IIE Transactions (2003) **35**, 419–432
0740-817X © 2003 "IIE"
0740-817X/03 $12.00 + .00
DOI: 10.1080/07408170390187870

Optimal and heuristic algorithms for the multi-location dynamic transshipment problem with fixed transshipment costs

YALE T. HERER[1,]* and MICHAL TZUR[2]

[1]*Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, 32000 Haifa, Israel*
E-mail: yale@techunix.technion.ac.il
[2]*Department of Industrial Engineering, Tel Aviv University, Tel Aviv 69978, Israel*
E-mail: tzur@eng.tau.ac.il

Received May 2000 and accepted January 2002

We consider centrally controlled multi-location systems, which coordinate their replenishment strategies through the use of transshipments. In a dynamic deterministic demand environment the problem is characterized by several locations, each of which has known demand for a single product for each period in a given finite horizon. We consider replenishment, transshipment and inventory holding costs at each location, where the first two have location-dependent fixed, as well as linear components, and the third is linear and identical to all locations. We prove that the resulting dynamic transshipment problem is NP-hard, identify a special structure which is satisfied by an optimal solution and develop, based on this structure, an exponential time algorithm to solve the problem optimally. In addition, we develop a heuristic algorithm, based on partitioning the time horizon, which is capable of solving larger instances than the optimal solution. Our computational tests demonstrate that the heuristic performs extremely well.

1. Introduction and literature review

Centrally controlled multi-location systems are capable of increasing their profitability through coordination. In recent years, organizations have started using novel logistic strategies such as internal and external coordination in order to achieve a competitive advantage. In this paper we consider coordination among facilities through replenishment strategies that take into consideration transshipments, that is, movement of a product between locations at the same echelon level. Consider, for example, a retail chain with several stores located across the country. Further, consider one of their products whose supplier is located outside the country. Every replenishment, by any of the stores, is associated with a fixed replenishment cost which, in the case of a distant supplier, can be quite significant. The stores forecast their demand for the product, according to which they must plan their replenishment strategy. In this context, coordinated replenishment strategies, accompanied by an appropriate transshipment strategy, can result in substantial cost savings.

*Corresponding author

This research was performed, in part, while the author was at the Department of Industrial Engineering, Tel Aviv University.

The main benefit often associated with transshipments is balancing inventory levels at the various locations through emergency stock transfers. In other words, when one location has surplus inventory and a second has a shortage, a transshipment will take place from the first to the second location. Indeed, many of the problems discussed in the literature which consider transshipments are associated with environments in which demand is stochastic, where such emergency transfers are meaningful. In the existing stochastic models demand is always static.

In this paper we consider a problem in which demand is deterministic and dynamic. In this environment, the benefit associated with transshipments is the saving of replenishment costs, as in the example above. This means that when one location (store) replenishes its stock, a large quantity may be requested, some of which is designated to be transshipped to another location. In this way, fixed (and possibly variable) replenishment costs are saved at the expense of transshipping the product between the locations.

Our Dynamic Transshipment Problem (DTP) is characterized by several locations, each of which has a known demand for a single product for each period in a given finite horizon. The requirements have to be satisfied without backlogging. The cost components are replenishment (from an outside supplier), transshipment and inventory holding costs at each location. The first two have both a fixed and a linear component which are location-dependent. The third is linear and identical to all locations.

We consider several mechanisms by which transshipments can take place. The mechanism is determined exogenous to the model. For example, the *Direct transshipment mechanism* requires that transshipments are sent directly from the sending location to the receiving locations. Another example is the *TSP transshipment mechanism* which requires that each time a location sends out a transshipment, a truck follows a traveling salesman tour through the sending location and each of the receiving locations. We consider the former in the body of the paper, while the latter is considered, along with other extensions, at the end of the paper. The combined decision of how much to transship and how these units are transferred from the sending to the receiving locations (given a transshipment mechanism) is referred to as the *transshipment strategy*. The problem is to determine the minimal cost replenishment quantities and transshipment strategy.

The dynamic transshipment problem was introduced by Herer and Tzur (2001). There, two locations were considered and location-dependent holding costs were allowed, giving rise to another reason for using transshipments. In this paper we extend the analysis to multiple locations with identical holding costs. Fixed replenishment costs were considered previously only by Herer and Tzur (2001), as mentioned above, and by Herer and Rashit (1999) in the single period stochastic setting. In addition to these references, much additional work has been carried out in a stochastic setting. Most of this work has concentrated on either the single period or the infinite horizon version of the problem. However, Robinson (1990) examined the finite horizon version as we do here. Additional recent work on transshipments include Archibald *et al.* (1997), Rudi *et al.* (1998) and Tagaras (1999).

We show that our problem is NP-hard and provide an exponential algorithm to solve it optimally. In addition, we develop a time partitioning heuristic based on the general framework of the design of time partitioning heuristics for dynamic lot sizing problems developed by Federgruen and Tzur (1999). The heuristic performs very well in an extensive experiment that we conducted.

The major contributions of this paper are in proving the NP-hardness of the multi-location DTP, identifying the structure satisfied by an optimal solution and developing optimal and heuristic algorithms for the problem. Both algorithms are based on the structure that was found, which is a generalization of related dynamic lot sizing results. In addition, we contribute to the modeling of the dynamic transshipment problem with multiple locations by presenting several possible transshipment mechanisms.

The DTP may be viewed as an extension of the classical single item dynamic lot sizing problem presented by Wagner and Whitin (1958). It is closely related to the Joint Replenishment Problem (see Section 4.3) studied by Joneja (1990), Federgruen and Tzur (1994a), and others, as well as the *one warehouse multi-retailer* problem, studied by Federgruen and Tzur (1999). These latter problems are known to be NP-hard (Arkin *et al.*, 1989). Note that none of these lot sizing problems considered the issue of transshipments.

This paper is organized as follows: in the next section we formulate our model. In Section 3 we prove structural properties of an optimal solution. These properties are then used in Section 4 to develop our optimal algorithm for the problem. In this section we also discuss the complexity of the DTP and our optimal algorithm. In Section 5 we develop our heuristic algorithm, and discuss its complexity while in Section 6 we describe our computational study. Finally, in Section 7 we discuss three extensions to the problem. Section 8 concludes the paper.

2. Model formulation

We follow Herer and Tzur (2001), extending the notation to L locations:

L = number of locations; ($i = 1, \dots, L$), we use p to refer to the location which replenishes from the outside supplier, hereafter referred to as the *replenishing location*, $p \in \{1, \dots, L\}$;

T = number of periods; ($t = 1, \dots, T$);

d_{it} = demand at location i in period t; (for ease of exposition we assume $d_{i1} > 0$ for all i);

h = holding cost per unit per period, independent of the location at which it is held. The assumption of identical holding costs is reasonable considering that a single product is being considered;

K_i = fixed cost incurred whenever location i replenishes (from the outside supplier);

c_i = replenishment cost per unit at location i. Typically we would expect c_i to be independent of i, however, this is not required;

A_{ij} = fixed cost incurred whenever the link between locations i and j is used in the transshipment strategy;

$A_p(\mathscr{I})$ = fixed costs of sending a transshipment from location p to the locations in the set $\mathscr{I} \setminus \{p\}$. Under the direct transshipment mechanism $A_p(\mathscr{I}) = \sum_{i \in \mathscr{I} \setminus \{p\}} A_{pi}$, $p \in \{1, \dots, L\}$, $I \subseteq \{1, \dots, L\}$.

