher F and M values of group average than RFM values of channel average.

Next, the K-means clustering method is used to further cluster heavy users into groups based on the similarity of their consumption behavior patterns. The similarity is measured by Pearson's correlation coefficient of the user–product category preference matrix, as shown in Table 2.

3.2. The recommendation engine

The proposed hybrid multiple channel method derives recommendations based on the association-rule and most-frequent-items approaches. For each group of users, two kinds of association rules are extracted, namely, product-level association rules and category-level association rules. The former are extracted from the product transactions; and the latter are extracted from category-level transactions, which are derived by replacing the products in product transactions with their respective categories. The recommendation engine is comprised of three components: the product association rules ($X_H^{PR_i} \rightarrow Y_H^{PR_i}$) component, the product category association rules ($X_H^{CR_j} \rightarrow Y_H^{CR_j}$) component, and the most frequent items (Y_H^{Mf}) component, as shown in Fig. 2. In the figure, H represents either M, T, C, or W, which denote the mobile, television, catalog and Web channels, respectively.

In the multiple channel approach, let $X_H^{PR_i} \rightarrow Y_H^{PR_i}, H \in \{M, T, C, W\}$ be the product-level association rules extracted from the product transactions of a group of channel users, comprised of mobile, television, catalog, and Web channel users; and let their associated confidence scores be $cf_M^{PR_i}, cf_T^{PR_i}, cf_C^{PR_i}$, and $cf_W^{PR_i}$, respectively. In addition, let X_u represent the previous set of products that the target user u browsed in the mobile channel; and let Y_u^{AR} be the set of candidate products generated from the union of $Y_H^{PR_i} - X_u$ according to all the association rules $X_H^{PR_i} \rightarrow Y_H^{PR_i}$ that satisfy $X_H^{PR_i} \subseteq X_u$. The products in Y_u^{AR} are ranked according to the weighted sum of their confidence scores.

$$cf^{PR_i} = w_M \times cf_M^{PR_i} + w_T \times cf_T^{PR_i} + w_C \times cf_C^{PR_i} + w_W \times cf_W^{PR_i}, \tag{2}$$

where w_M, w_T, w_C, and w_W are the weights assigned to the mobile, television, catalog, and Web channels, respectively.

Let $Y_H^{Mf}, H \in \{M, T, C, W\}$ denote the set of most frequent items derived from the user groups of target user u in multiple channels. The frequency count of an item v for a user group U_g is equal to the number of users in U_g that had browsed item v. Let $f_{v,M}^{Mf}, f_{v,T}^{Mf}, f_{v,C}^{Mf}$, and $f_{v,W}^{Mf}$ represent the frequency counts of an item v in Y_H^{Mf}, respectively. Let Y_u^{Mf} be the set of candidate products generated from the union of $Y_H^{Mf} - X_u$. The products in Y_u^{Mf} are ranked according to the weighted sum of their frequency counts calculated as Eq. (3).

$$f_v^{Mf} = w_M \times f_{v,M}^{Mf} + w_T \times f_{v,T}^{Mf} + w_C \times f_{v,C}^{Mf} + w_W \times f_{v,W}^{Mf} \tag{3}$$

Let $X_H^{CR_j} \rightarrow Y_H^{CR_j}, H \in \{M, T, C, W\}$ be the category-level association rules extracted from the category-level transactions of a group of channel users, comprised of mobile, television, catalog, and Web

channels; and let their associated confidence scores be $cf_M^{CR_j}, cf_T^{CR_j}, cf_C^{CR_j}$, and $cf_W^{CR_j}$, respectively. In addition, let X_u^C represent the set of product categories that the target user u browsed previously from the mobile channel; and let Y_u^C be the set of candidate product categories generated from the union of $Y_H^{CR_j}$ according to all the category-level association rules $X_H^{CR_j} \rightarrow Y_H^{CR_j}$ that satisfy $X_H^{CR_j} \subseteq X_u^C$. The categories in Y_u^C are ranked according to the weighted sum of their confidence scores (Eq. (4)).

$$cf^{CR_j} = w_M \times cf_M^{CR_j} + w_T \times cf_T^{CR_j} + w_C \times cf_C^{CR_j} + w_W \times cf_W^{CR_j} \tag{4}$$

Let Y_u^{CMf} denote the set of most frequent candidate items derived from the candidate product categories Y_u^C and most frequent candidate items Y_u^{Mf}. We note that Y_u^{Mf} is derived from the user groups of target user u in multiple channels. Y_u^{CMf} is the set of items in Y_u^{Mf} that also belong to the candidate categories in Y_u^C. Each item v in Y_u^{CMf} is associated with a pair of (cf^{C_k}, f_v^{Mf}), where cf^{C_k} is the associated confidence score of v's category C_k derived using Eq. (4), and f_v^{Mf} is the frequency count of item v calculated using Eq. (3) The product items in Y_u^{CMf} are ranked as follows. The items with the highest frequency counts in each category of Y_u^C are selected first and ranked according to their associated confidence scores. Then, the items with the highest frequency counts among the remaining items in each category are selected and ranked according to their associated confidence scores. The process repeats to select and rank items in Y_u^{CMf} by recommending most frequent items from diverse candidate categories.

