Proceedings of the 2011 International Conference on Industrial Engineering and Operations Management
Kuala Lumpur, Malaysia, January 22 – 24, 2011

Association Rule Based Approach for Improving Operation Efficiency in a Randomized Warehouse

Hau Ling Chan, King Wah Pang
Department of Logistics and Maritime Studies
The Hong Kong Polytechnic University, Hung Hom, Hong Kong

Ka Wing Li
ExIT Solution Company, Hong Kong

Abstract

Data mining has long been used in relationship extraction from large amount of data for a wide range of applications such as consumer behavior analysis in marketing. Some research studies have also extended the usage of this concept in warehousing operations management to determine the order picking policy by batching the orders to minimize the picking distance. Yet, not many research studies have considered the application of the data mining approach on storage location assignment decision to minimize the manual effort in put-away operation that contributes significantly to the overall cost of warehousing operations. We present a data mining approach for the storage location assignment problem in a randomized warehouse using association rules extraction algorithm. Result of the preliminary experimental study shows that our proposed storage location assignment algorithm is efficient in determining the correlated products storage location that minimizes the total travel distances of both order picking and put-away operations for a randomized less-than-unit-load warehouse.

Keywords
Warehousing operations, storage location assignment problem, order picking, data mining, association rules

1. Introduction

For efficient supply chain operation, warehousing management is of vital importance due to the fact that it acts as the intermediary that connects upstream suppliers with downstream customers along the supply chain. In order to enhance the competiveness, many companies attempt to achieve high volume production and distribution while keeping minimal inventories throughout the supply chain that products are to be delivered to customers within short time. Hence, efficient warehousing management and operations can significantly reduce the travelling distance and processing time of product movement inside a warehouse so as to response to customer requests faster and achieve cost savings eventually.

Warehousing operations involve all movements of goods within warehouses or distribution centers, i.e. receiving, transfer and put-away, storing, order picking/selection, accumulation/sortation, cross-docking and shipping. Extensive research studies have been focused on the improvement of the order picking operations, which is one of the most labor-intensive activities for manual system and capital-intensive for automated system in a warehouse as presented in [1]. In [2], they presented four methods to reduce the traveling and distance for order picking in a warehouse: (1) determining good picking route; (2) zoning the warehouse; (3) assigning products to strategic storage locations; and (4) combining orders for batch picking. In order to achieve higher operation efficiency in warehousing, the most critical decision is on storage location assignment prior to the order picking decision. The decision on the storage location of the products can be defined as a storage location assignment problem (SLAP) that involves the decision on assigning the incoming product, often defined as a Stock Keeping Unit (SKU), to a particular location in a warehousing system. The objective includes minimizing the distance traveled for order picking or maximizing the space utilization of a warehouse. According to [3], they presented three important factors that affect the storage location decision, which are inventory turnover rates, demand dependence between products and storage space requirements. An efficient warehousing management planning should consider these factors together as an integrated system in order to determine the storage location of the products in a warehouse.

In a warehouse, product storage approach can be categorized as randomized, dedicated and correlated storage. Randomized storage approach assigns the items randomly to the storage locations without considering the products popularity and interdependency. The incoming products are generally stored at the first available location closest to the I/O entry point. In [4], they showed that the randomized storage method results in better space utilization of the warehouse at the expense of increased travel distance for order picking operations. For dedicated storage approach, each storage location is pre-allocated to a specific product and all locations are reserved for particular products, even when some products are out of stock, other products will not be allocated to these spaces. According to [5-7], the researchers commented that dedicated storage approach gives better result in terms of order picking travel distance when compared with randomized storage approach for high space utilization warehouses. In the last two decades, correlated storage assignment problems have been studied extensively. For correlated storage approach, the products that are often requested simultaneously by customers are stored closely with each other [8-11]. In [8], they proposed a 4-step procedure for the determination of the storage location assignments of correlated products to reduce the traveling in order picking operations. Other studies on correlated product clustering for storage location assignment can be found in [3, 12, 13]. Yet, they have not explored the product correlation for storage location assignment that considers the manual effort on put-away execution, which also contributes significantly to the overall cost of warehousing operations. The reason is no matter processing order picking or replenishing the products in a warehouse, assuming the amount of items inflow equal to the amount of items outflow under equilibrium, the amount of items need to be handled are the same. Therefore, assigning product storage locations for the execution of order picking operation and put-away operation during replenishment should be considered simultaneously. These three decisions and operations are highly interdependent and affect the overall warehousing operation efficiency significantly.

