Expert Systems with Applications 37 (2010) 4066–4072

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A temporal data mining approach for shelf-space allocation with consideration of product price

Maryam Nafari *, Jamal Shahrabi

Amirkabir University of Technology, Department of Industrial Engineering, 424, Hafez Avenue, Tehran, Iran

ARTICLE INFO

Keywords:
Data mining
Shelf-space allocation
Retail knowledge
Pricing

ABSTRACT

Marketing research has suggested that the in-store stimuli such as shelf-space allocation and product assortment have great influence on customer buying behaviour and may induce sales by maximizing impulse buying and cross-selling. The previous studies, however, have ignored the effect of product price in shelf-space arrangement. That is, they study the relationship between products and their simultaneous sales in a static fashion, disregarding the dynamic changes of their prices. The changes in product price may change the association between products such as complementarity and substitutability relationships. Consequently, it would affect the applied strategies of shelf allocation. In this paper a new approach is developed to optimally select and price the products and allocate them to shelf space with consideration of their prices. This paper takes advantage of data mining techniques, association rules, to find relationships between products regarding to their prices. Finally, to show the efficiency and effectiveness of the proposed approach, the experiment on real world data is executed.

1. Introduction

With limited shelf space and abundance of current and new products, retailer must select, price, and allocate the products to the available shelf space. Many retailers are now turning to product assortment and shelf-space allocation models to maximize their total profit. In the past three decades, there is substantial literature on issues involving shelf-space allocation and product assortment. Rajaram (2001) defines product assortment planning as "the process to determine the number and types of products in a line, which is carried out by retailers". Shelf space is, also, an important resource for retail stores since a great quantity of products compete the limited shelf space for display. Product assortment and shelf-space allocation are two crucial issues in retailing which can affect the customers' purchasing decisions. Through the proficient shelf-space management, retailers can improve return on inventory and consumer's satisfaction, and therefore increase sales and margin profit.

Several models and approaches have been developed to deal with the product assortment and the shelf-space allocation problems. In many experimental studies, the space elasticity is one of the approaches which has been widely used to estimate the relationship between sales and allocated space. Space elasticity was defined as the ratio of relative change in unit sales to relative change in shelf space. Curhan (1972) took a large sample from store experiments, and found the average value of 0.212 for space elasticity. Doyle and Gidengil (1977) had summarized the results from many studies about space elasticity and pointed out the difficulties that might be encountered as these approaches are applied in retail practice. Therefore, commercial approaches and experimental approaches fail to evaluate the aggregate store performance of their allocation solution. Thus, optimization models with an application orientation are worthy of consideration. Anderson and Amato (1974) formulated the shelf-space management model as a knapsack problem and took only the direct elasticity into their model to simultaneously optimize the product assortment and shelf-space allocation. Also, Hansen and Heinsbroek (1979) proposed non-linear mathematical programming model which incorporated main demand effect with cost effect and made the model more complete. However, this study has, also, not considered the cross effect among products within the store. Corstjens and Doyle (1981) broadened the model to consider both space elasticity and cross-elasticity. They applied a polynomial functional form of demand, and they found a set of solutions by using signomial geometric programming. Also, an optimization model of Bultez and Naert (1988) utilized marginal analysis and took into account the interdependencies within product groups and across groups. Borin, Farris, and Freeland (1994) considered the main effects and cross effects of substitute items. In their constrained optimization models, objective function is the return on investment of inventory. Due to the complexity of model and non-linearity of objective function, they suggested a meta-heuristic,

* Corresponding author. Tel.: +98 2122497336.
E-mail addresses: Maryam_nafar@aut.ac.ir (M. Nafari), jamalshahrabi@aut.ac.ir, shahrabi@dal.ca (J. Shahrabi).

0957-4174/$ - see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.eswa.2009.11.045

simulated annealing, as a solution methodology. A critical drawback for applying this model is that it needs to estimate a large number of parameters. The number of estimated parameters in Borin et al. (1994) is $2n + n^2$, in which n is the number of possible products. After that Yang and Chen (1999) and Yang (2001) proposed a space allocation model, a type of multi-constraint knapsack problem, incorporating the main and cross effects of demand as well as the location effects. Only for simplified versions of the original model, he found an optimal solution. Rajaram (2001) applied demand forecasts derived from historical sales patterns, and also constructed a non-linear integer-programming model to make the product assortment planning. Due to the high complexity in the model, heuristics were developed by Rajaram to resolve this problem. Although some existing research papers on the product assortment and space allocation problems (e.g., Borin & Farris, 1995; Borin et al., 1994) use return on inventory as the objective and take stock outs into consideration, they ignore the inventory-related decisions and do not explicitly include the conventional inventory control decisions as variables (Urban, 1998). Urban (1998) proposed integrated models of inventory control and shelf-space allocation problems. The above-mentioned models, however, did not consider the location effects. Hwang, Choi, and Lee (2005) also proposed an integrated mathematical model, which combines the shelf-space allocation model and inventory control model with the objective of maximizing the retailer's profit. Due to the complexity of the integrated model, they proposed a gradient search heuristic and a genetic algorithm to resolve the model. Moreover, Dréze et al. (1994) made a series of field experiments and found that location of the product within a display, especially the level of shelf on which the product is displayed in case of multi-level shelf, has a significant effect on sales. On the other hand, he concluded that changes in the number of facings allocated to a brand had much less impact as long as a minimum stock is maintained.

