Available online at www.sciencedirect.com
SCIENCE DIRECT®
Omega 33 (2005) 333 – 343
omega
The International Journal of Management Science
www.elsevier.com/locate/dsw

# An association-based clustering approach to order batching considering customer demand patterns

Mu-Chen Chen[a,*], Hsiao-Pin Wu[b]

[a] Department of Business Management, Institute of Commerce Automation and Management, National Taipei University of Technology, 1, Section 3, Chung-Hsiao E. Road, Taipei 106, Taiwan, ROC
[b] Information System Department, Chinese Petroleum Corporation, Taiwan, ROC

Received 6 January 2003; accepted 8 May 2004

**Abstract**

Research on warehousing systems has gained interest since the 1980s, reflecting the fact that supply chain management has pursued a demand-driven organization with high product variety, small order sizes, and reliable short response times throughout the supply chain. This market trend has affected warehouse management and operations tremendously. Order batching in a warehouse attempts to achieve high-volume order processing operations by consolidating small orders into batches. Order batching is an essential operation of order processing in which several orders are grouped into batches. This paper describes the development of an order batching approach based on data mining and integer programming. It is valuable to discover the important associations between orders such that the occurrence of some orders in a batch will cause the occurrence of other orders in the same batch. An order-clustering model based on 0–1 integer programming can be formulated to maximize the associations between orders within each batch. From the results of several test problems, the proposed approach shows its ability to find quality solutions of order batching problems.


*Keywords:* Warehousing; Order batching; Data mining; Association rule; 0–1 integer programming

## 1. Introduction

Recently, warehouse managers have had to react constantly to market changes because customers have been gaining more power to influence the market structure. Warehouses develop an essential connection among suppliers, distributors and customers in the supply chain. The supply chain management has pursued a demand-driven organization with high product variety, small order sizes, and reliable short response times throughout the supply chain. This market trend has affected warehouse management and operations tremendously [1]. Warehousing involves all movement of goods within warehouses or distribution centers including receiving, storage, order picking, accumulation, sorting and shipping. In general, warehousing necessitates an enormous amount of product movement. The development of efficient warehousing operations may bring in considerable reduction in product movement.

Four methods are usually used to reduce travel times or distances by means of more efficient control mechanisms in warehouses [2]. They include (1) determining good order picking routes; (2) zoning the warehouse (i.e., an order picker picks only that part of an order that is in his/her assigned zone); (3) assigning products to the correct storage locations; and (4) assigning orders to batches. In an order picking operation, the order pickers may apply the *single-order picking* (picking one order at the same time) or *batch picking* (picking a group of orders simultaneously). A batch is a group of orders that is picked in a single tour [1]. The orders in a batch may not exceed the storage capacity of

* Corresponding author. Tel.: +886-2-2771-2171x,3417; fax: +886-2-2776-3964.
*E-mail address:* bmcchen@ntut.edu.tw (M.-C. Chen).

0305-0483/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.omega.2004.05.003

the order picking vehicle. A higher efficiency and productivity can be achieved by means of batch picking. For this approach, orders must be consolidated before the picking operations. Warehouse managers are interested in finding the most economical way of picking customer orders which minimizes the costs involved in terms of distance traveled and/or time spent.

Batch picking may be more cost-effective in warehousing as a result of the decrease of average travel time or average travel distance per order. However, a relatively small amount of research on the subject of order batching can be found in the literature [3]. Heuristic methods were commonly developed to resolve the problem of order batching since its mathematical formulation is very complicated. Some batching heuristics [3–9] have been introduced in the literature for minimizing the total distance traveled by operators and/or S/R (storage/retrieval) machines. Van den Berg [1] has made a survey of these batching heuristics. Most heuristics firstly pick a *seed order* (initial order) for a batch and afterward expand the batch with orders that have proximity to the seed order as long as the vehicle capacity is not exceeded.

