Volume 31 (2), pp. 61–76
http://orion.journals.ac.za

ORiON
ISSN 0259–191X (print)
ISSN 2224–0004 (online)

SKU assignment to unidirectional picking lines using correlations

J Matthews* SE Visagie†

Received: 21 January 2015; Revised: 07 April 2015; Accepted: 15 July 2015

Abstract

A real life order picking system consisting of a set of unidirectional picking lines is investigated. Batches of stock keeping units (SKUs) are processed in waves defined as a set of SKUs and their corresponding store requirements. Each wave is processed independently on one of the parallel picking lines as pickers walk in a clockwise direction picking stock. Once all the orders for a wave are completed a new mutually exclusive set of SKUs are brought to the picking line for a new wave. SKUs which differ only in size classification, for example small, medium and large shirts, are grouped together into distributions (DBNs) and must be picked in the same wave. The assignment of DBNs to available picking lines for a single day of picking is considered in this paper. Different assignments of DBNs to picking lines are evaluated using three measures, namely total walking distance, the number of resulting small cartons and work balance. Several approaches to assign DBNs to picking lines have been investigated in literature. All of these approaches seek to minimise walking distance only and include mathematical formulations and greedy heuristics. Four different correlation measure are introduced in this paper to reduce the number of small cartons produced and reduce walking distance simultaneously. These correlation measures are used in a greedy insertion algorithm. The correlation measures were compared to historical assignments as well as a greedy approach which is known to address walking distances effectively. Using correlation measures to assign DBNs to picking lines reduces the total walking distance of pickers by 20% compared to the historical assignments. This is similar to the greedy approach which only considers walking distance as an objective, however, using correlations reduced the number of small cartons produced by the greedy approach.

Key words: SKU assignment, order picking, assignment problems, combinatorial optimisation.

1 Introduction

Warehouses form a central part of supply chains. The role of warehouses is typically to match supply with demand and to consolidate product from multiple suppliers [2]. To

*Corresponding author: Department of Logistics, University of Stellenbosch, South Africa, email: jmatthews@sun.ac.za

†Department of Logistics, University of Stellenbosch, South Africa

http://dx.doi.org/10.5784/31-2-531

play this role, stock must be stored, moved and picked in the warehouse using one or more of a variety of different layouts, mechanisms, picking systems and storage equipment depending on product and market characteristics. One of the essential parts of warehouse management is the placing of stock into locations that improve stock movement and picking efficiency.

The warehouse slotting problem is described by Kim & Smith [8] as determining an assignment of stock keeping units (SKUs) to picking slots[1] to support carton picking. Ideally SKUs which are usually placed in the same carton should be placed as near to each other as possible. Warehouses typically do not re-slot in the short term because in most cases long term SKU correlations are used as a desirability measure to slot SKUs close to each other. Furthermore the time and cost involved to re-slot is often too high.

Following on a study by Matthews & Visagie [14] a distribution centre (DC) owned by Pep Stores Ltd[2] (PEP) is considered [16]. A major influence on the order pick system in use at their DCs is the centralised stock management of PEP. Stock levels for each store are managed collectively and stock outflows are planned centrally at SKU level creating a push system. During an outflow for a SKU all stores requiring that SKU are stocked with the SKU in a single pick operation. Distributions (DBNs) which consist of a set of SKUs that are of the same product but differ in size are used to achieve this. The make-up of a DBN is determined by the central planning department which sets how much of each SKU in the DBN should go to each store. These DBN pick instructions are issued (or released) to the DC. All the SKUs in the DBN are picked in the same batch once the DBN pick instructions are released.

A type of forward pick area – as described by Bartholdi & Hackman [2] – consisting of 12 picking lines is used to pick these DBNs. A picking line has 56 slots (or locations) each holding up to five pallet loads of the same SKU and is used for all piece picking in the DC. These picking lines are serviced by multiple pickers and are able to operate in parallel to each other. Figure 1 illustrates the layout of the picking lines in this forward pick area.

Picking lines operate in waves, defined as a set of SKUs and their corresponding store requirements which are to be picked in a batch. Each wave of SKUs is placed on its own picking line and all the store requirements for those SKUs are picked in a single operation. Once all the picking is completed any remaining stock is removed and a new wave of SKUs is brought to the picking line. All the SKUs associated with the same DBN are placed on the same picking line ensuring that all the SKUs in the same DBN arrive at the store at the same time. This process of populating, picking and clearing stock on a picking line may take anything from four hours to two days depending on the number and size of orders associated with, and the characteristics of, the SKUs assigned to that wave on the picking line.

Due to the varying rates at which picking lines are completed and the parallelisation of the picking line area the number of picking lines which become available for new waves during each day varies. DBNs are scheduled onto available picking lines using a first-in, first-out (FIFO) system in an attempt to ensure that all DBNs are processed within the

[1] A picking slot is a storage location which is directly accessible by pickers.

[2] Pep is the largest single brand retailer in South Africa.

desired threshold of seven days from receiving both the pick instructions from the planning department and the physical stock from the suppliers.

During the picking phase pickers walk in a clockwise direction around a picking line sequentially picking orders. Order consolidation is not performed to ensure that picking lines are managed independently from each other. Instead pickers pick directly into cartons placing the completed cartons onto a conveyor belt. New cartons as well as re-cycled cartons from suppliers are used and are available around the picking line. Cartons only hold stock from a single picking line and are closed and shipped as they come from the picking line. Moreover when pickers select a carton to hold the stock for an order they do not know what volume of stock is required for that order. Many cartons therefore have excess capacity as small volumes of stock required by stores from individual picking lines are placed in much larger cartons. These cartons are manually resized into smaller cartons to reduce volume and are undesirable as they increase the per volume handling cost throughout the DC. Completed cartons are then stored in buffer areas located in the outbound section of the DC which are emptied on a regular basis based on delivery schedules.