We compare the number of candidate products $|Y_u^{AR}|$ and the top-N recommendations. Note that Y_u^{AR} is the set of candidate products generated from the product-level association rules. If the number of candidate products $|Y_u^{AR}|$ is higher than the number of top-N recommendations ($|Y_u^{AR}| \geqslant N$), the system will recommend the top-N products from Y_u^{AR}. If the number of candidate products $|Y_u^{AR}|$ is less than the number of top-N recommendations ($|Y_u^{AR}| < N$), but $|Y_u^{AR} \cup Y_u^{CMf}|$ is larger than the number of top-N recommendations ($|Y_u^{AR} \cup Y_u^{CMf}| \geqslant N$), the system will recommend $|Y_u^{AR}|$ products from Y_u^{AR}. The remaining $N - |Y_u^{AR}|$ products for recommendation are selected from Y_u^{CMf}. Note that Y_u^{CMf} is the set of most frequent product items belonging to the associated product categories in Y_u^C.

If $|Y_u^{AR} \cup Y_u^{CMf}|$ is less than the number of top-N recommendations ($|Y_u^{AR} \cup Y_u^{CMf}| < N$), the remaining $N - |Y_u^{AR} \cup Y_u^{CMf}|$ products for recommendation are selected from $Y_u^{Mf} - (Y_u^{AR} \cup Y_u^{CMf})$, which is the set of most frequent items that the target user u has not browsed in the mobile channel and are not in $Y_u^{AR} \cup Y_u^{CMf}$. The products are ranked according to the weighted sum of the frequency counts of the products. (See Fig. 2 for an overview of the process.)

4. Experimental evaluation

4.1. Experiment setup and datasets

We use a dataset obtained from a multi-channel company to conduct our experiment evaluation. The company is a home shopping company that owns television, catalog and Web channels in Taiwan. Because of the rapid development of 3G mobile network in recent years, the company plans to develop a new mobile
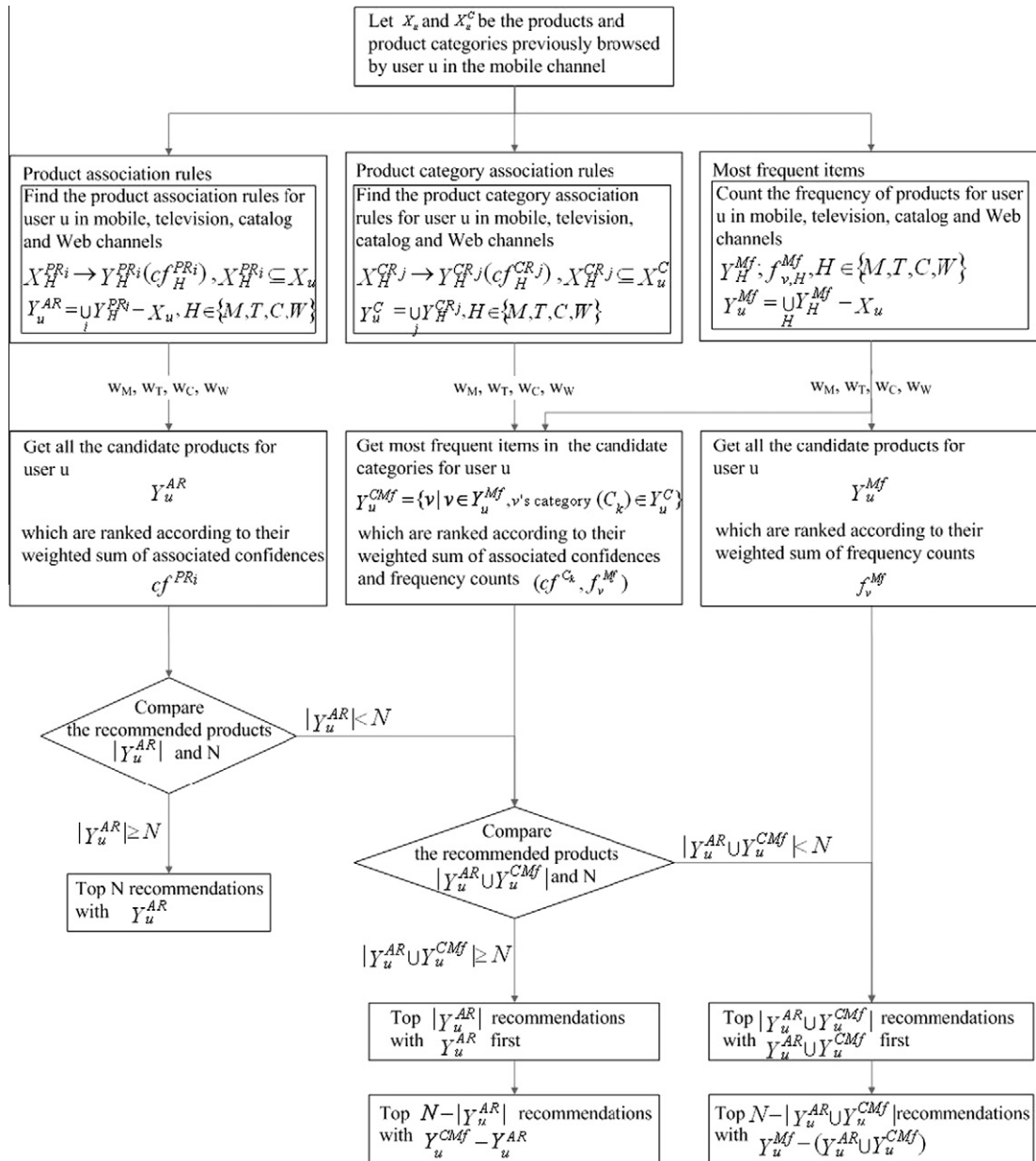
Fig. 2. The recommendation engine.

