Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 100 (2016) 136 – 146

Procedia Computer Science

Conference on ENTERprise Information Systems / International Conference on Project MANagement / Conference on Health and Social Care Information Systems and Technologies, CENTERIS / ProjMAN / HCist 2016, October 5-7, 2016

Product recommendation based on shared customer's behaviour

Fátima Rodrigues[a,b*], Bruno Ferreira[b]

[a]*GECAD, Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development*
[b]*Institute of Engineering – Polytechnic of Porto (ISEP/IPP)*

Abstract

Today consumers are exposed to an increasing variety of products and information never seen before. This leads to an increasing diversity of consumers' demand, turning into a challenge for a retail store to provide the right products accordingly to customer preferences. Recommender systems are a tool to cope with this challenge, through product recommendation it is possible to fulfill customers' needs and expectations, helping maintaining loyal customers while attracting new customers. However the huge size of transactional databases typical of retail business reduces the efficiency and quality of recommendations. In this paper a hybrid recommendation system that combines content-based, collaborative filtering and data mining techniques is proposed to surpass these difficulties. The recommendation algorithm starts to obtain similar groups of customers using customer lifetime value. Next an association rule mining approach based on similar shopping baskets of customers of the same cluster, in a specific time period is implemented in order to provide more assertive and personalized customer product recommendations. The algorithm was tested with data from a chain of perfumeries. The experimental results show that the proposed algorithm when compared with a base recommendation (made solely on past purchases of customers) can increase the value of the sales without losing recommendation accuracy.

Peer-review under responsibility of the organizing committee of CENTERIS 2016

Keywords: clustering; market basket; association rules; product recommendation

* Corresponding author. Tel.: +351 228 340 500; fax: +351 228 321 159.
E-mail address: mfc@isep.ipp.pt

Peer-review under responsibility of the organizing committee of CENTERIS 2016
doi:10.1016/j.procs.2016.09.133

1. Introduction

Consumers are permanently involved in multi-category decision-making. In a retail context, such multi-category decision processes result in the shopping-baskets that comprise the set of items that the consumers purchase on one visit to the store. Both on-line and off-line retailers are traditionally interested in understanding the composition of their customers' market baskets, since valuable insights for designing micro-marketing and/or targeted cross-selling programs can be derived[1]. Recommender systems are technologies that assist businesses to implement such strategies. Schafer[2] presented a detailed taxonomy of recommender systems in e-commerce, and determined how they can increase the probability of cross-selling; establish customer loyalty; and fulfill customer needs by discovering products in which they may be interested. The need to reduce information overload by retrieving the most relevant information and services from a huge amount of data, and also, the development of recommendation approaches and techniques, has determined a rapid proliferation of recommender systems grouped into eight main domains[3]: e-government, e-business, e-commerce/e-shopping, e-library, e-learning, e-tourism, e-resource services and e-group activities. Recommender systems are usually classified based on how recommendations are made[4]. A content-based recommender system is based on similar items to those a given user has liked in the past[5,6]. A collaborative filtering makes recommendations based on items owned by users whose taste is similar to those of the given user[7,8]. Combining content-based and collaborative recommendations originate hybrid approaches[9], which are commonly used, considering that both types of recommendations may complement each other.

In the recommender system here described two sources of information are used. First, it is used clustering to obtain groups of customers with similar interests based on prior purchase patterns. Second, rule association mining is performed on baskets of the same cluster, in order to derive relationships between products. Since these relationships are based on purchases of similar customers that have also purchased, in the same time period, at least one same product, it is expected to identify additional product relationships that are not captured using only the past baskets of the customer. In summary, this recommender system uses collaborative filtering combined with the ideas from content-based filtering.

There are usually quite a lot of products to be considered in a recommender system. It would be very inefficient if every product needs to be considered before making recommendations. Dimensionality reduction techniques have been incorporated to produce quickly quality recommendations for large-scale problems[10,11]. However these systems have some disadvantages, for example, require extra attributes about users or products to group the users into clusters and require the number of clusters be given in advance, which is a big burden on the user. With the proposed approach, clustering is done totally based only on derived attributes about products purchased by costumers, without the necessity of collecting extra attributes about customers and products. Besides, when selecting the baskets for recommending products, we consider only baskets of clients of the same cluster, bought in a specific time period, resulting in a much greater reduction of the number of products to consider. Due to dimensionality reduction on the number of products, the processing time for making recommendations by our approach is much reduced. Experimental results show that the proposed recommender system can enhance the recommendations with a good performance without compromising the recommendation quality.

The remainder of this paper is organized as follows: in section 2 a brief explanation of concepts and algorithms used to implement the recommender system is made. In the following section the hybrid recommendation algorithm is explained. In the next section details of the recommender system are provided. Section 6 presents the experimental results and in last section conclusions and suggestions for future work are disclosed.

2. Background

2.1. Customer Lifetime Value Analysis and RFM Evaluation

Customer lifetime value is typically used to identify profitable customers and to develop strategies to target customers. The RFM (recency, frequency and monetary) model is the most widely used model to characterize customers due to its simplicity and good predictive capabilities. "Recency" represents the time since the last purchase, a lower value corresponding to a higher probability of the customer making a repeat purchase. "Frequency" denotes the number of purchases within a specified time period; higher frequency indicates higher loyalty. "Monetary" means the amount of money spent in this specified time period, a higher value indicating a