Nowadays, the development of new information technology is growing rapidly for applications in supply chain management, i.e., enterprise resource planning (ERP) system, radio frequency identification (RFID) system, and warehousing management system (WMS) are widely used in companies for improving the operation efficiency. Besides, recording and gathering transaction data of customer orders is more timely and cost efficient than before. Knowing the customer order patterns from transaction history is crucial not only for increasing the sales volume in retail industry, but also for the facilitation of better decision making in warehousing management. Through capture the order transactions by data mining technique, companies can analyze the interdependence of the product selling pattern easily and effectively. One of the data mining techniques is the association rule extraction that aims at discovering those interesting association relations among the items from a large dataset. In [14], the researchers presented an association rule mining algorithm, named as Apriori algorithm, to determine the itemsets that appear frequent enough, and then derive those association rules which are strong enough to be considered interesting. Readers can refer to [14] for the details. With the analysis of these interesting relations among the items, company can use the information for different functions, such as marketing planning, warehousing management, inventory planning, etc. Regarding the applications of association rule for warehousing, an association rule based mining order batching heuristic is presented in [15] to reduce the order picking travel distance. Some researchers have further explored the application of association between orders and formulated the order clustering problem as an 0-1 integer programming model that solved by heuristic approaches [9, 10, 16]. These studies adopted the Apriori algorithm for association rule mining proposed in [14]. To the best of our knowledge, none of the above studies has considered the distance traveled in put-away operations during replenishment of products when the existing stock level in the warehouse is not enough to satisfy future customer orders. Therefore, in this paper, we present a decision support system that determines the storage location of products with correlated demand based on association rule approach that minimizes the total distance travel on both order-picking and product replenishment operations in a randomized warehouse. The paper is organized as follows: In Section 2, we present the proposed storage location assignment approach based on association rule mining. In Section 3, design of the experimental study is presented with a preliminary experimental study that compares the performance between our proposed method and two other storage location assignment approaches. Finally, the paper is summarized with some concluding remarks and future research direction in Section 4.

2. Methodology

In this study, the objectives of the proposed storage location assignment algorithm are to determine the association between products and recognize the popularity of each product based on order history. After that we determine the strategic storage location for the replenishment products such that they are placed closely with the correlated items in order to minimize the total travel distance for both order picking and put away operations.

2.1. Association rule mining

We follow the formulation of the Apriori algorithm developed in [13] for the association rule mining. Given a set of items defined as $\boldsymbol{I} = \{I_i, i = 1, 2, \ldots, m\}$. Let $\boldsymbol{D}$ be a set of transactions, where each transaction $\boldsymbol{T}$ is a set of items such that $\boldsymbol{T} \subseteq \boldsymbol{I}$. Associated with each transaction $\boldsymbol{T}$ is a unique identifier, called T_{ID}. We say that a transaction $\boldsymbol{T}$ contains $\boldsymbol{X}$, a set of some items in $\boldsymbol{I}$, if $\boldsymbol{X} \subseteq \boldsymbol{T}$. An association rule is an implication in the form of $\boldsymbol{X} \Rightarrow \boldsymbol{Y}$, where $\boldsymbol{X} \subseteq \boldsymbol{I}$, $\boldsymbol{Y} \subseteq \boldsymbol{I}$, and $\boldsymbol{X} \cap \boldsymbol{Y} = \emptyset$. Such rule implies that if customers buy itemset $\boldsymbol{X}$, they will also buy itemset $\boldsymbol{Y}$ simultaneously. Given a set of transaction $\boldsymbol{D}$, the mining of association rules is to generate all association rules that have support and confidence level greater than the user-specified thresholds. The support of an association rule $\boldsymbol{X} \Rightarrow \boldsymbol{Y}$ is the percentage of transactions in the database that contains $\mathbf{X} \cup \mathbf{Y}$ which means the rule has support s in the transaction set $\boldsymbol{D}$ if $s\%$ of transactions in $\boldsymbol{D}$ contains $\mathbf{X} \cup \mathbf{Y}$. For the confidence of an association rule $\boldsymbol{X} \Rightarrow \boldsymbol{Y}$, it is the ratio of the number of transactions that contains $\mathbf{X} \cup \mathbf{Y}$ to the number of transactions that contains $\boldsymbol{X}$. That means the rule $\boldsymbol{X} \Rightarrow \boldsymbol{Y}$ holds in the transaction set $\boldsymbol{D}$ with confidence c, if $c\%$ of transactions in $\boldsymbol{D}$ that contains $\boldsymbol{X}$ also contains $\boldsymbol{Y}$. Both support and confidence are jointly taken as measures of association between any set of items and reflect how often the association occurs and how strong is the rule in the dataset respectively. For the mining of association rules between products, we adopt the Apriori algorithm proposed in [14] by setting the support and confidence level threshold so that strong association rules are determined.