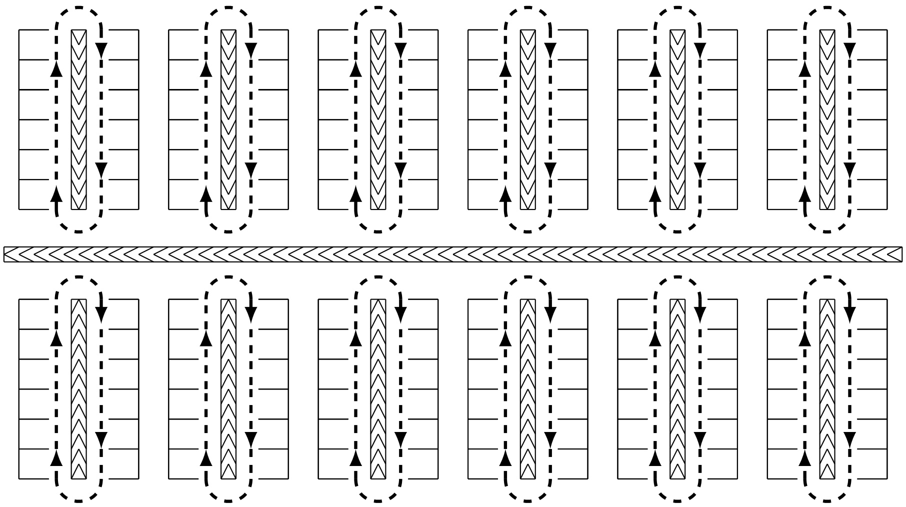

Figure 1: *A schematic representation of the layout of the 12 picking lines in the DC, six on either side of the main conveyor. The dashed lines indicate the movement of the pickers around the conveyor belts [14]*

Matthews & Visagie [14] suggested approaches to assign DBNs to picking lines while minimising the total walking distance of pickers by using the concept of a maximal SKU[3]. Significant improvements were made on the historical method by using both integer programming (IP) and heuristic approaches. They pointed out that focusing on walking distance alone resulted increased the number of small cartons produced, as many stores required a small volume of stock from certain picking lines. In addition, operational risk was increased as more picking lines required excessively large volumes of stock which increased the overall time which a picking line is occupied by a single wave of picking and might result in a need to replenish stock during a wave of picking. The focus of this paper

[3]A maximal SKU is a SKU within a set of SKUs with the highest number of stores requiring it.

is to address these two additional issues by using SKU correlations while still achieving satisfactory walking distances.

The remainder of this paper is structured as follows. A discussion of related work in existing literature is given in §2. The four solution approaches using correlations are introduced in §3 with the results presented in §4. The paper is concluded in §5.

2 Literature review

Accorsi *et al.* [1] addressed both the storage allocation and storage assignment problems simultaneously. The storage allocation problem focuses on the amount of stock stored in each location, typically addressing the issues around replenishment costs, while the storage assignment problem focuses on the physical location of stock in an effort to minimise order picking costs. Three main problems were identified when optimising order picking systems with a forward pick area, namely:

1. Which SKUs should be in the forward pick area?
2. How much of each SKU should be in the forward pick area?
3. Where should each SKU be stored?

Bartholdi & Hackman [2] addressed the first problem and introduced three approaches for the second problem namely the equal space, equal time and optimal allocation strategies. Accorsi *et al.* [1] addressed the final storage assignment problem and identified three main approaches, namely the class-based, ranked-index-based and correlation storage assignment policies. The clustering storage assignment policy was further expanded as three sequential steps, namely correlation analysis, clustering and priority list determination with cluster assignment.

Accorsi *et al.* [1] also proposed a top down hierarchical procedure for overall order picking optimisation which was applied to a case study. Numerous combinations of different approaches to each problem were considered including different allocation rules, correlation measures and clustering algorithms. It was shown that considering both SKU allocation and SKU assignment simultaneously yields better overall order picking costs compared to sequentially solving each problem.

Although Accorsi *et al.* [1] proposes a generalised framework for order picking optimisation the detail of the SKU assignment approaches have often been governed by DC layouts and management practices. Chiang *et al.* [4] used data mining techniques to assign newly arriving SKUs to available slots in a rectangular DC with a S-shaped picking strategy. An association index was developed between SKUs and available locations based on association rule mining, SKU turnover rates and the distance from a location to the exit. The association rule mining used SKU correlations between already assigned and unassigned SKUs. A generalised assignment problem was formulated to assign SKUs to available locations which maximised the sum of these association indices.

Bindi *et al.* [3] also investigated storage allocation rules for a rectangular DC with parallel shelves. Two processes were identified, namely family grouping and storage allocation. A proposed similarity measure, based on SKU correlations and a stock turn coefficient (total stock movement over average stock level) was developed and compared to a Jaccard statistic. Several storage assignment rules were also tested as part of the storage allocation process. Extensive testing by means of what-if analysis for a case study showed that correlation measures significantly improve overall throughput of the DC.