$\hat{c}_{ij}$ = direct variable transshipment cost per unit transshipped from location i to location j;

c_{ij} = effective variable transshipment cost, or simply the variable transshipment cost, per unit transshipped from location i to location j; $c_{ij} = \hat{c}_{ij} + c_i - c_j$.

The direct variable transshipment cost $\hat{c}_{ij}$ may represent the per unit cost of loading the unit onto the vehicle at location i and unloading it at location j. Note that c_{ij} is considered the effective variable transshipment cost because when a unit is transshipped from location i to location j we pay, in addition to the direct variable transshipment cost, a cost of c_i instead of c_j to satisfy a unit of demand at location j. From now on we ignore the variable replenishment cost, as it is included in the variable transshipment cost. (Therefore, the constant $\sum_{i=1}^{L} c_i \sum_{t=1}^{T} d_{it}$ has to be added to the total cost that we obtain in order to get the true cost of a given solution.) Even though c_{ij} can be negative, we observe that $c_{ij} + c_{ji} = \hat{c}_{ij} + \hat{c}_{ji} \geq 0$, thus it is suboptimal to transship items back-and-forth. We assume with respect to the variable transshipment and fixed link costs that the triangle inequality is satisfied, that is: $c_{ij} + c_{jk} \geq c_{ik}$ and $A_{ij} + A_{jk} \geq A_{ik}$. In almost all practical situations this assumption is satisfied.

We say that there is a transshipment from location i to location j when, in a given period, there are items that originate at location i and their final destination is location j. A transshipment cost of c_{ij} is incurred for each unit transshipped in this way. We refer to the decision regarding how much to transship between each pair of locations as the *transshipment quantities* decision. Another decision, referred to as the *transshipment links* decision, specifies the links to be used in order to carry out the transshipment. The transshipment links decision is associated with the fixed costs of transshipments which is determined by the (given) transshipment mechanism. With the Direct transshipment mechanism, a transshipment from location i to location j means traversing the link between location i to j and incurring a cost of A_{ij}. The fixed link (transshipment) costs depend only on the identity of the sending and receiving locations and not on the quantities to be transshipped; thus, this cost is denoted $A_p(\mathscr{I})$. From now through to the end of Section 6 we consider the Direct transshipment mechanism. In Section 7.1 we discuss additional transshipment mechanisms and the cost structure resulting from each of them. Finally, we assume that the time to perform a transshipment is small in comparison to the length of a period.

The dynamic transshipment problem is to find replenishment quantities and transshipment strategies for all locations over the finite horizon, such that demand at every location in every period is satisfied and the sum of fixed replenishment costs, fixed and variable transshipment costs and variable holding costs is minimized. We refer to this problem as the DTP.

It is helpful to represent the flow of items in the DTP with the Direct transshipment mechanism as a fixed cost network flow problem in the following way (see Fig. 1). There are $LT + 1$ nodes: a source node denoted as node 0, and a node for each location i for every period t, denoted as node (i,t), $(1 \leq i \leq L$ and $1 \leq t \leq T)$. There is a demand of d_{it} units at node (i,t), and a supply at the source node of the sum of all demands. The set of arcs consist of replenishment arcs, inventory arcs and transshipment arcs. Replenishment arcs exist between the source node and every other node; inventory arcs exist between node (i,t) to $(i,t+1)$ for $1 \leq i \leq L$ and all $1 \leq t < T$; two transshipment arcs (one in each direction) exist between nodes (i,t) and (j,t) for every pair of nodes $1 \leq i \neq j \leq L$ and $1 \leq t \leq T$. The costs of the flows on these arcs are the replenishment costs (fixed), holding costs (variable) and transshipment costs (fixed plus variable), respectively. We refer to this network as the *replenishment network*. A feasible flow (i.e., a flow that satisfies the demand requirements of the nodes) in this network corresponds with a feasible solution to the DTP and every feasible solution can be represented as a feasible flow.

3. The structure of the optimal solution

In this section we identify the structure of an optimal solution. This structure will form the backbone of both our optimal and heuristic solution procedures. The following lemma is closely related to a well-known result for minimum cost uncapacitated network flow problems with concave costs (Denardo, 1982). While the DTP with the Direct transshipment mechanism is in fact such a problem (thus the result follows from the above reference), this is not the case for the transshipment mechanisms presented in Section 7.1. Therefore, a short proof is provided.

Lemma 1. *There exists an optimal solution to the DTP in which every node in the replenishment network has only one source of supply.*

Proof. For each node with more than one source of supply, choose the source with the least variable source to node unit cost (if several such sources exist, choose one of them arbitrarily). Transfer the flow from all other sources of this node to the flow from the chosen source; this does not increase the variable costs, and may even save on fixed replenishment or transshipment costs. ■

In the following definition of a block, s_i is used to denote the Start of the block at location i; similarly $e_i - 1$ is used to denote its End. In addition, in the following definition and throughout the rest of the paper, whenever $s_i = e_i$ the series $(i, s_i), \ldots, (i, e_i - 1)$ is considered empty.

Definition 1 (block). *A block denoted by* $(s_1, s_2, \ldots, s_L) \rightarrow (e_1, e_2, \ldots, e_L)$ $(1 \leq s_i \leq e_i \leq T + 1$ *for all* $i = 1, \ldots, L$ *and there exists a* $j \in \{1, \ldots, L\}$ *such that* $s_j < e_j)$*, or simply* $(s) \rightarrow (e)$ *is a set of nodes on the replenishment network of the form* $(1, s_1), \ldots, (1, e_1 - 1)$, $(2, s_2), \ldots,$ $(2, e_2 - 1)$, $\ldots\ldots$, $(L, s_L), \ldots, (L, e_L - 1)$ *whose demand is satisfied*

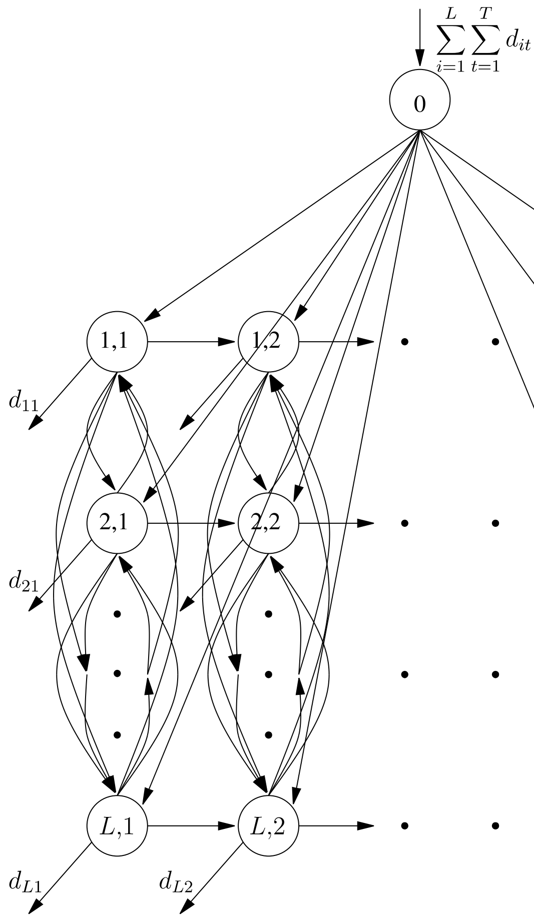

Fig. 1. The replenishment network.

from replenishment(s) within these nodes, and replenishment(s) within these nodes are not used to satisfy demand at nodes outside the block.