2.2. Storage location assignment

To formulate the storage location assignment problem, let $\boldsymbol{P} = \{1, 2, \ldots, |\boldsymbol{P}|\}$ be the set of items to be put away to the warehouse. We have a set of locations available in the warehouse for storing new items, defined as $\mathrm{W} = \{1, 2, \ldots, |\boldsymbol{W}|\}$. We further denote the location of the I/O point of the warehouse as location 0. For an item i stores next to another item j, we define an association strength as A_{ij} that measure the relationship between item i and item j based on whether their relationship is defined as strong relation extracted by the association rule mining algorithm defined above. In the determination of the fitness of a storage location for a particular item, we denote F_{ik} as the fitnress measure if item i is allocated to the location k. The value of the F_{ik} is affected by the following factors: First, the distance between the location k from the I/O entry point, location 0 defined as d_{0k}. Second, items are suggested to be allocated close to the other items with strong association. We measure the strength of the association of item i with the items currently stored near to the location k, defined as a summation of the association strength A_{ij} for all items j that stored at locations defined as neighbors of location k (note: $A_{ij} \neq A_{ji}$ as association rule $i \Rightarrow j$ may not equal to the association rule $j \Rightarrow i$, which means the association of orders that contain i always contains j may not equal to the association of order that contains j also contains i. In such case, $i \Rightarrow j$ can be a strong rule while $j \Rightarrow i$ is not). In order to define the neighbours of a location k, we find all locations that are within a predetermined distance from the location. Finally, we also consider the importance of product i based on ABC Classification, which defined as w_i in $[0..1]$. The value of w_i can be determined by the transaction history based on the concept of ABC classification of inventory. In inventory management, company items can be classified as Class A, B and C according to their popularity and contribution to the total company sale volume. According to statistical analysis of inventory contribution to company sales in the literatures, items in general follow "80-20" rule that means about 20% of the SKUs account for about 80% of the sales volume and these items are defined as Class A items. Another 30% of the SKUs account for about 15% of the sales volume and the items are defined as Class B items. The remaining 50% of the SKUs are defined as Class C items which contribute only 5% of the volume. Some variations of the "80-20" rule are "70-20" and "60-20" rules that depend on the nature of company products and the product market characteristics. The purpose of popularity classification is to identify the most frequently ordered products and allocate them in strategic locations so that the distance traveled for both put-away and picking of these products is reduced.

To formulate the fitness measure in assigning an item i to a location k, we define two values α_A and α_B as the important weightings assigned to the association strength and the weighted distance respectively. The fitness can be measured by the following simple formula:

$$F_{ik} = \alpha_A \sum_{j \neq i} A_{ij} - \alpha_B \frac{d_{0k}}{w_i} \tag{1}$$

where

j is the location defined as neighbour of location k
α_A and α_B are the weightings assigned to the strength of association and the weighted distance
$\frac{d_{0k}}{w_i}$ is the weighted distance if product i is assigned to be stored at location k.

The value of the weighted distance component $\frac{d_{0k}}{w_i}$ in (1) is proposed to allocate the popular products, defined as Class A items with larger value of w_i, to a location closer to the I/O entry point than Class C items. Even for a location with the same travel distance from I/O entry point, the degree of fitness for a Class A item should be higher than that of a Class C item. A location k with larger value of F_{ik} is considered as more suitable to store item i. Based on this fitness measure, we determine the storage locations of all products that needed to be stored in the warehouse.