Manzini [9] minimised total picking time in a rectangular warehouse by using SKU correlations to arrange SKUs. In this case, however, the warehouse had two orthogonal sets of shelves. The DC employed a picker-to-parts system and picking vehicles with a finite capacity using a composite picking strategy. Three solution approaches based on correlations were proposed, namely a clustering approach, a parallel algorithm and a sequential algorithm. It was noted that the sequential approach, which used the last assigned SKU to determine the next SKU to assign to a cluster, had a risk of generating correlated couplets of SKUs instead of maximising overall correlations.

A SKU assignment problem in a synchronised zone order picking system was investigated by Jane & Laih [7] using correlation and clustering techniques. All zones processed the same order at the same time and only once an order has been picked in all zones can a new order begin for any zone. The completion time of an order was thus seen as the longest completion over all zones and an objective was therefore defined by using correlations which balanced the workloads of each order over all zones

The structure and layout of the DC play a role in determining appropriate SKU assignment techniques as seen by the previous studies. Kim & Smith [8] investigated a carton order picking system which has many similarities to the order picking system discussed in this paper. The DC considered by Kim & Smith [8] also had re-slotting which was performed on a daily basis. Different sets of SKUs were picked on different days of the week which created vastly different SKU correlations for each wave of picking. The time and cost required to re-slot was reasonable and re-slotting was therefore performed during the night followed by a wave of picking during the next day.

The picking area considered by Kim & Smith [8] consisted of a number of single aisle zones, each with a single picker. Cartons requiring SKUs from multiple zones are conveyed between zones which removes the need for later consolidation. Pickers receive new orders[4] for their zone from a designated starting point at one end of the aisle and proceed to pick all required SKUs before placing the carton on a conveyor belt and walking back to the start to begin a new order. In this way the distance walked by a picker to complete an order in a zone is equal to twice the distance from the start to the furthest required SKU.

Kim & Smith [8] considered the SKU slotting problem while minimising the total completion time to pick a wave. Orders typically require stock from other zones and thus the total completion time of a wave was determined by the zone with the longest completion time to pick all orders. An IP formulation was developed which minimised walking distance, but it was too complex to solve. Three further heuristic approaches were therefore introduced, namely a steepest descent neighbourhood slotting heuristic, a correlated slotting heuristic and a simulated annealing slotting heuristic. These approaches used correlations to determine good pairs of SKUs to switch by reasoning that SKUs with high correlations should be near each other. Once a switch is made the total walking distance is then re-calculated.

[4]New orders for a zone includes orders picked in other zones.

It was shown that the simulated annealing approach performed best.

Although the order picking system addressed in this paper has a similar structure and reslotting methodology to the study by Kim & Smith [8] the structure of each zone/picking line is fundamentally different. Instead of the linear zone the cyclical structure of a picking line shows many similarities to a carousel system. A carousel system is an automated storage and retrieval system with a set of shelves which rotate to present stock to pickers. Hassini [5] presents an extensive literature study on carousel systems. Hassini [5] noted that correlations between SKUs have received little attention in the carousel context. It is further suggested that SKU correlations should be used when assigning SKUs to carousels as well as locations within a carousel.

There has also been some attention paid in literature to the exact order picking setup discussed in this paper with some of the different decision tiers of this order picking system being addressed. These decisions include the sequence in which orders are passed to pickers, the arrangement of SKUs in a picking line and the assignment of DBNs to picking lines. Matthews & Visagie [11] considered the problem of sequencing orders for pickers in a unidirectional picking line, with fixed SKU slotting, while minimising the total distance travelled. The concept of a maximal cut was introduced and an IP formulation was developed which generated a solution within one pick cycle of a lower bound.

Matthews & Visagie [13] considered the SKU arrangement on a single picking line. SKUs which have already been assigned to the picking line are arranged while minimising the total walking distance using the maximal cut approach as suggested by Matthews & Visagie [11]. An IP formulation was presented which was shown not to be solvable in a realistic time frame for problem instances with more than 15 SKUs. Matthews & Visagie [13] also tested several heuristic methods including an organ pipe and a greedy approach, both of which are optimal for some carousel systems which have many similarities to unidirectional picking lines. These heuristics were tested against historical arrangements as well as a set of random solutions. It was found that the scope for optimisation when arranging SKUs on a single picking line was minimal. A lower bound for the number of cycles traversed was also identified by considering the SKU with the maximum number of stores requiring it, called a maximal SKU. At least one cycle would need to be traversed for each store requiring the maximal SKU (*i.e.* the size of the maximal SKU) which generated this lower bound.

Matthews & Visagie [14] generalised their study considering the assignment of DBNs over multiple picking lines. It was reasoned that the maximal SKU measure of a lower bound should be correlated with the actual number of cycles traversed. Matthews & Visagie [14] therefore minimised the sum of the sizes of the maximal SKUs for each picking line to reduce total walking distance. An IP formulation with this new objective was developed which was not solvable for problem instances with more than four picking lines. A further relaxation of this formulation was therefore developed which rounded the size of the maximal SKUs in an effort to reduce the computational effort of proving an optimal solution. This relaxation showed faster computational times but was still not solvable in a realistic time frame (within 10 minutes) for problem instances with more than six picking lines.

A greedy insertion approach based on the algorithm by Martello & Toth [10] for multiple knapsacks was therefore developed. It assigns DBNs to picking lines in a greedy fashion based on the size of the maximal SKUs and the available space in the picking line. Although this approach yielded good results in a short computational time, in some cases a feasible solution could not be found because all the DBNs were not assigned to a picking line. A phased insertion approach was thus developed which held certain small DBNs back for a second round of insertion to ensure feasibility. The results were similar in terms of walking distance and in all cases a feasible solution could be found.