channel. For the television channel, products are introduced on the channel and viewers can purchase the products by calling a toll-free number.

The mobile channel is an on-line experimental mobile shopping website. The objective is to gather information about the consumption behavior of the new mobile channel users who use their mobile phones to access the website via 2G, 3G, 3.5G or Wi–Fi networks. We collected data from the mobile website, and obtained the transaction data of existing channel users from the case company's CRM system for the period from October 2006 to January 2007. The dataset contained information about 1692 mobile users who owned 184 different models of mobile phones; and the website offered 1416 products divided into 194 categories. The most frequently browsed product categories were mobile phones, lingerie, digital cameras, skincare products, MP3 players, watches, living products, cosmetics, cordless phones and travel promotions. The products offered on the mobile channel were also available on the other three channels.

The dataset was divided as follows: 55% for training, 25% for preliminary analysis and 20% for testing. The training set was used to derive recommendation rules, while the preliminary analytical dataset was used to determine the hybrid weights assigned to mobile, television, catalog, and Web channels based on the quality of the recommendations. There were 1353 mobile users in the training dataset and 339 mobile users in the test dataset.

The support and confidence of the association rules are set to retrieve interesting patterns in datasets. Based on the characteristics of our dataset, the minimum support and confidence values of the association rules were set at 0.004 and 0.4 to find interesting patterns. Both values were higher than those used by Cooley et al. (1999), but lower than those used by Cho et al. (2005).

4.2. Evaluation metrics

Two metrics, precision and recall, are commonly used to measure the quality of a recommendation. They are also used in the

field of information retrieval (Van Rijsbergen 1979, Salton and McGill 1986). Product items can be classified into products that customers are interested in browsing and those that are of no interest. The recommendation method then suggests products of interest to the customers accordingly. The recall metric indicates the effectiveness of a method in locating products of interest, while the precision metric represents customers' levels of interest in the recommended product items.

Recall is the fraction of interesting product items located:

$$Recall = \frac{\text{number of correctly recommended items}}{\text{number of interesting items}} \quad (5)$$

Precision is the fraction of the recommended products that customers find interesting:

$$Precision = \frac{\text{number of correctly recommended items}}{\text{number of recommended items}} \quad (6)$$

The items deemed interesting to customers are the products that the customers browsed in the test set. Correctly recommended items are those that match the interesting items. Because increasing the number of recommended items tends to reduce the precision and increase the recall, the F1 metric is used to balance the tradeoff between precision and recall (Van Rijsbergen 1979). The F1 metric, which assigns equal weights to precision and recall, is calculated as follows:

$$F1 = \frac{2 \times recall \times precision}{recall + precision} \quad (7)$$

The three metrics are computed for each user, and the average value of each cluster is computed. Then, the overall average of all users is calculated to measure the quality of the recommendations.

4.3. Experiment results

4.3.1. The selection of heavy users of the existing channels

Groups of heavy users were selected by comparing the average RFM values of the groups with the average RFM values of a channel. Recall that heavy user groups have smaller average R values and larger average F and M values than the total average RFM values of each channel. First, the R, F, and M values of every user of television, catalog, and Web channels were calculated, and the users of each channel were clustered into groups. When comparing the group average to the total average of a channel, the group average may be higher (↑) or lower (↓) than the total average. Because each R, F, and M value of a group can have two alternative values (i.e., higher (↑) or lower (↓) than the total average) we clustered users based on the three R, F, and M values into 8 groups (2 × 2 × 2).