To overcome the high cost of conducting experiments to measure parameters in space elasticity, Brijs, Swinnen, Vanhoof, and Wets (1999) proposed an association rule based approach to select the most interesting products in convenience stores with consideration of their cross-selling. Brijs, Goethals, Swinnen, Vanhoof, and Wets (2000) further generalized his model to deal with large baskets and category management in practice. However, Brijs et al. (1999, 2000) only explored the product assortment problem. Therefore, they did not take the shelf space requirement of selected products. Chen and Lin (2007) took advantage of Brijs's model and proposed an approach to resolve product assortment and shelf-space allocation. They discovered the relationship between products items, between product subcategories and between product categories by the use of multiple-level association rules mining. Then, in the process of product assortment the profits of frequent itemsets are considered. In their approach, the products, subcategories, and categories which are frequently bought together should be displayed as close as possible. Finally, the product display locations were determined by considering the relationships between categories, subcategories, and between items.

According to the previous studies, in the procedure of shelf-space allocation, the product price is not taken into account. However, it would have great impact on consumer purchase behaviour. As it is illustrated in experimental study of this paper, properly product arrangement based on their relationship and their prices would definitely have positive effects on cross-selling. It is necessary to say that, the presented approach is fundamentally dynamics. That is, the products display is changing dynamically based on the changes in product prices and consequently their relationship. In this paper, we utilize the previous studies and present a new approach in which products are allocated to shelf space according to their both relationships and prices. The proposed shelf-space management procedure begins with discovering multi-level association rules so that the relationship between product category, subcategory and associations between items with consideration of their price are exploited. We make use of the algorithm proposed by Nafari and Shahrabi (2008) called Apriori-TdMl and we alter it in a way that just in the first and second level, category and subcategory, the association rules are discovered disregarding their price. On the contrary, in the third level, the association rules are found with consideration of items price. After finding association rules in each level, the proposed approach by Brijs et al. (2000) is used and the products which are worth investing are select. However, in the new approach the price information is added to Brijs model to find the best products with consideration of their price. Ultimately, the subcategories and categories are allocated to shelves based on their relationships and items are arranged based on their specific relationship with specific price combination. In fact, eventually subcategories and categories which are frequently bought together can be displayed much closer. Moreover, items can appropriately be allocated with respect to their price.

The rest of the paper is organized as follows. In Section 2, a modified algorithm for discovering multi-level association rules is presented. In Section 3, a new approach for finding best products and their optimal price is developed. In Section 4, the shelf-allocation procedure with consideration of products price is proposed. In Section 5, an experimental study was carried out and the results are discussed. Finally, in Section 6 the conclusion is drawn.

2. Discovering association rules

Association rules approach studies the frequency of items occurring together in transactional database based on two thresholds called *support* and *confidence* which identifies the frequent item sets and association rules, respectively. An association rule is expressed as the form $A \rightarrow B$, where A and B are sets of items, such that the presence of A in a transaction will imply the presence of B in the same transaction (Agrawal et al., 1993; Agrawal & Srikant, 1994).

Since for displaying product we need the association between product categories and subcategories, the algorithm for discovering multi-level association rules, Apriori-TdMl, (Nafari & Shahrabi, 2008) is modified so that it effectively find the patterns in the two first levels disregarding product price. But, we use the Apriori-TdMl method in the third level and find the pattern with consideration of their price. The reason is that, just in the third level the product price is concrete and it is not relative. Since in the refined algorithm price information is only added to items in the third level, its run time is much shorter. We call the refined algorithm, RApriori-TdMl and it is illustrated in Fig. 1 and its parameter is defined in Table 1. In our new algorithm, frequent itemsets and association rules ignoring their price are discovered in the two first levels. Discovering price patterns and association rules in the third level is helpful for properly arranging the products. In fact, the shelf-space management can change the product location based on their prices because of its effects on the relationship between products. Moreover, it considered the temporal characteristics of product price. The strategy of reduced minimum support which is generally used in mining multi-level association rules help us to find enough rules in each level, especially in the third level, which price information is added to patterns and makes the retail transaction data become sparse.

3. Product selection

By the use of price patterns which is found in the third level, we are able to determine which product with how much price are