For all the studies mentioned above, the performance of solution approaches are only compared based on a single measure, namely picking speed and efficiency. It is clear from the studies discussed in this section that for many DC configurations using SKU correlations to assign SKUs to slots improves the speed of the order pick operation. In most cases only correlations between adjacent SKUs and not a broader neighbourhood are included in the objective function. No correlation approaches have yet been adapted for or applied to the order picking system presented in this paper.

3 Solution approaches

When applying the top down hierarchical procedure by Accorsi *et al.* [1] to the order picking system at PEP only the storage assignment phase needs to be applied as all piece picking must be processed on a picking line. Replenishment while picking is in progress is not present at PEP because all required stock for a pick wave is stored in the picking line before a wave of picking begins.

The storage assignment phase may further be simplified by only considering a clustering based approach. An index based approach for the storage assignment problem, which typically addresses restock travel distances, is not appropriate for the order picking system at PEP as all picking lines may be viewed as equidistant from the reserve area and restocks are rare. Only the correlation and clustering steps are required when assigning SKUs to picking lines as the number of SKUs assigned to each cluster should match the number of available locations for an available picking line. This removes the need for the priority list determination and cluster assignment.

Using the simplified top down hierarchical procedure by Accorsi *et al.* [1] as well as the study by Matthews & Visagie [14] the assignment of SKUs may be seen as two phased. Firstly each DBN (d) in the set of DBNs ($\mathcal{D}$) needs to be assigned to a picking line. Once a set of DBNs has been assigned to a picking line the SKUs associated with those DBNs should be arranged by assigning them to individual locations in a SKU arrangement phase. The walking distance of the pickers can be calculated only once each SKU has been assigned to a location. The first phase is illustrated on the left and the second phase on the right hand side of Figure 2.

Matthews & Visagie [13] investigated approaches for the SKU arrangement phase for a single picking line and made use of the maximal cut approach described in Matthews & Visagie [11] to evaluate the resulting walking distances of the different arrangements. They showed that savings were minimal during this phase and that the problem was too complex to be solved exactly suggesting that the two assignment phases need to be handled

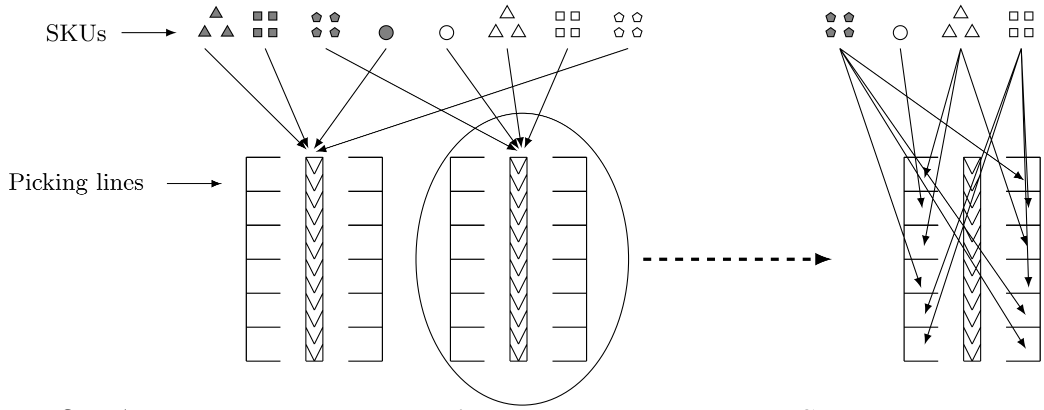

Figure 2: *A schematic representation of the slotting phases in the DC. Each shape represents a SKU and clusters of the same shape with the same shading represent DBNs.*

independently. The SKU arrangement phase will therefore be solved separately using the greedy approach by Matthews & Visagie [13] as it is fast, easy to implement and is known to achieve good results. The focus therefore moves specifically to the assignment of DBNs to a set of picking lines ($\mathcal{L}$).

In the picking line assignment phase each DBN requires a number of locations ($|d|$) in a picking lines. Each picking line (l) in the set of available picking lines ($\mathcal{L}$) has a number of available locations ($|l|$). The current approach used to assign DBNs is to spread work, measured by volume of stock, evenly between available picking lines which does not take into account walking distances of pickers. An underlying principle of each approach by Kim & Smith [8] was to interchange SKUs between slots followed by an objective function re-evaluation. Several characteristics of the problem considered here points against the use of SKU interchanges when considering unidirectional picking lines. DBNs which vary in size (number of required locations) would need to be interchanged in their entirety between picking lines. This creates more complexity as either only DBNs of the same size can be interchanged or sets of DBNs with the same number of SKUs collectively need to be interchanged. A phased greedy insertion approach was therefore introduced by Matthews & Visagie [14] to insert DBNs into available picking lines. Here DBNs are ranked according to some desirability measure and inserted sequentially into available picking lines. If a feasible solution is not found the DBNs are segmented into two subsets according to their size (number of SKUs and number of stores). These different subsets are then inserted into the available picking lines in two phases. These subsets iteratively change in size until a feasible solution is found.