Second, the heavy user groups were checked (✔), as shown in Table 3, because their average R score was lower (↓) than the total average score of each channel, but the F and M scores were higher

(↑). The clustering results are considered significant (p < 0.05) based on the R, F and M variable differences for television, catalog and Web channels. For example, the clusters of heavy users of the television channel are groups 4 and 5 because: (1) their average R scores (32 and 23) are lower (↓) than the total average R (44); and (2) their average F scores (5 and 9) are higher (↑) than the total average F (4); and their average M scores (16,831 and 40,772) are higher (↑) than the total average M (15,431) of the television channel. Similarly, based on the selection criteria, the clusters of heavy users of the catalog channel are clusters 3 and 6, and the groups of heavy users of the Web channel are groups 2 and 7, as shown by (✔) in Table 3.

Based on the selection criteria, there are two heavy user groups for each channel. Other datasets may yield a different number of groups. The final selection results are shown in Table 4. In this study, we are interested in the major consumption behavior patterns of heavy users in different channels.

4.3.2. Determining channel weights for the hybrid recommendation scheme

The hybrid channel recommendation scheme is based on the weight ratios of the mobile (w_M), television (w_T), catalog (w_C), and Web (w_W) channels ($w_M + w_T + w_C + w_W = 100\%$). The weights are derived as follows. First, the dataset is divided into a training dataset (55%), preliminary analytical dataset (25%) and a test dataset (20%). We use the training dataset to derive the association rules, and use the preliminary analytical data to derive the weights. Second, the weights are determined according to the best recommendation quality that can be achieved under different combinations of weight assignments for the preliminary analytical data. The average number of products browsed in the mobile channel is 3.87, as calculated from the dataset. We use the top four recommendations to determine the hybrid weights of the multiple channels. We adjust the values of the channel weights systematically in increments of 1%. The qualities of the top four hybrid recommendations according to different hybrid weight combinations (w_M, w_T, w_C, w_W) are shown in Fig. 3. The best recommendation quality F1 metric of 0.1573 for the top four recommendations is derived when (w_M, w_T, w_C, w_W) = (60%, 1%, 33%, 6%). We use these weights as the weight ratios of the hybrid recommendation scheme in the experiments.

Table 4
Clusters of heavy users selected in each channel.

Television		Catalog		Web	
Group ID	Users	Group ID	Users	Group ID	Users
4	5694	3	83	2	216
5	4534	6	178	7	325
(4, 5)	10,228	(3, 6)	261	(2, 7)	541

Table 3
The R, F, and M values of the groups of users in each channel.

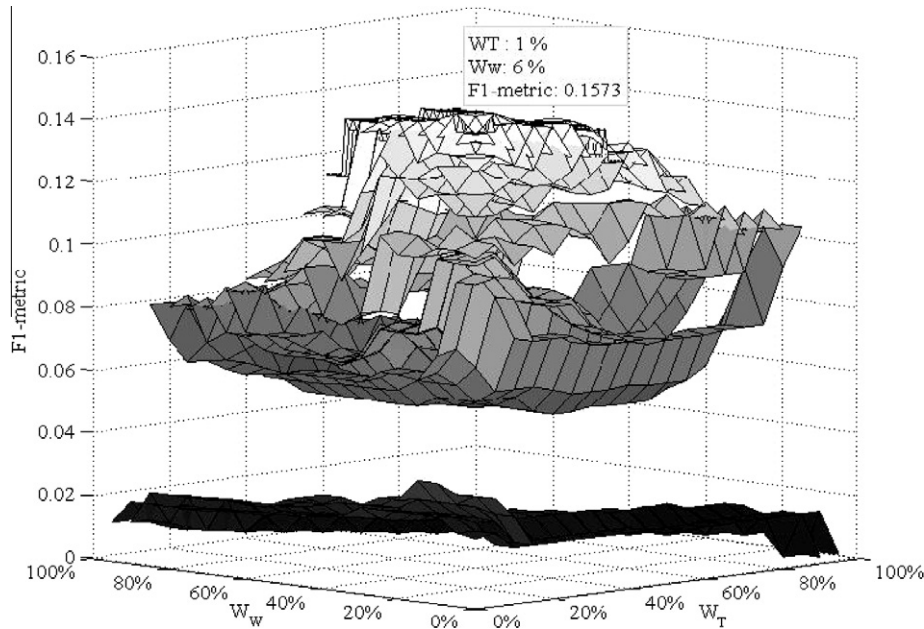| Channel	Television[*]				Catalog[*]				Web[*]			
Group ID	Users	R	F	M	Users	R	F	M	Users	R	F	M
0	1156	80↑	2↓	3932↓	132	54↑	2↓	3951↓	26	82↑	2↓	2677↓
1	4844	40↓	4↑	10,366↓	187	40↓	3↑	5990↓	235	40↓	3↓	5174↓
2	562	93↑	2↓	3059↓	61	63↑	2↓	2917↓	✔ 216	16↓	14↑	38,158↑
3	2013	69↑	3↓	4969↓	✔ 83	19↓	7↑	23,019↑	155	52↑	3↓	4260↓
4	✔ 5694	32↓	5↑	16,831↑	21	68↑	2↓	2566↓	321	28↓	4↓	9822↓
5	✔ 4534	23↓	9↑	40,772↑	101	60↑	2↓	3331↓	262	37↑	3↓	7610↓
6	4032	47↑	3↓	7556↓	✔ 178	32↓	4↑	9101↑	67	61↑	2↓	3069↓
7	2765	60↑	3↓	6287↓	140	47↑	3↑	4577↓	✔ 325	22↓	6↑	16,012↑
Total	25,600	44	4	15,431	903	44	3	7067	1607	33	5	12,909