2.3. Routing for put away and order picking operations
No matter executing order picking operation or put away operation functions, the warehouse worker requires to determine the routing inside the warehouse that travels from and returns to the I/O entry point for the picking or put away operation of the products. For the put-away operations to replenish products in a warehouse, we simply based on the nearest neighbor search to construct the route with shortest travel distance which goes through all the assigned storage locations of the products. According to the fitness measure defined in (1), an item i will be allocated to the available location nearest to the I/O entry point and allocated near to its correlated products. The logic of this storage assignment is beneficial to order picking operations as these correlated products which are frequently requested together in the same customer order are stored nearby, thus, reducing the total distance traveled to pick all these requested products for a customer order. For the picking of a customer purchase order, we propose to identify the least popular product in this picking order based on the ABC classification first, i.e., product Z, and construct the route towards the I/O entry point that visits the locations of all items requested by this order. It is because the amount of product Z stored in the warehouse should be the least. Hence, the number of possible routes to be generated for consideration is lesser compares with the routes constructed by starting with the most popular product. For product Z that currently exists in the warehouse, we construct the route that visits the locations of all products requested in this customer order by the same nearest neighbor searching algorithm. Then, we identify all routes constructed and pick the one which gives the shortest total travel distance as the final order picking route.

3. Experimental study and computational results
In order to evaluate the performance of our proposed storage location assignment algorithm, we have conducted extensive computation experiments to simulate the warehouse operations of real life applications, and analyzed the total distance traveled for order picking and put away operation in a randomized warehouse

The configuration of our designed warehouse has several characteristics. First, each picking tour begins and ends at the I/O entry point located at the front end of the leftmost aisle in the warehouse. Second, the shelves in the warehouse are two-sided that products can be stored on both sides with single-block only. In addition, each item occupies exactly one storage location and all storage locations are with the same size. To start up the preliminary experiment study, we have built up a warehouse with dimensions 30 shelve columns and each shelve has 30 storage space on each side , i.e. maximum storage capacity is 1800 SKUs .We generated 1000 synthetic customer orders as order history with reference to [14] for the association rule extraction between products that uses the Apriori algorithm. In this initialization stage, the popularity of products is determined by “60-20” rule of the ABC classification, so that 20% of the SKUs account for 60% of the company sales volume. Products are allocated to the storage location according to their popularity and correlation with other products to fill up the warehouse. After that, we have generated another 300 synthetic customer orders as the incoming customer purchase orders and started the simulation with picking the requested products to fulfill the orders. The order picking operation is based on discrete order picking with single-command cycle policy that either order picking or put away operation is carried out in a trip. In case the amount of stocks available in the warehouse cannot fulfill the incoming customer order, a replenishment order is placed. The reorder quantity for replenishment is determined by the Economic Order Quantity (EOQ) model due to its simplicity in formulation. We have assumed that there is no lead time for placing each replenishment order, i.e. the required products can be replenished immediately. Products with higher popularity and contribution to the total company sale volume will be allocated to the warehouse first so that they are stored nearest to the I/O entry point which can reduce the total distance travel for both order picking and put away operations. Hence, the storage location assignment algorithm is triggered to determine the suitable locations to store these replenished items.

The parameters setting for the simulation is illustrated in Table 1. We have assumed that there are 19 different kinds of product defined as Stock-Keeping Unit (SKU), with 1,000 customer orders as order history for association rule extraction, and each order has an average of 30 units which follow normal distribution. In order to test the reliability of our proposed approach, five random samples with the above settings are generated for the performance comparison of our proposed algorithm with other two classical storage location approach, namely purely random and purely dedicated methods. In those 5 samples, the number of products in a customer order follows normal distribution with average of 30 units and average standard deviation of 5.417 units. By combining different settings of the parameters, we can simulate different scenarios for the analysis of the potential cost saving attainable by using the proposed storage location assignment algorithm.