Elsayed and Unal [6] developed timesaving batching heuristics. Elsayed and Stern [5] considered various proximity measures for batching heuristics, and they found that none of them produces consistently superior results through experimentations. In a comparative study of batching heuristics, Pan and Liu [10] recommend a batching heuristic developed by Hwang and Lee [9]. Their algorithm clusters orders according to proximity, which is measured by attribute vectors. Additionally, Elsayed and Lee [11] and Elsayed et al. [12] developed batching heuristics that minimize the earliness and tardiness for an order picking operation with man aboard S/R machines.

Previous studies of order batching focused on developing heuristics for assigning a small amount of orders (e.g., 5–30 orders) to batches [3]. Of these previous studies, Gibson and Sharp [7] and Rosenwein [3] presented heuristics that obtain results for a more realistic warehouse environment. Rosenwein [3] compared various approximated distance metrics that provide a quantitative basis for assigning orders to batches. However, Rosenwein [3] went along with Gibson and Sharp [7] in labeling the approximate location of each product item. They assumed that each item's location in a warehouse is its aisle index, but did not consider an item's specific location within an aisle. Additionally, most of the previous batching methods focused on the order-clustering problem in warehouses with a single-aisle layout. These batching methods were rarely applied for the case of parallel-aisle situations [13]. Examples are Elsayed [4], Elsayed and Stern [5], Elsayed and Unal [6], Hwang et al. [8], Hwang and Lee [9] and Pan and Liu [10]. The order batching methods developed and compared by Rosenwein [3] and Gibson and Sharp [7] are most applicable in the large warehouses with a parallel-aisle layout.

The optimal solutions of order batching problems are very difficult and time consuming to obtain because the travel distance or travel time of assigning a specific order to a tour is dependent on the other orders that are assigned to the tour [3]. The optimization-based approach would seem to be impractical since the formulation of the mathematical model for order batching was very complicated. The reported results for optimization-based batching methods were also limited to problems with a small set of orders. Vinod [14] presented integer programs for addressing the batching problem. One illustrative example was given in which seven batches are grouped from only 14 orders. Kusiak et al. [15] presented the batching results of eight orders by using an integer quadratic program. Armstrong et al. [16] studied the order batching problem in a semi-automated order picking system. A mixed-integer program was formulated with the objective to minimize processing time, and it was solved by the Bender's decomposition method. In the literature, only a very limited amount of research developed the optimization-based batching procedures.

In this paper, a clustering model based on 0–1 integer programming is formulated to maximize the associations between orders within each batch. The associations of customer demand can be found from the order database. When arranging order batches in warehousing, data mining can be employed to beneficially discover the association information from the order database. In data mining, association rules [17,18] are descriptive patterns of the form $\mathbf{X} \Rightarrow \mathbf{Y}$, where $\mathbf{X}$ and $\mathbf{Y}$ are sets of objects. The well-known application of the association rule is the *market basket analysis*, in which the market basket represents the set of products purchased by a customer in a single store visit. In the problem of order batching studied herein, association rules can be adopted to identify the connections between orders, which reflect the demand relationships of customers. The inclusion of one product in an order (Order $O_i$) when another order (Order $O_j$) includes the product represents an association rule ($O_i \rightarrow O_j$) for demand relationships. The association information offers a quantitative basis for order batching, and it is flexible in incorporating diverse order data into the batching procedure.

By using association rule mining, the associations between orders are determined in terms of *support* and *confidence*. The order associations imply the customer demand relationships, and they can be extracted from the product items contained in customer orders. These demand patterns indicate groups of product items that are frequently ordered together. The orders with similar demand patterns can be grouped into batches to minimize the picking distance in warehouses. The association function can be applied to find relationships between orders. The orders with higher associations may have more similar product items ordered simultaneously. By using the 0–1 integer programming, a clustering method which maximizes the customer demand relationships (associations between orders) is used to group orders into batches.

As mentioned above, previous studies usually used approximated distance and time metrics since it is complicated