[*] Significant at the 0.05 level.

Fig. 3. The weight combinations of the hybrid recommendation scheme.

The weight of the television channel is the lowest. To gain insight into the weighting ratios of the channels, we compare the overlap of products purchased from the mobile channel and each of the other channels. Let S_M be a set of products browsed by mobile channel users, and let S_T be a set of products purchased from the television channel by heavy users. The product overlap ratio between the television channel and mobile channel is the ratio of the total number of common products in S_M and S_T to the number of products in S_M. The product overlap ratio between the mobile channel and the other channels is derived similarly. In this case, the product overlap ratios between the mobile channel and the television, Web, and catalog channels are 10.9%, 13.9%, and 18.4%, respectively. Because the ratio between the television and mobile channels is the lowest, we conclude that the consumption behavior of television and mobile channel users is the most dissimilar. Thus, the television channel contributes the least to enhancing the recommendation quality of the mobile channel. Consequently, the weight of the television channel is the smallest.

4.3.3. Evaluation of the hybrid multiple channel recommendation method

We compare the proposed hybrid multiple channel (HMC) recommendation method, with three methods, namely, SC-PCAR, SC-PAR, and KNN-MFI methods. The HMC method recommends products based on the product-level and category-level association rules extracted from multiple channels. The SC-PCAR method is a single channel approach that recommends products based on the product-level and category-level association rules extracted from the mobile channel. The SC-PAR method is a single channel approach that recommends products based on the product-level association rules derived from the mobile channel. Note that if the number of candidate products selected from the association rules is less than N for the top-N recommendations, the HMC, SC-PCAR and SC-PAR methods recommend remaining products based on the most frequently occurring items. The KNN-MFI method is a typical k-NN CF method that recommends the top-N most frequently occurring products of the k-nearest neighbors (similar users) in the mobile channel. Because the average number of users in a user group is 232.5 (=930/4), we choose $k = 200$ as the number

of nearest neighbors. Note that the HMC and SC-PCAR methods cluster users into groups based on the users' similarity derived from the user–product category preference matrix; while the SC-PAR and KNN-MFI methods cluster users into groups based on the users' similarity derived from the user–product preference matrix.

Fig. 4 shows the evaluation results of the four recommendation methods. The SC-PCAR method outperforms the SC-PAR method because the user–product category preference matrix is not as sparse as the user–product preference matrix. Thus, it is possible to find more similar users by using the category preference-based approach. The HMC method generates recommendations based on multiple channels, including the mobile, television, catalog, and Web channels, with the hybrid weighting ratio, that was defined earlier, set at $(w_M, w_T, w_C, w_W) = (60\%, 1\%, 33\%, 6\%)$ for the top-N recommendations.

As shown in Fig. 4, the HMC method outperforms the SC-PCAR, SC-PAR and KNN-MFI methods. In general, the recommendation quality of HMC, SC-PCAR and SC-PAR methods declines after the top four recommendations, as the number of recommended products increases. Recall that association rule-based recommendations are based on the items users browsed previously. In our study, there are only a few recommended products because the average number of products browsed was 3.87. Therefore, the most frequent item recommendations are used to support the association rule recommendations if the number of recommended products is not sufficient. However, the most frequent item-based method does not perform better than the association rule-based recommendation methods, so the recommendation quality deteriorates after the top four recommendations.