Matthews & Visagie [14] used a maximal SKU measure with the phased greedy insertion algorithm (GP) which minimised the sum of the sizes of the maximal SKUs. For each DBN the SKU which has the highest number of stores requiring it (referred to as a maximal SKU) is considered and DBNs are ranked according to the size of this maximal SKU denoted as $\lceil d \rceil$. It is, however, proposed that correlations between DBNs should be considered to reduce the number of small cartons produced while still maintaining good walking distances. A correlation measure is therefore introduced as $\mathcal{B}_a \cap \mathcal{B}_b$, where $\mathcal{B}_a$ represents the set of stores requiring at least one SKU from DBN set $\mathcal{D}_a \subseteq \mathcal{D}$. By assigning DBNs with strong correlations in terms of this measure to the same picking line it would be expected that both the walking distance would be shorter and the size of each order in a picking line would be larger because more SKUs that have to be picked for the same stores will be grouped together in the same wave of picking.

Four possible desirability scores which use correlation measures were used to rank DBNs during the phased greedy insertion procedure. The first desirability score denoted as ADT considers the number of stores required by the candidate DBN and which requires at least one DBN already assigned to the picking line. This is achieved by merging all assigned DBNs in a picking line and considering them as a single DBN. The intersection of the set of stores requiring the candidate DBN and this new merged DBN (or correlation between the two DBNs) is then calculated. This desirability score is defined as

$$S(\mathcal{D}_l, d) = |\mathcal{B}_l \cap \mathcal{B}_d|. \tag{1}$$

The second desirability score (ADS), defined as

$$S(\mathcal{D}_l, d) = \sum_{a \in \mathcal{D}_l} |\mathcal{B}_a \cap \mathcal{B}_d|, \tag{2}$$

considers the correlations of a candidate DBN with all preassigned DBNs individually. This is achieved by calculating the sum of all the correlations between assigned DBNs and the candidate DBN. By assigning DBNs using these desirability scores picking lines should have fewer stores which only require one or two SKUs resulting in fewer small cartons being produced. In addition by increasing the number of shared stores the physical pick density (picks per store) of each store should increase which should create efficient pick cycles as pickers will be picking from more locations per cycle.

Bindi *et al.* [3] proposed a similarity measure which used both an adjusted Jaccard statistic and a stock turn coefficient, defined as the ratio between the total stock movement and average stock quantity. The nature of the order picking system considered in this paper does not, however, lend itself to the use of stock turn in a similarity measure due to the wave principle and the frequency at which picking lines are built. A third desirability score, defined as

$$S(\mathcal{D}_l, d) = \frac{\mathcal{B}_l \cap \mathcal{B}_d}{\mathcal{B}_l \cup \mathcal{B}_d}, \tag{3}$$

is based on the Jaccard statistic (JCT) and is included in the tests. Finally a desirability score (JCS)

$$S(\mathcal{D}_l, d) = \sum_{a \in \mathcal{D}_l} \frac{\mathcal{B}_a \cap \mathcal{B}_d}{\mathcal{B}_a \cup \mathcal{B}_d} \tag{4}$$

is introduced. The JCS measure is similar to the ADS measure, but scaled relative to the number of DBNs in the two subsets. Here the sum of the Jaccard statistics between all DBNs and the candidate DBN is calculated. The GP algorithm used to insert DBNs based on a desirability measure is illustrated in Algorithm 1.

Procedure 1: A partial greedy insertion of DBNs using a desirability measure.

Data: A set of picking lines $\mathcal{L}$ in descending order according to $|l|$
A set of DBNs $\mathcal{D}$
A set of pre-assigned DBNs $\mathcal{D}_l$ associated with each picking line
Result: A final set of assigned DBNs $\mathcal{D}_l$ associated with each picking line

for *Each picking line* $l \in \mathcal{L}$ **do**
  while *an unassigned DBN exists which fits into the remainging locations of picking line* l **do**
    Select the DBN with the largest desirability score, $S(\mathcal{D}_l, d)$, which fits into set $\mathcal{D}_l$
    Assign this DBN to set $\mathcal{D}_l$
  end
end

Algorithm 1: A sequential phased insertion of DBNs using a desirability measure.

Data: A set of picking lines $\mathcal{L}$
A set of DBNs $\mathcal{D}$
Result: An assignment of DBNs to picking lines

$\beta = 0$
while *an unassigned DBN exists* **do**
  Clear all assignments of DBNs
  Insert all DBNs where $|d| > 1$ or $\lceil d \rceil > \beta$ using Procedure 1
  Insert all remaining DBNs using Procedure 1
  $\beta = \max_{d/|d|>\beta} |d|$
end

4 Results

The four proposed desirability scores were tested using a phased greedy insertion approach on seven scenarios from real life historical data. Each problem instance comprised of a number of picking lines which were scheduled for the same historical day and the historical DBNs assigned to them. Each test scenario consisted of a set of these problem instances each with the same number of picking lines per day which allows for easier comparison. All the scenarios are available on-line [12]. A summary of the properties of these scenarios is given in Table 1. All testing was performed on an Intel i7 2 GHz processor with eight GB ram running the Windows 7 operating system. All mathematical formulations were solved with CPLEX 12.3 and coded in AIMMS 3.12 [6, 15].

Number of lines per day	Number of problem instances	Number of DBNs
2	61	2592
3	53	3437
4	49	4146
5	38	4109
6	32	4161
7	22	3177
8	14	2148

Table 1: *The composition of the scenarios from historical data.*

The results of all the approaches were compared to the maximal SKU phased greedy insertion approach (GP) by Matthews & Visagie [14]. The performance of the approaches are compared using three measures, namely walking distance, the number of small cartons produced and volume distribution. Table 2 illustrates the total distance walked for each scenario for each approach. It is clear that the GP approach performs the best in terms of walking distance, while the ADT approach shows the best results for approaches using correlations. All approaches using correlations have marginally longer walking distances (within 5%) compared to the GP. All the presented approaches still improve on the historical results by approximately 20%. In both cases the summed correlation measures (ADS, JCS) perform worse than their parent scores (ADT, JCT).