For example the block $(1,3,4,4) \rightarrow (3,3,6,5)$ denotes that the demands at location 1 in periods 1 and 2, the demand at location 3 in periods 4 and 5, and the demand at location 4 in period 4 (i.e., nodes (1,1), (1,2), (3,4), (3,5), (4,4) of the replenishment network) are all met through replenishment at these same location-period pairs.

Definition 2 (belong to a block). *Given a block, let $\mathscr{I} = \{i | s_i < e_i\}$ be the set of locations having some of their demand satisfied by a replenishment within the block; if $i \in \mathscr{I}$, we say that location i belongs to the block.*

Using the same example as above, locations 1, 3, and 4 belong to the block $(1,3,4,4) \rightarrow (3,3,6,5)$, and thus $\mathscr{I} = \{1,3,4\}$.

Definition 3 (immediate transshipment). *A transshipment is called immediate if and only if it is performed in the period in which the transshipped items arrive at the replenishing location.*

Definition 4 (special block). $(s_1, s_2, \ldots, s_L) \rightarrow (e_1, e_2, \ldots, e_L)$ *is a special block if the following three conditions are satisfied***:**

1. $(s_1, s_2, \ldots, s_L) \rightarrow (e_1, e_2, \ldots, e_L)$ *is a block.*
2. *All the demand of all nodes in* $(s_1, s_2, \ldots, s_L) \rightarrow (e_1, e_2, \ldots, e_L)$ *is satisfied by one replenishment.*
3. *If the replenishment is at location p, then there is an immediate transshipment to every $i \in \mathscr{I} \backslash \{p\}$ and there are no other transshipments in the block.*

Special blocks have the following properties.

Property 1. If $(s_1, s_2, \ldots, s_L) \rightarrow (e_1, e_2, \ldots, e_L)$ is a special block, then a replenishment occurs at location p in period s_p for some $p \in \mathscr{I}$ such that $s_p = \min_{i \in \mathscr{I}} s_i$. This property holds since a special block contains only one replenishment and no backorders are allowed.

Property 2. If $(s_1, s_2, \ldots, s_L) \rightarrow (e_1, e_2, \ldots, e_L)$ is a special block in a solution that satisfies Lemma 1, then $s_i = s_p$ for all $i \in \mathscr{I}$. This property holds because otherwise, due to Condition 3 of Definition 4, node (i, s_p) would have two sources of supply.

The following lemma, together with Lemma 1, will be needed in the proof of the theorem which characterizes the structure of an optimal solution.

Lemma 2. *There exists an optimal solution to the DTP in which all transshipments are immediate.*

Proof. Given any solution which satisfies Lemma 1 we show how to transform it, without increasing costs, into a solution in which all transshipments are immediate. Consider the earliest replenishment in the solution, and the final destination of all the units included in it. If all units arrive at the location of their final destination in the same period in which they were replenished, then all transshipments associated with this replenishment are immediate, and we proceed to the next replenishment.

On the other hand, if some units arrive at their final destination later than the replenishment period, then we push all the transshipments associated with this replenishment back in time to the replenishment period, thus making them immediate. Furthermore, since there is only one source of supply for each node, if some units are pushed backward, then all units that supply the node in question are pushed backward. The total holding costs remain the same since the holding costs at all locations are identical. The variable transshipment costs are unaffected by the timing in which the transshipments are performed. Moreover, the transshipments may now come directly from the replenishing location rather than spending time at another location, which is less costly by the triangle inequality on the c_{ij}'s. Similarly, the fixed transshipment costs are less costly by the triangle inequality on the A_{ij}'s. Note that a node may have more than one source in the replenishment period in the transformed solution, but since we are examining the replenishments from earlier to later this is of no consequence to the proof.

Repeating this transformation for all replenishments from the earliest to the latest results in all transshipments becoming immediate, with a cost value no larger than the original cost value. ■

Note that we have not yet shown that there is an optimal solution to the DTP that satisfies *both* Lemma 1 and Lemma 2. This fact is proven within the proof of Theorem 1.

Theorem 1. *There exists an optimal solution to the DTP that can be described as a collection of disjoint special blocks.*

Proof. Given any solution we first show how to transform it into a solution which satisfies both Lemmas 1 and 2 without increasing costs. Then we show that the resulting solution is a collection of disjoint special blocks. The transformation consists of two steps:

Step 1. Transform the solution into one in which every node has only one source of supply, as in the proof of Lemma 1.

Step 2. Transform the solution resulting from Step 1 into one in which all transshipments are immediate, as in the proof of Lemma 2.

At the completion of Step 2 some nodes may have more than one source of supply, and therefore Step 1 may need to be repeated. Whenever we perform Step 1, some transshipments may become non-immediate because some replenishments may be eliminated and the transshipments associated with these replenishments may no longer be immediate. As a result, Step 2 may need to be repeated at the completion of Step 1. Therefore, Steps 1 and 2 are executed one after the other until no nodes have more than one source of supply and all transshipments are immediate. This transformation process is finite since in Step 2 transshipments are always pushed backwards in time (and time is discrete), and in Step 1 no new transshipments are formed.

We now show that a solution which satisfies the above two lemmas, satisfies the condition of the theorem. Consider any replenishment and the set of locations which receive units from this replenishment. By Lemma 2 the transshipments to these locations are immediate. For a given location, we know from Lemma 1 that all nodes in the replenishment network that receive supply from this replenishment are consecutive in time and that they are not supplied from anywhere else in the network. Therefore, the nodes that are supplied from this replenishment satisfy the conditions of Definition 4 and therefore form a special block. Since every replenishment in the optimal solution is associated with a special block, Theorem 1 is proved. ■

To illustrate Theorem 1 consider a four location five time period problem. For this problem the following collection of disjoint special blocks forms a solution to the DTP: $(1,1,1,1) \to (1,1,3,2)$, $(1,1,3,2) \to (3,3,3,2)$, $(3,3,3,2) \to (6,3,3,2)$, $(6,3,3,2) \to (6,3,3,6)$, and $(6,3,3,6) \to (6,6,6,6)$.

4. Finding an optimal solution

As a result of Theorem 1 we consider only solutions which are a collection of special blocks. The cost of such a solution is simply the sum of the costs of all the blocks that are included in it, since there is no interaction between blocks. Given the cost of each special block, we show in Section 4.2 how to find an optimal collection of special blocks while in Section 4.3 we discuss the complexity of this algorithm. First, however, we describe how to calculate the cost of a special block.