A flowchart that illustrates the simulation of the experimental study is presented in Figure 1 below. To evaluate the effectiveness of the proposed SLAP method that based on the association rule mining approach, we compare the performance of our proposed approach with the purely dedicated storage and purely randomized storage approaches based on the total travel distance, distance travel for order picking operation, distance travel for put away operation in replenishment as well as the computation time. In this study, only horizontal travel distance is considered for the performance measure. For the vertical movement of up and down the shelves on different levels for picking or put away an item, we can easily extend the model to cover this scenario in practice. The experimental results of each storage approach are compared as presented in Table 2. The result shows that the distance traveled for put away operation contributes significantly to the overall operation cost of warehousing operation, it contributes about 70~80% of the total travel distance. In terms of the order picking distance measure, our proposed approach can save 21.19% and 39.24% when compared with the random approach and dedicated approach respectively. Thus, efficient storage location assignment management is critical to reduce order picking operations. Besides, the total distance travelled by using our proposed approach outperforms the randomized and dedicated storage approaches by 2.77% and 16.67% respectively, even though the put away distance for our proposed approach is longer than the other two approaches. This situation is well expected because we intentionally decide to store a product that is not closest to the I/O point, but store it close to other related products that facilitate faster order picking operations for customer order. In terms of computation time, all three approaches can make the storage location assignment decision within reasonable time so that the proposed algorithm is suitable to implement as a real time decision support system in warehousing management.

Table 1. Parameters of the experimental study

Experiment Setting	
Size of warehouses	30 x 30 x 2
ABC Characteristics	20/60 Rule
Number of SKUs	19
Average Order Size	30 units
Number of Picking Orders	300

Table 2. Results of the experimental study

	Our Approach	Purely Random		Purely Dedicated	
Average Total Distance	105995.6	108934.2	(2.77%)	123660	(16.67%)
Average Order Picking Distance	72780	88200	(21.19%)	101340	(39.24%)
Average Put Away Distance	33215.6	20734.2	(-60.20%)	22320	(-48.82%)
Computing Time	152.6	202	(32.37%)	220.8	(44.69%)

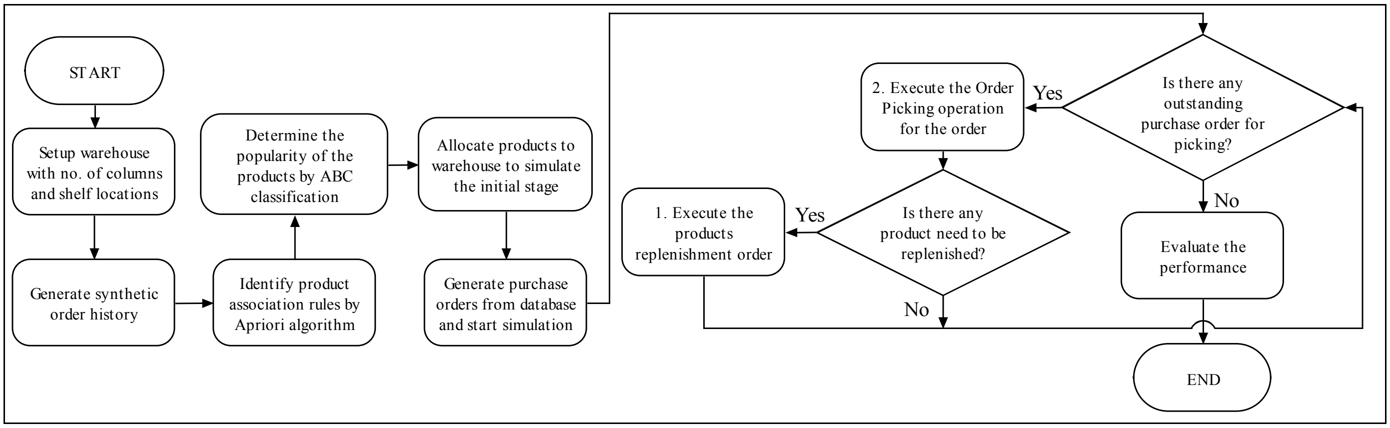

Figure 1. Flowchart of the simulation model for the experimental study

4. Conclusion

In this paper, we present a data mining technique to identify the correlated products and determine the storage location decision of the less-than-unit-load items in a randomized warehouse. The logic of this storage approach is not only beneficial to order picking activity, but also put away operations as those correlated products which are frequently requested together in the same order are stored nearby, thus, minimizing the manual efforts in the warehousing operations by reducing the total distance traveled.

