ELSEVIER

Expert Systems with Applications 28 (2005) 359–369

Expert Systems with Applications

www.elsevier.com/locate/eswa

Mining changes in customer buying behavior for collaborative recommendations

Yeong Bin Cho[a,*], Yoon Ho Cho[b,1], Soung Hie Kim[c]

[a]*Department of e-Business, Far East University, 5 san Wangjang, Gamgok, Eumsung, Chungbuk, 369-851, South Korea*
[b]*School of e-Business, Kookmin University, 861-1 Jungnung-dong, Sungbuk-gu, Seoul 136-702, South Korea*
[c]*Graduate School of Management, Korea Advanced Institute of Science and Technology, 207-43 Cheongryangri-Dong, Dongdaemun, Seoul, 130-012, South Korea*

Abstract

The preferences of customers change over time. However, existing collaborative filtering (CF) systems are static, since they only incorporate information regarding whether a customer buys a product during a certain period and do not make use of the purchase sequences of customers. Therefore, the quality of the recommendations of the typical CF could be improved through the use of information on such sequences.

In this study, we propose a new methodology for enhancing the quality of CF recommendation that uses customer purchase sequences. The proposed methodology is applied to a large department store in Korea and compared to existing CF techniques. Various experiments using real-world data demonstrate that the proposed methodology provides higher quality recommendations than do typical CF techniques, with better performance, especially with regard to heavy users.

Keywords: Recommender systems; Purchase sequence; Collaborative filtering; SOM; Association rules

1. Introduction

Recommender systems have been a recent focus of researchers and practitioners. Many companies hope that the use of recommender systems may be a means of surviving in a competitive environment. Recommender systems are particularly suited to retail business, as compared to other types of business, since retail markets are distinguished by several characteristics, such as repeated buying over a particular time horizon, large numbers of customers, and a wealth of information detailing past customer purchases (Schmittlein & Peterson, 1994).

In general, retail companies operate purchase databases in a longitudinal way, such that all transactions are stored in chronological order. A record-of-transaction database typically contains the transaction date for and the products bought in the course of, a given transaction. Usually, each record also contains a customer ID, particularly when the purchase was made using a credit card or a frequent-buyer card. Therefore, the purchasing sequence of a customer in the database who has made repeat purchases can easily be determined. This purchase sequence provides a description of the changes in a customer's preferences over time. However, in our domain of knowledge, there has been little study of the question of whether recommendations based on purchase sequences may be more accurate than existing recommender system predictions, based on non-sequential patterns. In this study, for the purpose of enhancing the quality of recommendations, we propose a new methodology that considers the way in which a customer's purchase sequence evolves over time.

1.1. Motivation

To date, a variety of recommender systems (Balabanović & Shoham, 1997; Basu, Hirsh, & Cohen, 1998; Hill, Stead, Rosenstein, & Furnas, 1995; Lawrence, Almasi, Kotlyar, Viveros, & Duri, 2001; Resnick, Iacovou, Suchak, Bergstrom, & Riedl, 1994; Sarwar, Karypis, Konstan, & Riedl,

* Corresponding author. Tel.: +82 2 958 3864; fax: +82 2 958 3604.
E-mail addresses: ybcho@kgsm.kaist.ac.kr (Y.B. Cho), www4u@kookmin.ac.kr (Y.H. Cho), seekim@kgsm.kaist.ac.kr (S.H. Kim).
[1] Tel.: +82 2 910 4950; fax: +82 2 910 4519.

0957-4174/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2004.10.015

2001; and Shardanand & Maes, 1995) has been developed. Collaborative filtering (CF) has thus far been the most successful recommendation technique and has been used in a number of different applications, such as in the recommendation of web pages, movies, articles and products (Hill et al., 1995; Resnick et al., 1994; and Shardanand & Maes, 1995). Collaborative filtering works by recommending products to a target customer through a process of identifying people who share similar preferences for products and looking for those products that target customers are most likely to purchase.

The recommendation processes of typical collaborative filtering in retail business consist of the following three steps (Sarwar, Karypis, Konstan, & Riedl, 2000, 2001).

(1) *Customer profile construction*

The purchase transaction records of a customer for a certain period are used to build a customer profile describing his or her likes and dislikes. The system represents the customer profile, A, such that a_{ij} is one if a customer i has purchased a product j, and zero, otherwise.

(2) *Neighborhood formation*

This is the most important part of the CF-based recommender systems. The system finds a set of customers, known as neighbors, who, in the past, have exhibited similar behavior (i.e. bought a similar set of products), through calculating the correlations among customers for the customer profile. A set of K customers is usually found (a neighborhood of size K), which is formed according to the degree of similarity between each of the neighbors and a target customer.

(3) *Recommendation generation*

Once a neighborhood is formed for a target customer, the system generates a set of the top N products that the target customer is most likely to purchase, by searching for products that the neighbors have purchased and that the target customer has not yet purchased.

As mentioned above, typical collaborative filtering is static, since it only makes use of information relating to whether the customer bought a product during a certain period, and does not use information on the purchase sequences of customers in the determination of the neighbors of a target customer. However, customers in retail business are not static, and their buying behavior changes over time. Thus, the quality of the recommendation of the typical CF could be further improved through the use of the available information on the purchase sequences of customers.

To illustrate the importance of this potential improvement in accuracy, let us consider the following example. Table 1 presents typical transaction records for a retail company and the customer profile is provided in Table 2. This example determines products that target customer ID011 is likely to buy, using transaction records for consumers CID001 through CID011.

Assume that the typical CF algorithm is used for solving the problem, that the neighborhood size (K) is three, and that the number of products recommended (N) is two. The typical collaborative filtering algorithm considers the correlation of preferences between the target customer and the other customers. All of the four customers, CID011, CID001, CID002 and CID003, commonly bought 'Perfumes' and 'Skincare Products.' The similarities between CID011 and the other three customers are equivalent; that is, the Pearson correlation coefficient is 0.67. Therefore, a recommender system based on the collaborative filtering algorithm will determine that CID001, CID002 and CID003 are the nearest neighbors and have the same preferences as the target customer. However, it is quite difficult to select two products that should be recommended to CID011, because CID001, CID002 and CID003 each purchased different additional products: 'Dresses,' 'Shoes,' and 'Knits,', respectively. In this case, two different products to be recommended to CID011 would have to be selected randomly. Accordingly, the recommendations would not necessarily be very appropriate for the preferences of the target customer.

Table 1
Transactions for illustrative example

Customer ID	Transaction time	Products category bought
001	July 11 2000	Perfumes
001	August 17 2000	Skincares
001	September 14 2000	Dresses
002	July 15 2000	Perfumes
002	August 13 2000	Shoes
002	September 25 2000	Skincares
003	July 19 2000	Skincares
003	August 22 2000	Perfumes
003	September 18 2000	Knits
:	:	:
:	:	:
010	September 27 2000	Dresses
011	July 22 2000	Perfumes
011	August 26 2000	Skincares

Table 2
A customer profile for typical CF

CID	Perfumes	Skincares	Knits	Dresses	Shoes
001	1	1	0	1	0
002	1	1	0	0	1
003	1	1	1	0	0
004	1	0	0	1	0
005	1	0	0	0	1
006	0	1	1	0	0
007	0	1	0	1	0
008	0	1	0	0	1
009	0	0	1	0	0
010	0	0	0	1	0
011	1	1	0	0	0