Available online at www.sciencedirect.com
ScienceDirect
Decision Support Systems 42 (2006) 1503–1520
Decision Support Systems
www.elsevier.com/locate/dss

A data mining approach for retail knowledge discovery with consideration of the effect of shelf-space adjacency on sales

Yen-Liang Chen [a,*], Jen-Ming Chen [b,1], Ching-Wen Tung [a]

[a] Department of Information Management, National Central University, 300 Jhongda Road, Jhongli City, Taiwan 32001, Republic of China
[b] Institute of Industrial Management, National Central University, 300 Jhongda Road, Jhongli City, Taiwan 32001, Republic of China

Received 3 December 2003; received in revised form 3 December 2005; accepted 28 December 2005
Available online 21 February 2006

Abstract

Recent marketing research has suggested that in-store environmental stimuli, such as shelf-space allocation and product display, has a great influence upon consumer buying behavior and may induce substantial demand. Prior work in this area, however, has not considered the effect of spatial relationships, such as the shelf-space adjacencies of distinct items, on unit sales. This paper, motivated in great part by the prominent beer and diapers example, uses data mining techniques to discover the implicit, yet meaningful, relationship between the relative spatial distance of displayed products and the items' unit sales in a retailer's store. The purpose of the developed mining scheme is to identify and classify the effects of such relationships. The managerial implications of the discovered knowledge are crucial to the retailer's strategic formation in merchandising goods. This paper proposes a novel representation scheme and develops a robust algorithm based on association analysis. To show its efficiency and effectiveness, an intensive experimental study using self-defined simulation data was conducted. The authors believe that this is the first academically researched attempt at exploring this emerging area of the merchandising problem using data mining.

Keywords: Data mining; Association rules; Shelf-space management; Retailing management; Merchandising

1. Introduction

Over the past three decades, merchandisers have relied heavily on marketing stimuli to increase their sales volume. Marketing research has suggested that the in-store stimuli (such as display, layout, atmosphere, and shelf-space arrangement) has a great influence on consumer buying behavior and may encourage sales by maximizing impulse buying and cross-selling. For example, a factorial experiment revealed that an in-store display of an item creates excitement and increases the average amount purchased [13]. A case study showed that effective store layout could stimulate demand to the point of doubling the sales rate by making it easier to find items and creating a positive image or feeling [43]. Psychological experiments show that elements of a store's atmosphere, like lighting, color, music, and aisle width, may have a greater influence on shopping behavior than characteristics of the product itself [19,27]. The positive curvilinear relationship between

* Corresponding author. Tel.: +886 3 426 7266; fax: +886 3 425 4604.
E-mail addresses: ylchen@mgt.ncu.edu.tw (Y.-L. Chen), jmchen@mgt.ncu.edu.tw (J.-M. Chen).
[1] Tel.: +886 3 425 8192; fax: +886 3 425 8197.

0167-9236/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2005.12.004

an item's shelf-space and its sales has been verified empirically for a wide variety of consumer goods [16,24].

Prior studies on in-store environmental stimuli, or merchandising techniques, have not considered the effect of spatial relationships, such as the shelf-space adjacencies of distinct items, on unit sales. As the famous beer and diapers example reveals [4], not considering the effects of side-by-side displays of items commonly purchased together may cause a retailer to miss out on tremendous revenue potential. The visual effect of adjacency can stimulate impulse purchases that account for 70% of buying decisions in a supermarket [5]. In light of this potential, this paper attempts to discover the implicit, yet meaningful, relationship between the relative spatial "distance" of displayed products and the items' unit sales in a retail store using data mining techniques. Special focus is placed on building a novel representation scheme for the historical transaction data and on developing an efficient and robust algorithm for knowledge mining. The proposed approaches measure and classify the effects of spatial adjacency of distinct items on increased sales.

Our data mining approach differs from the well-known market basket analysis in several aspects. Our approach takes the product-to-shelf assignment information into account and incorporates the transaction time into the data stream dynamically. In contrast, the market basket analysis mainly determines what products customers purchase together in a static fashion, disregarding the product-to-shelf information. Therefore, our approach requires a differently formatted data-warehouse that must have spatial and temporal contents, and demands a more sophisticated algorithm for mining the dynamic transaction data effectively and efficiently. For the purposes of this study, we have assumed that all the required historical data is readily available and has been stored in the data-warehouse. To manage sales, a retail company must continue storing information in its databases on when the items are on the shelf and where they are placed. By properly pre-processing and integrating this data using the standard ETL tools provided by data warehouse software, one can obtain all this information in the format specified in this paper. Based on the simplified scenario, the problem is defined, the representation scheme is proposed, and the mining algorithm is developed using the association rules.

The more sophisticated mining techniques, like the one discussed in this paper, are superior to traditional approaches in retail knowledge discovery, such as the market basket analysis or frequent-buyer program [7]. In some cases, the fact that items sell well together is obvious, such as laundry detergent and fabric softener [34], greeting cards and seasonal candy, or coffee and coffee makers. Occasionally, however, the fact that certain items would sell well together is far from obvious, such as in the case of diapers and beer [4] or bottled juice and cold remedies [7]. The true reason behind such purchase patterns remain unclear; it may be due to their close proximity in shelf location or other consumer behavior we have yet to discover. In this regard, the market basket analysis or frequent-buyer program is unable to provide satisfactory results. The proposed scheme attempts to dig for obscure clues by introducing the spatial relationship and transaction time information into the mining techniques.

These approaches are separated not only by function and required data content but also by their managerial implications. The frequent-buyer program focuses on a consumer-level analysis to investigate individual purchase habits, such as a customer's affinity analysis. The market basket analysis, on the other hand, is mainly devoted to a market-level analysis. Due to the crossover of consumer traffic among stores, market-level analysis classifies the demand relationships across product categories into complement, independent, or substitute within the consumer choice process [39]. In contrast, our mining scheme is a store-level merchandise technique that identifies the effects of shelf-space proximity on unit sales over a finite time horizon and classifies the patterns as positive, independent, or negative. A positive pattern refers to a positive effect of shelf arrangement of distinct products on sales, meaning that placing specific product assortments side-by-side or in close proximity will trigger supplemental sales due to factors such as increased impulse purchasing and cross-selling. The negative pattern refers to a negative effect of such a spatial relationship.

The discovered relationship between shelf patterns and unit sales is crucial for effective decision-making and strategic planning in merchandising goods. For example, retailers can rearrange their shelf-space to increase impulse buying, and the store manager can measure the effect on revenue. The beer and diapers example has suggested the potential of utilizing spatial relationships [4]. In another related example [30], Seven-Eleven Japan has a policy of adjusting its store layout and product placement multiple times every day to reflect the changing purchase patterns at different hours of the day, so that customers can easily find their favorite items. In this regard, this paper is very likely the first academic research that explores this emerging, high-potential area.