4.1. *Calculating the cost of a special block*

Recall that the block $(s_1, s_2, \ldots, s_L) \to (e_1, e_2, \ldots, e_L)$ is also denoted by $(s) \to (e)$ and that $\mathscr{I} = \{i | s_i \neq e_i\}$ and $s_i = s_p$ for every $i \in \mathscr{I}$. As a result, and according to the definition of a special block, a replenishment occurs in period s_p for one of the locations in the set $\mathscr{I}$. Denote by $M_p((s) \to (e))$ the minimal cost of the special block given that location p is the replenishing location. $M_p((s) \to (e))$ consists of four parts, the replenishment cost at location p (K_p), the holding cost at all locations ($h \sum_{i \in \mathscr{I}} \sum_{t=s_p}^{e_i-2} \sum_{r=t+1}^{e_i-1} d_{ir} = h \sum_{i \in \mathscr{I}} \sum_{t=s_p+1}^{e_i-1} (t - s_p) d_{it}$), the variable transshipment cost ($\sum_{i \in \mathscr{I} \setminus \{p\}} c_{pi} \sum_{t=s_p}^{e_i-1} d_{it}$), and the fixed transshipment costs ($\sum_{i \in \mathscr{I} \setminus \{p\}} A_{pi}$).

It follows that the cost expression for a special block (replacing $\sum_{i \in \mathscr{I} \setminus \{p\}} A_{pi}$ by $A_p(\mathscr{I})$) is:

$$M_p((s) \to (e)) = K_p + h \sum_{i \in \mathscr{I}} \sum_{t=s_p+1}^{e_i-1} (t - s_p) d_{it} + \sum_{i \in \mathscr{I} \setminus \{p\}} c_{pi} \sum_{t=s_p}^{e_i-1} d_{it} + A_p(\mathscr{I}). \qquad (1)$$

To find the cost of a block, we consider all possible locations as the replenishing location, and choose the best one. Let $M((s) \to (e))$ denote the minimum cost of a special block, then:

$$M((s) \to (e)) = \min_{p \in \mathscr{I}} M_p((s) \to (e)). \qquad (2)$$

For example, consider the special block $(7, 9, 7, 7, 6) \to (9, 9, 8, 7, 6)$. We say that locations 1 and 3 belong to the block and its cost is: $\min[K_1 + hd_{18} + c_{13}d_{37} + A_{13}, K_3 + hd_{18} + c_{31}(d_{17} + d_{18}) + A_{31}]$.

4.2. *Optimal algorithm for solving the DTP*

In this section we present an algorithm for finding an optimal solution by identifying an optimal collection of special blocks. Similarly to the single location Wagner and Whitin (1958) and the two location Herer and Tzur (2001) versions of the problem, we solve our problem by finding a shortest path in an appropriately defined network.

We refer to this network as the *block network* and it is based on an L-dimensional grid which has $T + 1$ nodes in each dimension. We identify a node in this network by an L-tuple, $(s_1, s_2, \ldots, s_L)$ which indicates that the starting inventory at location i in period s_i, $i = 1, \ldots, L$, is zero. An arc from node $(s_1, s_2, \ldots, s_L)$ to node $(e_1, e_2, \ldots, e_L)$ corresponds to the block $(s_1, s_2, \ldots, s_L) \to (e_1, e_2, \ldots, e_L)$. In fact we build the network such that there is an arc from node $(s_1, s_2, \ldots, s_L)$ to node $(e_1, e_2, \ldots, e_L)$ if and only if the block $(s_1, s_2, \ldots, s_L) \to (e_1, e_2, \ldots, e_L)$ is a well defined special block, i.e., $s_i \leq e_i$ for all i (and $s_i < e_i$ for some i) and for each $i \in \mathscr{I}$, $s_i = s_p$. Moreover, we use the notation $(s_1, s_2, \ldots, s_L) \to (e_1, e_2, \ldots, e_L)$(and $(s) \to (e)$) to refer both to the block and to the associated arc. The cost of an arc is set equal to the cost of the associated special block, i.e. $M((s) \to (e))$, see Equation (2).

We note that as a result of the way the block network was constructed:

1. Every path in the block network from node $(1, 1, \ldots, 1)$ to node $(T + 1, T + 1, \ldots, T + 1)$ corresponds to a feasible solution for the DTP. The cost of the path equals the cost of the corresponding DTP solution. With cost equal to the sum of the cost of the arcs.
2. There exists a path in the block network from node $(1, 1, \ldots, 1)$ to node $(T + 1, T + 1, \ldots, T + 1)$ that corresponds to an optimal solution to the DTP (Theorem 1).

As a result of these two observations we see that the DTP can be solved by finding the shortest path in the block network from node $(1, 1, \ldots, 1)$ to node $(T + 1, T + 1, \ldots, T + 1)$.

4.3. *Complexity of the optimal algorithm for solving the DTP*

Herer and Tzur (2001) described a polynomial ($O(T^4)$) algorithm to solve the dynamic transshipment problem with two locations when holding costs are location-dependent. Here we show that with L locations the problem is NP-hard even with identical holding costs at all locations.

Theorem 2. *The DTP is NP-hard.*

Proof. Consider the decision version of the Joint Replenishment Problem (JRP):

Instance JRP: a demand $\bar{d}_{it}$ is specified for $\bar{L}$ products for $\bar{T}$ periods. We incur a fixed replenishment cost $\bar{K}_i$ in every period in which item i is replenished as well as a variable inventory holding cost $\bar{h}_i$ in every period for each unit of item i held in inventory. In addition, we incur a joint fixed replenishment cost $\bar{K}_0$ whenever at least one item is replenished, regardless of the exact set of items being replenished.

The decision version of the JRP was shown to be NP-complete by Arkin *et al.* (1989) for item-dependent holding cost rates. By multiplying the demand of each item by its holding cost and setting the holding cost of each item to one, the result also applies for item-independent holding cost rates. Hence, we assume from now on that $\bar{h}_i = 1$ for all $i = 1, \ldots, \bar{L}$. For a given instance of the decision version of the JRP we define an instance of the decision version of the DTP, as follows: $T = \bar{T}$; $L = \bar{L}$ and each location is associated with an item; $d_{it} = \bar{d}_{it}$, for all i, t; $h = 1$; $c_{ij} = 0$, for all i, j; $A_{ij} = \bar{K}_j$, for all i, j; $K_i = \bar{K}_0 + \bar{K}_i$, for all i.

With this transformation one can see that the total fixed cost of replenishing a set of items $\mathscr{I}$ in the JRP (i.e., $\bar{K}_0 + \sum_{i\in\mathscr{I}} \bar{K}_i$) is the same as the total fixed cost of replenishing any location p in $\mathscr{I}$ and transshipping to the rest of the locations in $\mathscr{I}$. Similarly, the other cost components in a replenishment strategy for the JRP are identical to the costs of the associated replenishment and transshipment strategy for the DTP. Since the reduction is polynomial, this completes the proof. ■

Remark. While we have used the unit scaling technique in order to refer, without loss of generality, to the JRP with item-independent holding costs, a similar scaling is not applicable for the DTP. This is because the cost of transshipping one unit from location i to location j is not well-defined when these two locations have different unit scales.

We now turn our attention to determining the complexity of our optimal algorithm for solving the DTP. The algorithm contains two stages:

1. Building the block network.
2. Finding the shortest path in the block network.

To build the network we have to create its nodes and arcs. Since the block network is L-dimensional with $T+1$ nodes in each dimension, it has $(T+1)^L$ nodes. As there is a one-to-one correspondence between arcs in the block network and special blocks, we determine the number of arcs by counting the number of special blocks. Any subset of locations can make up the members of a special block, thus there are $2^L - 1$ possible combinations of locations that can belong to a special block. As mentioned in Property 2, the starting period is identical for all locations in the special block, and hence takes a value between one and T. The run out time $(e_i - 1)$ for each location i that belongs to the special block can be any one of $O(T)$ periods. Furthermore, since the number of locations that belong to a special block is $O(L)$, the special block may end in any one of $O(T^L)$ ways. Combining these observations we see that there are $O(2^L T^{L+1})$ arcs.