5. Discussion

To provide recommendations for users of a new mobile channel, the hybrid weights are determined according to the best recommendation quality that can be derived under different combinations of weight assignments for the preliminary analytical data. The derived hybrid weights (w_M, w_T, w_C, and w_W) of the multiple channels are 60%, 1%, 33%, and 6%, respectively. The weight of
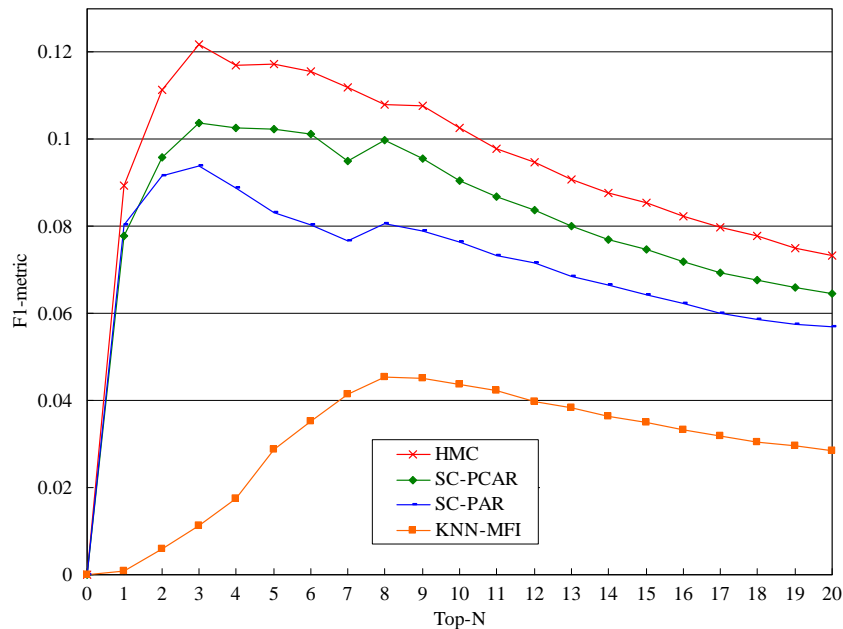
Fig. 4. Evaluation of the HMC, SC-PCAR, SC-PAR, and KNN-MFI methods.

the mobile channel is the largest (60%) because mobile users in the same sales environment (the mobile channel) have the most similar behavior.

The quality of recommendations made to mobile users can be enhanced by referring to the consumption behavior of other channels' users. The relative importance of the recommendations based on the consumption behavior of the users in other channels is 1% for the television channel, 33% for the catalog channel and 6% for the Web channel. There are two possible reasons why the weight of the catalog channel is larger than that of the Web channel. First, unlike the Web channel, the catalog channel ran advertising campaigns to promote the new mobile channel. Thus, more mobile channel users may have migrated from the catalog channel than from the Web channel. The second possible reason is that a mobile phone's interface is limited. Web channel users prefer to surf websites free-of-charge on a large computer screen, rather than paying to browse the same product websites on a small mobile phone screen.

Furthermore, the overlap ratio of products purchased from the catalog and mobile channels is 18.4%, which is higher than the ratio between the Web and mobile channels (13.9%). This implies that the consumption behavior of catalog and mobile channel users is more similar than the consumption behavior of Web and mobile channel users. Thus, the catalog channel contributes more than the Web channel to improving the recommendation quality of the mobile channel. Consequently, the weight of the catalog channel is larger than that of the Web channel.

Because the experimental platform provides a good environment to run trials, it is not necessary to invest a large amount of money in the initial stage (e.g., in advertising and marketing campaigns). It also allows a company to collect information about users' consumption behavior and use it as a reference to develop a commercial run. In the trial run, the retailer does not know about users' product preferences in the early stages of a new channel's development. Deriving the weight composition of multiple channels (e.g., 60%, 1%, 33%, and 6%) for a new channel's users (e.g. a mobile channel) from CRM analysis makes it possible to select products for the new channel. This can be done based on the weights of the multiple channels. For example, the selection of

products for the new channel from the television, Web and catalog channels should be in proportion 1%, 6%, and 33%; the remaining 60% of products should be new lines obtained from other sources. Furthermore, when developing the commercial run of a new channel, the retailer could form a task force to operate the channel. As the other channels are familiar with the product preferences of their respective users as well as their own marketing campaigns, the new task force could be formed based on weights of the existing channels. For example, staff for the new channel should be selected from the television, Web and catalog channels in the following proportions: 1%, 6%, and 33%; and the remaining 60% could be recruited from other companies in the mobile industry (e.g., telecommunications companies).

Although the hybrid multiple channel method outperforms all the other methods, it is more computationally intensive. We compared the tradeoff between recommendation quality and computational time across the four methods. The recommender system includes two subsystems: the off-line batch run and on-line recommendation subsystems. The off-line subsystem deals with data pre-processing, user clustering and association rule mining. When a target user browses the mobile Web, the system will recommend products based on the stored association rules of clusters in the on-line recommendation subsystem. The computation times were compared based on the on-line recommendation phase, as shown in Table 5.