Scenario	His	GP	ADT	ADS	JCT	JCS
2	7168	6019	6193	6212	6191	6224
3	9289	7515	7710	7749	7712	7791
4	11532	9015	9282	9311	9289	9363
5	11259	8618	8922	8932	8923	8988
6	10706	8129	8363	8384	8371	8442
7	8412	6366	6540	6559	6551	6608
8	6011	4715	4847	4856	4852	4870

Table 2: *The total number of kilometres walked in each scenario for all the solution approaches (ADT, ADS, JCT, JCS) as well as the historical assignment (His) and GP approach by [14].*

A summary of the number of small cartons produced by each approach is given in Table 3. Small orders (that cause cartons to have excess volume capacity) have less than 0.006 m^3 volume of stock from a picking line assigned to them [17]. In terms of number of small orders the worst performing approach is the GP approach. All the proposed correlation measures show lower numbers of small cartons produced compared to the GP approach as the number of small orders is roughly halved. These approaches still perform worse compared to the historical assignments, but this is offset by the improvement in walking distance.

Scenario	His	GP	ADT	ADS	JCT	JCS
2	0.05%	0.14%	0.09%	0.09%	0.09%	0.08%
3	0.04%	0.13%	0.08%	0.08%	0.08%	0.08%
4	0.03%	0.12%	0.07%	0.07%	0.07%	0.07%
5	0.03%	0.13%	0.06%	0.07%	0.07%	0.08%
6	0.04%	0.12%	0.06%	0.07%	0.06%	0.07%
7	0.03%	0.13%	0.06%	0.07%	0.06%	0.07%
8	0.03%	0.13%	0.06%	0.06%	0.06%	0.06%

Table 3: *The proportion of total volume of stock attributed to small orders (i.e. orders with less than 0.006 m^3 of stock). For the historical assignment (His) the GP approach by [14] and the correlation approaches (ADT, ADS, JCT, JCS).*

A scatter plot between these two measures is given in Figure 3 to better visualise the trade off between walking distance and the number of small cartons produced. Each marker indicates the total walking distance, in kilometres, as well as the proportion of total picked volume attributed to small orders. The historical assignments forms a cluster of solutions with long walking distance and good number of small cartons produced while the GP approach shows many more solutions with poor number of small cartons produced and shorter walking distances. The solutions obtained using the ADT approach are clustered

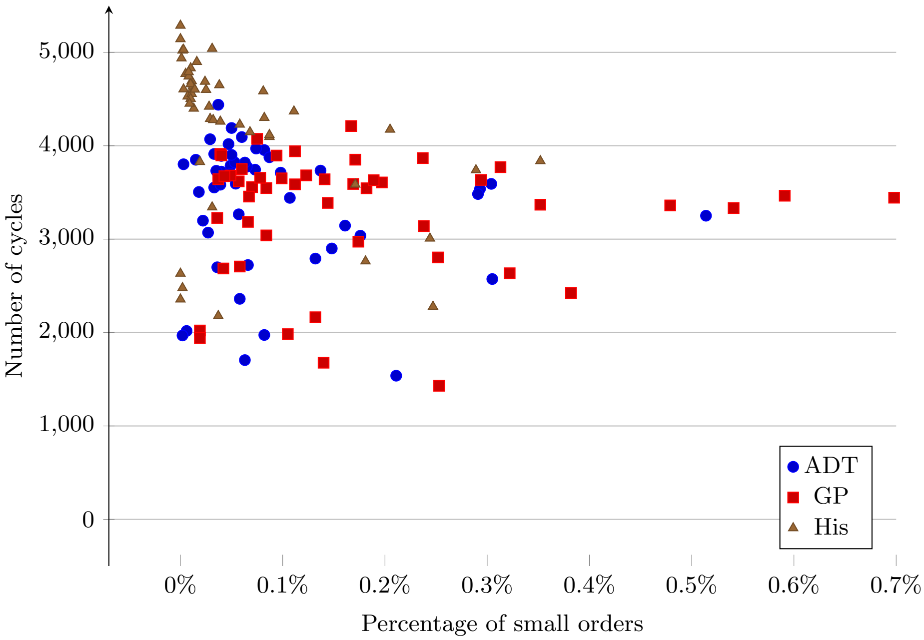

Figure 3: *A scatter plot between the number of cycles traversed and the percentage of the total picked volume attributed to small orders. Each marker represents a single problem instance with four picking lines.*

with shorter walking distances and good number of small cartons produced relative to the historical solutions.

A summary of the computational times required for each approach is given in Table 4. The use of correlations in a desirability score increases the computation times compared to the GP approach, which only considers maximal SKUs. This is attributed to the need to dynamically change the desirability score after each insertion of a DBN. The two measures which consider the sum of correlations (ADS, JCS) have shorter computational times than their parents (ADT, JCT). This is due to the ability to calculate $\mathcal{B}_a \cap \mathcal{B}_d$ and $\mathcal{B}_a \cup \mathcal{B}_d$ for each pair of DBNs only once and use this pre-calculated value for each iteration of the ADS and JCS approaches. It is also noted that the maximum computation time for problem instances with eight picking lines is high. This is due to the number of additional insertion phases required to find a feasible solution.