To complete the construction of the block network we need to evaluate the cost of each of the arcs. To do this we first perform an $O(LT^2)$[1] preprocessing step to calculate the following quantities for all $1 \le i \le L$ and $1 \le s < e \le T+1$:

$D_i(s,e)$ = the demand at location i for periods s through $e-1$ inclusive $(=\sum_{t=s}^{e-1} d_{it})$;

$H_i(s,e)$ = the holding cost at location i for periods s through $e-1$ if we replenish in period s for periods s through $e-1$ $(=h\sum_{t=s}^{e-2} D_i(t+1, e-1) = h\sum_{t=s+1}^{e-1}(t-s)d_{it})$.

[1] The naive complexity is $O(LT^3)$, but $O(LT^2)$ can be achieved with minimal care.

Once these calculations are performed we can rewrite Equation (1) as:

$$M_p((s) \to (e)) = K_p + \sum_{i\in\mathscr{I}} H_i(s_i, e_i) + \sum_{i\in\mathscr{I}\setminus\{p\}} c_{pi} D_i(s_i, e_i) + A_p(\mathscr{I}). \qquad (3)$$

Equation (3) can be evaluated in $O(L)$ time and thus evaluating $M((s) \to (e))$, requires $O(L^2)$ time. Since there are $O(2^L T^{L+1})$ arcs, building the network requires $O(L^2 2^L T^{L+1})$ time. Finally, since a shortest path in any network can be calculated in $O(\text{number of arcs})$ time, that is, $O(2^L T^{L+1})$ in our case, the complexity of the algorithm is $O(L^2 2^L T^{L+1})$.

5. Time partitioning heuristic for the multi-location DTP

The algorithm developed in Section 4 may be used to solve instances of the multi-location DTP of limited size. In this section we propose a heuristic which is capable of solving larger instances.

5.1. *Heuristic algorithm for solving the DTP*

The suggested heuristic is based on the idea of *time partitioning*, suggested by Federgruen and Tzur (1994a), initially for the Joint Replenishment Problem. It was later extended to a general framework for the design of time partitioning heuristics for dynamic lot sizing problems, see Federgruen and Tzur (1999). This approach is motivated by forecast horizon results for the single item dynamic lot sizing problem which suggest that optimal or close to optimal initial decisions can be determined on the basis of relatively short horizons (see e.g., Bensoussan *et al*., 1991; Federgruen and Tzur, 1994b). Therefore, in time partitioning heuristics for dynamic lot sizing problems, the complete horizon is partitioned into smaller intervals, such that an instance of the problem is defined for each interval and may be solved to optimality. The problems defined on the intervals are solved sequentially, where the results of an earlier interval determine the starting conditions of the subsequent interval, until a complete solution is constructed.

Following the general steps of the design of time partitioning heuristics (Federgruen and Tzur, 1999), our suggested time partitioning heuristic for the multi-location DTP is as follows:

Step 1. "Identify the collection of intervals into which the full horizon is to be partitioned." We partition the full horizon into non-overlapping intervals. A trade-off exists in determining the length of each interval. On the one hand, an important guideline is to choose a length for each interval small enough such that the multi-location DTP instance defined on it can be solved by our optimal algorithm in a reasonable amount of time. On the other hand, we expect that longer intervals will generally produce better results since then less truncations of the full horizon occur. These truncations are the only reason for the loss of optimality in applying the heuristic procedure. It is recommended to set the interval length to be the longest possible within a reasonable time constraint on the heuristic's run time.

Step 2. "Define initial conditions for the intervals." The subproblem associated with each interval is a slightly modified instance of the multi-location DTP. The parameters of each period in the subproblem are the original parameters of the corresponding intermediate period in the original problem. However, to avoid requiring zero initial and ending inventories in each subproblem, initial conditions are defined, which allow an option for starting inventory (which is the ending inventory of the previous interval). As a result a slightly modified instance of the multi-location DTP is created, see below. In Step 3 we discuss how these differences are handled algorithmically.

The initial conditions of each subproblem (except for the first subproblem which remains unaltered), in essence, add an extra source to period 1, denoted as period 0. The purpose and actual meaning of this extra source is to allow additional units to be replenished in previous subproblems, such that the initial inventory of the current subproblem may be positive. We restrict our attention to adding replenishment to each location only in the latest period (within the previous subproblems) in which replenishment already exists for that location. For some locations the last replenishment period means a direct replenishment of that location from the outside supplier; for other locations it means a period in which transshipment has occurred to that location. In this way, additional replenishment and transshipment quantities are added to previous subproblems. The timing of replenishments and transshipments in previous subproblems remains unchanged. In this way the "extra source" of supply associated with period zero is associated with variable costs only, since we have already accounted for the fixed costs. Within the variable costs, special care has to be taken with respect to holding cost calculations. A unit added to the previous subproblems has to be carried in inventory for as many periods as there are until the beginning of the new subproblem. This procedure is formalized below.

Let DTP_g denote the multi-location DTP associated with the gth interval, and n_g be its length, as determined in Step 1 above. We denote $N_g = \sum_{k=1}^{g} n_k$, i.e., N_g is the original index of the last period in the gth interval. For $g = 1$, DTP_g consists of the first n_1 periods of the original problem. To specify DTP_g for some $g \geq 2$, let $\ell_i(g-1)$ denote the last period in which location i replenished its inventory, either by ordering itself or by receiving a transshipment from some other location in the partial solution of the first $g-1$ intervals. Period 0, which represents the extra source of period 1 in each subproblem, has the following parameters (we add (0) to the notation of all parameters which are constant over time, to denote their value in period 0; for the holding cost rate we distinguish among locations by their index):

- $d_{i0} = 0$ for all i; there is no demand associated with the additional period.
- $A_{ij}(0) = c_{ij}(0) = \infty$ for all i and j; i.e., no transshipments are allowed in period 0.
- $K_i(0) = 0$ for all i; as explained above, additional units may be ordered at each location from the extra source, without re-incurring the fixed replenishment costs.
- $h_i(0) = (N_{g-1} - \ell_i(g-1) + 1)h$; this represents the cost of carrying one unit in inventory from the last time location i was replenished until the beginning of the new subproblem.
- $c_i(0) = 0$ for all i, if in period $\ell_i(g-1)$ location i replenished from the outside supplier, and $c_i(0) = c_{ji}$ if in period $\ell_i(g-1)$ a transshipment occurred from location j to location i; this cost represents the cost of purchasing an additional unit, in the same way it was purchased in its last replenishment period.

It may be observed that in the resulting subproblem, the cost parameters of period 0 are not equal to those of the rest of the periods, and that the demand in period 0 is zero. In this way an instance of a subproblem differs from the original definition of the multi-location DTP.

Step 3. "Apply or develop an exact procedure to solve the subproblem associated with each interval and solve the subproblems sequentially." Here we use the optimal dynamic programming algorithm developed in Section 4, modified to account for the differences of the subproblems from the original problem definition.