The evaluation was performed on a PC with an Intel Core 2 Quad 2.4 GHz CPU and 4 GB RAM. Table 5 shows the average recommendation qualities and computation times per target user from the top-1 to the top-20 recommendations. The computation times of the HMC, SC-PCAR, SC-PAR, and the KNN-MFI methods were 0.27, 0.16, 0.09, and 0.36 s, respectively. The computation time of the HMC method was longer than those of the single channel methods, but shorter than that of the KNN-MFI method. The multiple-channel method, HMC, required more time because it needed to match more association rules for the multiple channels. The recommendation quality of the hybrid multiple channel method is better than the single channel methods, but the tradeoff for better recommendation quality is an increase in the computation time. However, as the recommendation quality is important to a recommender, the additional computation time is acceptable.

Table 5
Computation times and recommendation qualities of the compared methods.

Method	Computation time	Recommendation quality
HMC	0.27	0.10
SC-PCAR	0.16	0.09
SC-PAR	0.09	0.07
KNN-MFI	0.36	0.03

6. Conclusion

Multiple-channel companies may encounter a number of difficulties when they develop new mobile channels due to the lack of knowledge about users' consumption behavior. Most companies use advertising and marketing campaigns to gather information about users' consumption behavior on a new channel. However, businesses could also obtain such information from the CRM systems of existing channels. In the early stages of new channel's development, there are insufficient purchase orders to determine users' consumption behavior patterns. Thus, it is difficult to find similar users because of the *sparsity problem* inherent in the CF method.

In this paper, we have proposed a hybrid multiple channel method to address the lack of knowledge about the consumption behavior of new channel users and the difficulty of finding similar users. It is assumed that the consumption behavior of new channel users correlates with the consumption behavior of users in existing channels. We conducted experiments to compare the hybrid multiple channel method, two single channel methods, and a typical KNN-based CF method. The experiment results demonstrate that the proposed method outperforms the compared methods. The proposed method mitigates the sparsity problem and improves the recommendation quality by finding more similar users based on the consumption behavior patterns of users in existing channels.

Our study has some limitations. First, we use browsing data rather than purchasing data because there are insufficient purchase orders in a new mobile channel for analysis. Of course, it would be better to make recommendations based on purchasing data, since our objective is to understand users' consumption behavior. Second, users' demographic data, such as gender, age and education, may affect the weighting ratios of channels. However, the system could not identify the consumption behavior patterns of the different demographic groups in each channel because users' demographic data was not available. Third, the reader should note that if a company runs another promotion on a different channel, the weights that we discussed earlier will change. This will have implications for product selection and staff composition, which should be adjusted accordingly.

In the future, we will investigate the reasons that users migrate from one channel to another, as well as how different factors, such as channel advertisements and interfaces, affect users' channel migration decisions. Moreover, the proposed approach could be applied to existing channels (television and catalog channels) to make better recommendations for other channels (the Web). In other words, television and catalog channels could act as auxiliary channels that recommend products for the Web. Such an approach would improve the recommendation quality of electronic commerce conducted by existing channels.

Acknowledgments

References

Agrawal, R., Imielinski, T., and Swami, A. Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, Washington, DC, United States, ACM, 1993, 207–216.

Agrawal, R., and Srikant, R. Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases*, Santiago de Chile, Chile, September 12–15, 1994, Morgan Kaufmann Publishers, San Francisco, CA, 1994, 487–499.

Ansari, A., Mela, C. F., and Neslin, S. A. Customer channel migration. *Journal of Marketing Research*, 45, 1, 2008, 60–76.

Balabanović, M., and Shoham, Y. Fab: content-based, collaborative recommendation. *Communications of the ACM*, 40, 3, 1997, 66–72.

Brunato, M., and Battiti, R. PILGRIM: A location broker and mobility-aware recommendation system. In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, Fort Worth, TX, March 23–26, 2003, IEEE Computer Society Press, Los Alamitos, CA, 2003, 265–272.

Bult, J., and Wansbeek, T. Optimal selection for direct mail. *Marketing Science*, 14, 4, 1995, 378–394.

Chen, M. S., Han, J., and Yu, P. S. Data mining: an overview from a database perspective. *IEEE Transactions on Knowledge and Data Engineering*, 8, 6, 1996, 866–883.

Chircu, A. M., and Mahajan, V. Managing electronic commerce retail transaction costs for customer value. *Decision Support Systems*, 42, 2, 2006, 898–914.

Cho, Y. B., Cho, Y. H., and Kim, S. H. Mining changes in customer buying behavior for collaborative recommendations. *Expert Systems with Applications*, 28, 2, 2005, 359–369.

Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., and Sartin, M. Combining content-based and collaborative filters in an online newspaper. In *Proceedings of the 1999 ACM SIGIR Workshop on Recommender Systems: Algorithms and Evaluation*, University of California, Berkeley, CA, August 19, 1999, ACM Press, New York, NY, 1999.

Cooley, R., Mobasher, B., and Srivastava, J. Data preparation for mining world wide web browsing patterns. *Knowledge and Information Systems*, 1, 1, 1999, 5–32.

Gensler, S., Dekimpe, M. G., and Skiera, B. Evaluating channel performance in multi-channel environments. *Journal of Retailing and Consumer Services*, 14, 1, 2007, 17–23.

Hill, W., Stead, L., Rosenstein, M., and Furnas, G. Recommending and evaluating choices in a virtual community of use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Denver, CO, May 7–11, 1995, ACM Press, New York, NY, 1995, 194–201.

Ho, S. Y., and Kwok, S. H. The attraction of personalized service for users in mobile commerce: an empirical study. *SIGEcom Exchange*, 3, 4, 2003, 10–18.

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

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

Kim, C. Y., Lee, J. K., Cho, Y. H., and Kim, D. H. Viscors: a visual-content recommender for the mobile web. *IEEE Intelligent Systems*, 19, 6, 2004, 32–39.

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.

Lang, K. Newsweeder: Learning to filter netnews. In A. Prieditis, and S. J. Russell (eds.), *Proceedings of the 12th International Conference on Machine Learning*, Tahoe City, CA, July 9–12, 1995, Morgan Kauffman, San Mateo, CA, 1995, 331–339.

Lee, H. J., and Park, S. J. MONERS: a news recommender for the mobile web. *Expert Systems with Applications*, 32, 1, 2007, 143–150.

Li, L., Lee, F., Chen, Y., and Cheng, C. A multi-stage collaborative filtering approach for mobile recommendation. In *Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication*, Suwon, South Korea, ACM Press, New York, NY, 2009, 88–97.

Lim, J., Currim, I., and Andrews, R. Consumer heterogeneity in the longer-term effects of price promotions. *International Journal of Research in Marketing*, 22, 4, 2005, 441–457.

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

Liou, C. H., and Liu, D. R. MPF-based hybrid recommendations for mobile commerce. In *International Conference on Information and Knowledge Engineering*, Las Vegas, NV, July 13–16, 2009, CSREA Press, 2009, 475–480.

MacQueen, J. Some methods for classification and analysis of multivariate observations. In *Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Statistical Laboratory of the University of California, Berkeley, University of California Press, Berkeley, CA, 1967, 281–297.

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

Ngai, E. W. T., and Gunasekaran, A. A review for mobile commerce research and applications. *Decision Support Systems*, 43, 1, 2007, 3–15.

Omran, M. G. H., Engelbrecht, A. P., and Salman, A. An overview of clustering methods. *Intelligent Data Analysis*, 11, 6, 2007, 583–605.

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

Punj, G., and Stewart, D. W. Cluster analysis in marketing research: review and suggestions for application. *Journal of Marketing Research*, 20, 1983, 134–148.

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*, Chapel Hill, NC, USA, ACM Press, New York, NY, 1994, 175–186.

Salton, G., and McGill, M. J. *Introduction to Modern Information Retrieval*. McGraw-Hill, New York, NY, 1986.

Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. Analysis of recommendation algorithms for e-commerce. In *Proceedings of the Second ACM Conference on Electronic Commerce*, Minneapolis, MN, ACM Press, New York, NY, 2000, 158–167.

Schröder, H., and Zaharia, S. Linking multi-channel customer behavior with shopping motives: an empirical investigation of a German retailer. *Journal of Retailing and Consumer Services*, 15, 6, 2008, 452–468.

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*, Denver, CO, ACM Press, New York, NY, 1995, 210–217.

Thomas, J., and Sullivan, U. Managing marketing communications with multichannel customers. *Journal of Marketing*, 69, 4, 2005, 239–251.

Tiernan, B. *The Hybrid Company: Reach All Your Customers through Multi-Channels Anytime, Anywhere*. Dearborn Trade, Chicago, IL, 2001.

Van den Poel, D., and Leunis, J. Consumer acceptance of the internet as a channel of distribution. *Journal of Business Research*, 45, 3, 1999, 249–256.

Van Rijsbergen, C. J. *Information Retrieval*. Buttersworth, London, 1979.

Venkatesh, V., Ramesh, V., and Massey, A. P. Understanding usability in mobile commerce. *Communications of the ACM*, 46, 12, 2003, 53–56.

Wu, J. H., and Wang, Y. M. Development of a tool for selecting mobile shopping site: a customer perspective. *Electronic Commerce Research and Applications*, 5, 3, 2006, 192–200.

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