A comparison of the size of the picking lines in terms of volume was also performed. Figure 4 illustrates the distribution of volume over all the lines for scenarios with four and five lines per problem instance respectively. Similar results were obtained for the other problem instances. It is clear that the approaches using correlations have reduced the size of largest picking lines with respect to volume. For the scenario with five picking lines per problem instance the spread of volume over lines is aligned to that of the historical case. Similar patterns were observed for the other scenarios. Correlation measures provide the best trade offs when used to assign DBNs if all three measures are taken into account. Following all the results it is proposed that the ADT approach should be used to assign DBNs.

Algorithm	# Lines/instance	μ	σ	Max	Q_1	Median	Q_3	Min
ADT	2	1.86	0.59	3.07	2.35	1.92	1.41	0.65
	3	3.88	1.34	7.11	4.86	3.82	2.72	1.46
	4	6.49	2.23	13.11	7.61	5.96	5.31	1.98
	5	14.35	39.72	252.40	9.74	7.14	6.25	3.93
	6	19.31	50.63	296.30	13.02	9.97	8.29	5.99
	7	40.13	94.95	357.99	12.01	11.14	9.44	6.22
	8	87.58	206.04	734.06	14.36	12.81	11.45	7.08
ADS	2	0.99	0.42	1.96	1.26	0.98	0.72	0.06
	3	1.96	0.90	4.82	2.45	1.84	1.36	0.16
	4	3.06	1.63	8.32	3.78	2.88	1.90	0.26
	5	2.71	1.28	6.29	3.78	2.64	1.57	0.50
	6	3.25	1.85	7.81	4.63	2.81	1.74	0.76
	7	2.52	0.97	4.16	3.41	2.50	1.63	0.98
	8	3.13	4.35	16.02	2.60	1.52	0.89	0.61
GP	2	0.01	0.00	0.02	0.01	0.01	0.01	0.00
	3	0.03	0.09	0.69	0.02	0.01	0.01	0.00
	4	0.07	0.30	2.13	0.03	0.02	0.02	0.01
	5	0.07	0.23	1.43	0.04	0.03	0.02	0.02
	6	0.09	0.30	1.75	0.05	0.03	0.03	0.02
	7	0.32	0.71	2.53	0.05	0.05	0.03	0.02
	8	0.39	0.87	2.53	0.06	0.04	0.03	0.02
JCT	2	1.83	0.56	2.74	2.35	1.86	1.45	0.69
	3	4.58	1.52	8.09	5.58	4.68	3.59	1.86
	4	6.83	2.19	13.76	8.03	6.27	5.48	2.31
	5	9.87	2.90	17.53	11.82	8.94	7.63	4.77
	6	13.22	3.88	22.41	15.42	12.73	10.60	6.87
	7	14.00	4.74	25.08	15.15	13.21	12.17	7.53
	8	60.14	42.64	204.77	56.21	49.38	45.03	28.89
JCS	2	0.88	0.38	1.73	1.14	0.87	0.64	0.06
	3	1.92	0.87	4.59	2.40	1.80	1.34	0.15
	4	7.47	6.99	28.96	12.00	3.87	2.53	0.24
	5	2.84	1.26	5.01	4.07	2.83	1.68	0.51
	6	3.26	1.68	6.52	4.58	2.81	1.90	0.83
	7	2.37	0.97	4.15	3.12	2.40	1.52	0.90
	8	98.75	101.89	296.06	171.08	72.19	9.39	0.61

Table 4: *A comparison of computational times in seconds between the different solution approaches for each scenario. Both the average times (μ) and the standard deviation (σ) thereof are presented. Standard measures of spread are also presented with the maximum time (Max) the 25^{th} percentile (Q_1), the 50^{th} percentile (Median), the 75^{th} percentile (Q_3) and the minimum time (Min) given.*

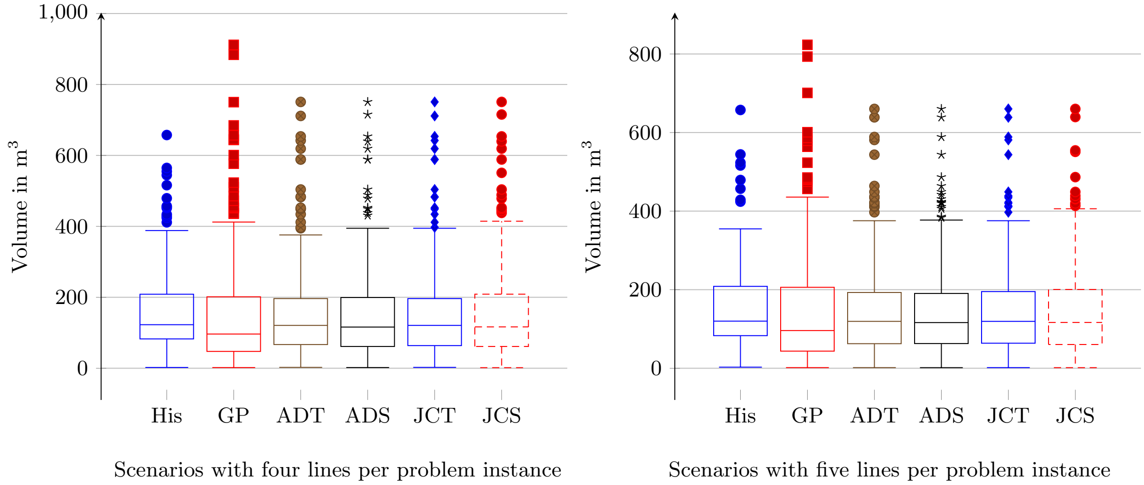

Figure 4: *A graphical box-plot representation of the distribution of the total volume for each line after scheduling scenarios with four and five lines per problem instance. The median (50^{th} percentile), Q1 (25^{th} percentile) and Q3 (75^{th} percentile) are represented as the horizontal lines in the closed box. The individually plotted coordinates are associated with the outliers which are either 1.5 times the inter quartile range (IQR) smaller than Q1 or larger than Q3. The whisker lines indicate the minimum and maximum number of cycles for non-outliers.*