Consider the block network associated with an arbitrary subproblem. We represent period 0 in this network by an *additional* arc (beyond the existing one), emanating from node $(1, 1, \ldots, 1)$ to every other node in the network. The cost of the additional arc from node $(1, 1, \ldots, 1)$ to node $(e_1, e_2, \ldots, e_L)$ represents the cost of using period 0 to replenish every location i for which $e_i > 1$. This replenishment will be used to satisfy demand in periods $1, \ldots, e_i - 1$. That is, the cost of adding units to the last replenishment period of location i. This cost is separable by location, since no new replenishments or transshipments, which incur fixed costs, can be made. Therefore the cost of the additional arc is as follows (the indices used in the expression are the indices of the periods in the subproblem): $\sum_{i\in\mathscr{I}}\left[\sum_{t=1}^{e_i-1} d_{it}[c_i(0)+h_i(0)+(t-1)h]\right]$. The first two terms within the inner parentheses represent the purchasing and holding cost up to the beginning of the sub-problem of each unit replenished in period 0; the third term represents the holding costs within the periods of the subproblem for these units. The existing (regular) arc from node $(1,1,\dots,1)$ to node $(e_1,e_2,\dots,e_L)$ represents purchasing the same units in period 1 of the sub-problem. Therefore, the alternative with the least cost should be used in the block network. The rest of the algorithm proceeds as described in Section 4.

It is interesting to note that the form of both the optimal solution and the solution obtained from the above heuristic are the same, i.e., they are collections of disjoint special blocks (cf. Theorem 1).

5.2. *Complexity of the heuristic algorithm for solving the DTP*

The complexity of the heuristic is derived from the complexity of the optimal algorithm for the problem and the number of subproblems solved. Defining the subproblems appropriately and constructing the complete solution at the end of the heuristic takes $O(LT)$ time, which is negligible relative to the rest of the complexity. The complexity of the optimal dynamic programming algorithm as discussed in Section 4.3, is $O(L^2 2^L T^{L+1})$. The modification of the subproblem involves the addition of $O(T^L)$ arcs, whose costs can (with a little bit of care) be calculated in $O(T^L)$ time. Thus, the overall complexity of the algorithm is not affected by the modifications described above. Assume, for simplicity, that all subproblems are of equal length, n, independent of T and that T/n is an integer. In this case, the complexity of the optimal algorithm for each subproblem is $O(L^2 2^L n^{L+1})$ and the complexity of the entire heuristic (which consists of solving T/n subproblems) is $O((T/n)L^2 2^L n^{L+1}) = O(TL^2 2^L n^L)$.

The complexity is now linear in the number of periods, but still exponential in the number of locations, due to the factor $2^L n^L$, where n may be very small (see our computational results where n is chosen to be as small as three). The reason is that the heuristic partitioned the periods, which affected only the complexity associated with the number of periods. Due to the reduced complexity of the heuristic it will run in a reasonable amount of time for values of L and T which are considerably larger than is possible in the optimal algorithm. If needed, the heuristic may be generalized to include a partitioning of the locations as well.

6. Computational tests

To test the effectiveness and behavior of our heuristic, we conducted an extensive experiment which consisted of a total of 1808 runs of our heuristic. In total, nine attributes of the problem were considered; the first two involved the size of the problem, the third was the way in which the horizon was spilt into intervals, and the last six involved various aspects of the problem investigated.

The two attributes that determined the size of the problem are the number of locations (L) and the number of periods (T). The size of the problems investigated was limited since we needed to find the optimal solution for each problem. We used a limit of 12 hours of computer time to find the optimal solution as a guide in limiting the problem size. The problem sizes investigated, hereinafter denoted (L, T), were (1,8), (1,12), (1,24), (2,8), (2,12), (2,24), (3,8), (3,12), (3,24), (4,8), (4,12), (4,16), (4,20), (4,24), (5,8), (5,12), (6,8). Problems with one, two, and three locations were only used to examine trends in the heuristic performance with respect to the problem size; they are not interesting multiple location problems in themselves. Note that the problem size was not limited by our heuristic, but by the optimal algorithm which was needed for comparison.

The third attribute was how the problem was split into subproblems. Problems consisting of eight periods were split into two problems of four periods each (denoted 4-4). Problems consisting of 12, 16, 20, and 24 periods were split either into two equal halves, 6-6, 8-8, 10-10, and 12-12, respectively, or a series of problems each having four periods, 4-4-4, 4-4-4-4, 4-4-4-4-4, and 4-4-4-4-4-4, respectively. In addition, problems having five locations and 12 periods were split into four sub-problems having three periods each, 3-3-3-3, and problems having four locations and 24 periods were split in three additional ways—3-3-3-3-3-3-3-3, 6-6-6-6, and 8-8-8.

For each problem size and each method of splitting the horizon into intervals we compared the heuristic solution to the optimal for 64 problems representing a full factorial experiment of the last six attributes, each tested at two levels. In all problems the holding cost was set equal to one. The six attributes of the problem and values tested are as follows:

1. Mean demand—10 and 50. Note that even though demand is deterministic, it is dynamic. Thus, we randomly drew the demands for each location in each period from a uniform distribution.
2. Coefficient of variation (CV) of demand—0.2 (stable demand) and 0.5 (unstable demand).
3. Number of identical locations in terms of the demand parameters—L and $L-1$. When all locations were not identical the last location took its two demand parameters from the other level of the factorial design. This factor was chosen to examine if the heuristic's performance was affected by the 'identicalness' of the locations.
4. Placement of locations—uniform and separated. In all experiments the locations were confined to the unit

circle. When the placement was uniform, the locations were placed uniformly within the unit circle. When the placement was separated, all locations were placed at a distance of at least 0.5 from the center of the circle. $L-1$ locations were placed together (between 75° and 105°) and the last location was placed by itself (between 180° and 360°). In all cases the locations were placed uniformly in the allowable areas. Note that when the locations were separated and the locations were not identical, the non-identical location was the location that was chosen to be separated from the group. This factor was also chosen to vary the 'identicalness' of the locations. We denote the distance between locations i and j by δ_{ij}.
5. Fixed replenishment cost (K_i)—30 and 60.
6. Fixed and variable transshipment costs between locations i and j (A_{ij}, c_{ij})—($5\delta_{ij}, 1\delta_{ij}$) and ($20\delta_{ij}$, $4\delta_{ij}$). These two attributes were allowed to vary together because they both represent the attribute of transshipment costs.

We split our analysis into two parts. In the first part we examine, for a fixed problem size, the effects of the other seven attributes of the problem. In order to more clearly and concisely present this data we will concentrate on the biggest problem sizes (4,24), (5,12), (6,8). The results for the other problem sizes are qualitatively the same. In Table 1 we present the average values of the ratio of the heuristic solution to the optimal solution[2], organized by the attributes being investigated. We also report, in parentheses after the average, the percentage of the problem instances for which the heuristic solution coincided with the optimal solution. Note that each number presented (except for the last row and last column) represents the average of 32 problems instances.