The performance and effectiveness of the proposed methodology are examined by measuring the total travel distance for both order-picking and put-away operations and comparing the result with randomized and dedicated storage approach. The preliminary results show that our proposed storage approach is effective to improve the overall warehousing operation efficiency when compares with the randomized and dedicated storage approaches. However, our proposed algorithm has some limitations that have to be considered for future extension. First, a mathematical model with an objective function that considers both put-away and order picking travel distance will be formulated for performance comparison by using mathematical programme optimizers, such as ILOG CPLEX. Second, the routing decision to pick up the products after identified all the product locations to be visited is the same as a Traveling Salesmen Problem (TSP) presented in [17], which the decision maker has to determine the routing for a single vehicle to visit each assigned customer exactly once, with shortest possible travel distance. With the TSP formulation of the routing problem and being solved by corresponding algorithms, solutions with shorter total travel distance could be obtained, which further enhance the operation efficiency of the warehouses. Finally, an effective updating mechanism can also be incorporated to the decision support system in order to capture the trend of market changes on customer preferences. This extension can help enhance the system to dynamically determine the storage location of the replenishment items that minimizes the travel distances, as some items are becoming more popular while the sale volume of some items are dropping towards the end of their product life cycles.

References

1. Koster, R., Le-Duc, T., and Roodbergen, K.J., 2007, "Design and control of warehouse order picking: A literature review," European Journal of Operational Research, 182, 481-501.
2. Roodbergen, K.J., and De Koster, R., 2001, "Routing order-pickers in a warehouse with a middle aisle," European Journal of Operational Research, 133(1), 32–43.
3. Kim, K.H., 1993, "A joint determination of storage locations and space requirements for correlated items in a miniload automated storage-retrieval system," International Journal of Production Research, 31, 2649-2659.
4. Choe, K., and Sharp, G.P., 1991, "Small parts order picking: design and operation," Material Handling Research Center Technical Report. Georgia Tech. August.
5. Linn, R.J., and Wysk, R.A., 1987, "An analysis of control strategies for automated storage and retrieval systems," INFOR, 25, 66-83.
6. Malmborg, C.J., 1996, "Storage assignment policy tradeoffs," International Journal of Production Research, 34, 363–378.
7. Gu, J., 2005, "The forward reserve warehouse sizing and dimensioning problem," Ph.D. Dissertation, Georgia Institute of Technology, Atlanta, GA.

8. Frazelle, E.A., and Sharp, G.P., 1989, "Correlated assignment strategy can improve any order picking operation," Industrial Engineering, 21(4), 33-37.
9. Chen, M.C., and Wu, H.P., 2005, "An association-based clustering approach to order batching considering customer demand patterns," Omega: International Journal of Management Science, 33(4), 333-343.
10. Chiang D.M.H., Lin, C.P., and Chen, M.C., 2008, "A product allocation approach based on association rule mining," In Proceedings of the 9th Asia Pacific Industrial Engineering and Management Systems Conference, 2418-2423.
11. Xiao, J., and Zheng, L., 2009, "A correlated storage location assignment problem in a singleblock- multi-aisles warehouse considering BOM information," International Journal of Production Research, 48(5), 1321-1338.
12. Liu, C.M., 1999, "Clustering techniques for stock location and order-picking in a distribution center," Computers and Operations Research, 26, 989–1002.
13. Bindi, F., Manzini, R., Pareschi, A., and Regattieri, A., 2007, "Similarity coefficients and clustering techniques for the correlated assignment problem in warehousing systems," In Proceedings of 19th International Conference on Production Research.
14. Agrawal, R., Imielinski, T., and Swami, A., 1993, "Mining association rules between sets of items in large databases," ACM SIGMOD Conference, Washington DC, USA, 254-259.
15. Chen, M.C., Huang, C.L., Wu, H.P., Hsu, M.F., and Hsu, F.H., 2005, "A data mining technique to grouping customer orders in warehouse management system," In Proceedings of WSTST'05, California, USA.
16. Chen, M.C., Huang, C.L., Chen, K.Y., and Wu, H.P., 2005, "Aggregation of orders in distribution centers using data mining," Expert Systems with Applications, 28(3), 453–460.
17. Hwang H., Baek, W.J., Lee, M.-K., 1988, "Clustering algorithms for order picking in an automated storage and retrieval system," International Journal of Production Research, 26(2), 189 – 201.