5 Conclusion

A real life order picking system where re-slotting is performed on a daily basis as implemented by PEP was investigated. This investigation follows on a study by Matthews & Visagie [14]. The order picking system consisted of unidirectional picking lines in a forward pick area where all the piece picking is processed. SKUs, which are grouped together into DBNs by PEP, are batched into waves and processed in a single operation on a picking line. The number of picking lines which became available for the assignment of DBNs each day vary as the time required to stock, pick and clear picking lines varies. The assignment of DBNs to available picking lines forms the focus of this study. Assignments are evaluated in terms of the distance walked to pick all orders, the number of small cartons produced as well as the spread of volume over picking lines.

Matthews & Visagie [14] used a phased greedy insertion technique to minimise the sum of the maximal SKUs in an effort to minimise the walking distance of pickers. It was shown that this objective had negative effects on other operational areas such as the number of small cartons produced and volume distribution. Many approaches in literature use correlations to assign SKUs to locations although the main objective was still to reduce total picking time. It is therefore proposed to use correlations between DBNs as a measure to assign DBNs to picking lines to reduce the number of small cartons produced while still maintaining acceptable walking distances.

Four desirability scores were tested and compared to the historical case as well as the maximal SKU approach (GP) by Matthews & Visagie [14]. The first two scores (ADT and ADS) considered the total number of stores sharing SKUs in DBNs. The second two approaches (JCT and JCS) use the Jaccard statistic as a measure of correlation.

It was shown that the total walking distance marginally increased in comparison to the GP approach while still significantly improving on the historical case. In addition the number of small orders generated was roughly half that of the GP approach. It was also shown that using correlations resulted in a slightly better distribution of volume over picking lines, although, the large picking lines are still undesirable. It is recommended that the ADT desirability scores be used to assign DBNs to picking lines.

Using correlation measures have reduced the number of small cartons produced and improved volume distribution slightly with only a marginal increase in walking distance. Future work may include approaches to reduce these large picking lines using capacity constraints or goal programming techniques.

Acknowledgement

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF.

References

[1] Accorsi R, Manzini R & Bortolini M, 2012, *A hierarchical procedure for storage allocation and assignment within an order-picking system. a case study*, International Journal of Logistics: Research and Applications, **15(6)**, pp. 351–364.

[2] Bartholdi JJ & Hackman ST, 2011, *Warehouse and Distribution Science*, The Supply Chain and Logistics Institute School of Industrial and Systems Engineering Georgia Institute of Technology Atlanta, GA 30332-0205 USA, Available from www.warehouse-science.com.

[3] Bindi F, Manzini R, Pareschi A & Regattieri A, 2009, *Similarity-based storage allocation rules in an order picking system: an application to the food service industry*, International Journal of Logistics Research and Applications, **12(4)**, pp. 233–247.

[4] Chiang DMH, Lin CP & Chen MC, 2011, *The adaptive approach for storage assignment by mining data of warehouse management system for distribution centres*, Enterprise Information Systems, **5(2)**, pp. 219–234.

[5] Hassini E, 2009, *One-dimensional carousel storage problems: Applications, review and generalizations*, INFOR, **47(2)**, pp. 81–92.

[6] IBM, *CPLEX*, Available from www.ibm.com.

[7] Jane CC & Laih YW, 2005, *A clustering algorithm for item assignment in a synchronized zone order picking system*, European Journal of Operational Research, **166(2)**, pp. 489–496.

[8] Kim BS & Smith JS, 2012, *Slotting methodology using correlated improvement for a zone-based carton picking distribution system*, Computers & Industrial Engineering, **62(1)**, pp. 286–295.

[9] Manzini R, 2006, *Correlated storage assignment in an order picking system*, International Journal of Industrial Engineering, **13(4)**, pp. 384–394.

[10] Martello S & Toth P, 1990, *Knapsack Problems Algorithms and Computer Implementations*, Johhn Wiley & Sons, West Sussex England.

[11] Matthews J & Visagie SE, 2013, *Order sequencing on a unidirectional cyclical picking line*, European Journal of Operations Research, **231(1)**, pp. 79–87.

[12] Matthews J & Visagie SE, *Picking line data*, Available from http://scholar.sun.ac.za/handle/10019.1/86110.

[13] Matthews J & Visagie SE, 2014, *SKU arrangement on a unidirectional picking line*, Submitted, **0(0)**, pp. 0–0.

[14] Matthews J & Visagie SE, 2015, *SKU assignment to parallel unidirectional picking lines*, South African Journal of Industrial Engineering, **26(1)**, pp. 235–251.

[15] Paragon Decision Technology, *AIMMS*, Available from www.aimms.com.

[16] Pep Stores Ltd, 2013, *PEP*, Available from www.pepstores.com.

[17] Serdyn H, *Carton sizes*, personal communication. Contactable at:.henniese@pepkorit.com