In the discussion below we will use the term "significant" in its statistical sense. However, the term can also be used in its practical sense and it is questionable whether any of the differences found below are practically significant. The greatest ratio in Table 1 is 1.010, which represents a relative error of only 1%! We performed an ANOVA analysis with a significance level of 95% (see Table 1). We used ANOVA because of its widespread recognition despite the fact that the residuals are not normally distributed with mean zero (as can be inferred from Table 1). We also performed a nonparametric analysis of the results and obtained qualitatively the same results with only slightly different p values.

[2]The true cost of any policy includes the constant $\sum_{i=1}^{L} c_i \sum_{t=1}^{T} d_{it}$ (see Section 2). However, we exclude this constant from our computational study with the note that this is a conservative decision (the inclusion of the constant would decrease the ratio being considered).

Of the six attributes of the problem, three were found to significantly affect the performance of the heuristic. The heuristic performs better with higher mean demands, lower fixed replenishment costs, and lower transshipment costs. These three trends can be viewed as one: each tends to increase the replenishment frequency (Harris, 1915). With a high frequency of replenishments it is more likely that a period in which a replenishment occurs in the optimal solution would be contained within a given interval considered by the heuristic, and therefore would not be "missed" by our heuristic.

One interesting significant second order interaction was found, that being between the mean demand and the CV of demand. A high CV is synonymous with having demand spikes which is sometimes called lumpy demand. In the presence of demand spikes our heuristic is more likely to identify optimal decisions which are involved with fixed costs since the cost associated with non-optimal decisions is relatively much higher. For example, often it is preferable to replenish in a period with a demand spike; the items needed to satisfy the unusually large demand do not have to be held in inventory. Moreover, when the mean demand is also high, the absolute magnitude of the demand spikes increases, thus there is a two-way interaction.

The way the horizon was split into intervals was found to significantly affect the performance of the heuristic. The longer the intervals, the better the heuristic performed, since less "truncations" of the horizon occurred. However, even with a very short interval (three periods) the results can be termed exceptional.

The problem size consists of two attributes: number of locations and number of time periods. We investigated them separately. First we fixed the number of periods at eight, 12 and 24 and allowed the number of locations to vary from one to the point at which computing the optimal solution required more than 12 hours (see Table 2). The length of the subproblems was set to four periods. In Table 2 we see that the quality of the solution generally increased as the number of locations grew and this trend was indeed shown to be significant. We believe the reason for this is that when more locations are involved, there are more possibilities to "correct" a suboptimal decision. A correction can be made, for example, through a transshipment from another location, which makes a direct replenishment; with more locations, there are more such opportunities.

To investigate the effect of the length of the horizon we fixed the number of locations and allowed the horizon length to vary (see Table 3). The sizes of the subproblems were again set to four periods. In the table we see that the quality of the solution generally decreased with the length of the horizon and this trend was indeed shown to be significant. One possible reason for this phenomenon is that for the first subproblem our heuristic finds the optimal solution and 'mistakes' can only creep in with subsequent intervals.

Table 1. Average ratio of heuristic solution to optimal (percent of optimal solutions found)—examining the effects of the various attributes of the problem

Problem size	*(4,24)*					*(5,12)*			*(6,8)*	
Heuristic	*3-3-3-3-3-3-3-3*	*4-4-4-4-4-4*	*6-6-6-6*	*8-8-8*	*12-12*	*3-3-3-3*	*4-4-4*	*6-6*	*4-4*	*Overall*
Mean demand 10	1.010 (28)	1.006 (38)	1.004 (47)	1.002 (53)	1.002 (63)	1.007 (34)	1.006 (47)	1.002 (56)	1.002 (66)	1.005 (48)
Mean demand 50	1.002 (59)	1.001 (69)	1.001 (66)	1.001 (84)	1.000 (88)	1.001 (81)	1.001 (78)	1.000 (94)	1 (100)	1.001 (80)
CV of demand 0.2	1.006 (38)	1.003 (59)	1.003 (63)	1.001 (72)	1.001 (81)	1.004 (66)	1.003 (63)	1.001 (78)	1.001 (78)	1.002 (66)
CV of demand 0.5	1.006 (50)	1.004 (47)	1.003 (50)	1.001 (66)	1.001 (69)	1.004 (50)	1.005 (63)	1.002 (72)	1.001 (88)	1.003 (61)
Identical locations	1.008 (47)	1.004 (63)	1.004 (59)	1.001 (75)	1.001 (75)	1.004 (56)	1.005 (56)	1.002 (69)	1.001 (81)	1.003 (65)
One location different	1.004 (41)	1.003 (44)	1.002 (53)	1.002 (63)	1.001 (75)	1.004 (59)	1.002 (69)	1.001 (81)	1.001 (84)	1.002 (63)
Locations together	1.007 (44)	1.005 (44)	1.003 (47)	1.002 (56)	1.001 (63)	1.006 (50)	1.003 (56)	1.002 (63)	1.001 (84)	1.003 (56)
One location separated	1.005 (44)	1.002 (63)	1.002 (66)	1.001 (81)	1.000 (88)	1.002 (66)	1.004 (69)	1.000 (88)	1.001 (81)	1.002 (72)
Fixed replen. cost 30	1.005 (56)	1.002 (56)	1.002 (59)	1.001 (78)	1.001 (78)	1.002 (69)	1.002 (72)	1.001 (78)	1.001 (84)	1.002 (70)
Fixed replen. cost 60	1.007 (31)	1.005 (50)	1.003 (53)	1.002 (59)	1.001 (72)	1.006 (47)	1.005 (53)	1.001 (72)	1.001 (81)	1.004 (58)
Trans. costs $(5\delta_{ij}, 1\delta_{ij})$	1.004 (66)	1.002 (72)	1.001 (78)	1.000 (84)	1.001 (84)	1.002 (75)	1.003 (81)	1.000 (88)	1.001 (94)	1.001 (80)
Trans. costs $(20\delta_{ij}, 4\delta_{ij})$	1.008 (22)	1.005 (34)	1.004 (34)	1.002 (53)	1.001 (66)	1.006 (41)	1.004 (44)	1.002 (63)	1.001 (72)	1.004 (48)
Overall	1.006 (44)	1.004 (53)	1.003 (56)	1.001 (69)	1.001 (75)	1.004 (58)	1.004 (63)	1.001 (75)	1.001 (83)	1.003 (64)

Table 2. The effect of the number of locations: in all cases the heuristic divided the horizon into subproblems each having four time periods

L	*T*	*Ratio of heuristic solution to optimal solution*	*Percent of optimal solutions found (%)*	*Run time of heuristic (in seconds)*
1	8	1.004	75	<1
2	8	1.003	70	<1
3	8	1.002	75	<1
4	8	1.001	78	<1
5	8	1.001	84	<1
6	8	1.001	83	53
1	12	1.004	75	<1
2	12	1.003	66	<1
3	12	1.003	69	<1
4	12	1.002	66	<1
5	12	1.004	63	3
1	24	1.011	31	<1
2	24	1.006	47	<1
3	24	1.004	50	<1
4	24	1.004	53	<1

Table 3. The effect of the horizon length: in all cases the heuristic divided the horizon into subproblems each having four time periods

L	*T*	*Ratio of heuristic solution to optimal solution*	*Percent of optimal solutions found (%)*	*Run time of heuristic (in seconds)*
3	8	1.002	75	<1
3	12	1.003	69	<1
3	24	1.004	50	<1
4	8	1.001	78	<1
4	12	1.002	66	<1
4	16	1.003	70	<1
4	20	1.003	47	<1
4	24	1.004	53	<1
5	8	1.001	84	<1
5	12	1.004	63	3

7. Extensions

In this section we discuss three extensions of the problem. The first is when different transshipment mechanisms are applied within the transshipment strategy (Section 7.1), the second is when the variable replenishment and transshipment costs may be general concave functions (Section 7.2), and the third is constant supplier lead time. In both cases we demonstrate that our analysis may be extended to these more general cases.

7.1. *Other transshipment mechanisms*

In the analysis of Sections 2–6 we assumed that the Direct transshipment mechanism is used. Indeed, we believe

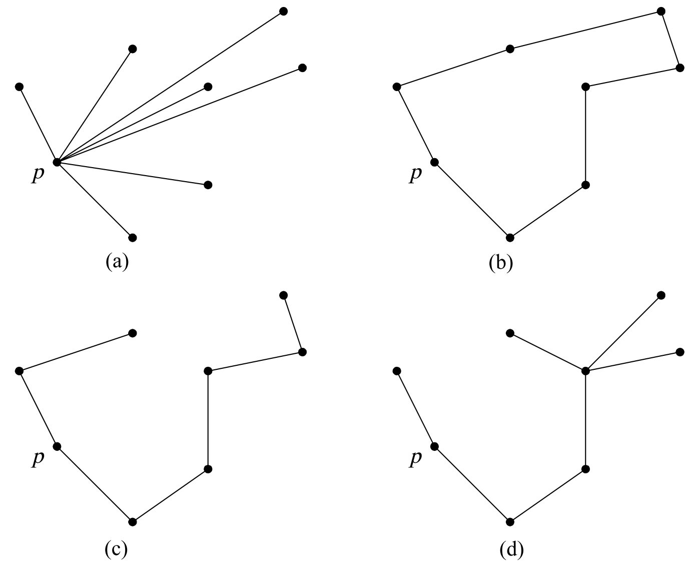

Fig. 2. Illustration of the four transshipment mechanisms: (a) Direct; (b) TSP; (c) Multi-Path and (d) MST.

that this mechanism is the most important one. However, other transshipment mechanisms are also realistic.

We now present three other transshipment mechanisms (see Fig. 2(a–d)). We also consider the implications of these mechanisms on our analysis and conclude that all theoretic as well as algorithmic results in Sections 2–6 apply to all the mechanisms considered. In fact, our results hold as long as $A_p(\mathscr{I})$ is both monotone (i.e., $A_p(\mathscr{I}_1) \leq A_p(\mathscr{I}_1 \bigcup \mathscr{I}_2)$) and subadditive (i.e., $A_p(\mathscr{I}_1 \bigcup \mathscr{I}_2) \leq A_p(\mathscr{I}_1) + A_p(\mathscr{I}_2)$). These two properties are found in most (if not all) practical cases.

7.1.1. *The TSP transshipment mechanism*
This mechanism requires that each time a location sends out a transshipment, a truck follows a traveling salesman tour through the sending location and each of the receiving locations. In this case $A_p(\mathscr{I})$ is the cost of the links in the traveling salesman tour through the locations in the set $\mathscr{I}$.

7.1.2. *The multi-path mechanism*
In this mechanism, the transshipment reaches the receiving locations through a set of paths which originate at location p. Thus, $A_p(\mathscr{I})$ is found by identifying a minimum cost set of disjoint paths originating at location p and covering all locations in the set $\mathscr{I}$.

7.1.3. *The MST transshipment mechanism*
This mechanism requires one to move the items from location p to all other receiving locations by traversing a series of links emanating from the sending location. Its cost is the minimal cost of linking up all the locations into one connected graph using the subgraph induced by the set $\mathscr{I}$. If $A_{ij} = A_{ji}$ for all $i, j \in \mathscr{I}$, then $A_p(\mathscr{I})$ is the length of the minimal spanning tree in the subgraph induced by

the set $\mathscr{I}$. If $A_{ij} \neq A_{ji}$ for some $i, j \in \mathscr{I}$, then we need to solve the related problem of finding the minimal spanning arborescence rooted at p.

We now show that all the results obtained in Sections 2–6 for the Direct transshipment mechanism are valid for the other mechanisms as well. Whenever required, we extend the arguments given earlier. The replenishment network (Fig. 1) still represents the flow of items in the DTP. However, when the other transshipment mechanisms are used, the fixed transshipment costs cannot be associated with the transshipment arcs. This is because the link between the sending location and a receiving location is not necessarily used. On the other hand, other links may be used, as explained above. The fixed link (transshipment) costs still depend only on the identity of the sending and receiving locations and are given by the set of functions $A_p(\mathscr{I})$. Therefore, the replenishment network does not represent a network flow problem any more, and yet the results associated with this network still hold. In particular:

- A feasible flow in the replenishment network corresponds with a feasible solution to the DTP and every feasible solution can be represented as a feasible flow.
- Lemma 1 and its proof hold with no changes.
- Lemma 2 holds with no change. In its proof, the only change is associated with claiming that the fixed transshipment costs are less costly in the transformed solution than in the original solution. This is true since $A_p(\mathscr{I})$ is subadditive.
- Theorem 1 and its proof hold with no changes.
- Equation (1) holds with no change.
- Theorem 2 holds with no change. This generalization is clearly true for the TSP transshipment mechanism, since the TSP itself is NP-hard. For the other two mechanisms, it is still true (as in the proof of Theorem 2) that the total fixed cost of replenishing a set of items $\mathscr{I}$ in the JRP is the same as the total fixed cost of replenishing any location p in $\mathscr{I}$ and transshipping to the rest of the locations in $\mathscr{I}$.
- To efficiently implement Equation (3) we evaluate $A_p(\mathscr{I})$ for all $\mathscr{I}$ in a preprocessing step. Thus the time required to evaluate Equation (3) remains unchanged. However, the time required to evaluate $A_p(\mathscr{I})$ for all $\mathscr{I}$ in the preprocessing step depends on the transshipment mechanism used. If the TSP transshipment mechanism is used, the value of $A_p(\mathscr{I})$ is independent of the replenishing location, p. With this mechanism the preprocessing step may be accomplished in $O(L^2 2^L)$ time by using the standard dynamic program for finding the length of the minimal traveling salesman tour through the set $\{1, \ldots, L\}$ (Held and Karp, 1962). A similar dynamic program, with the same complexity, can be used to evaluate $A_p(\mathscr{I})$ for all $\mathscr{I}$ and a given p for the Multi-path transshipment mechanism. Since p can take on $O(L)$ values, the total complexity of the preprocessing step is $O(L^3 2^L)$. If the MST transshipment mechanism is used with symmetric costs, then $A_p(\mathscr{I})$ can be calculated in $O(L^2)$ time because this is the number of arcs (see e.g., Gondran and Minoux (1989)).
- Thus, since $A_p(\mathscr{I})$ is independent of p, the complexity of the preprocessing step is $O(L^2 2^L)$. If the arc costs are not symmetric, then $A_p(\mathscr{I})$ does depend on location p. The cost of the resulting rooted arborescence problem can also be found in $O(L^2)$ time, see e.g., Gondran and Minoux (1989), thus the total complexity of the preprocessing step is $O(L^3 2^L)$. Were these algorithms more closely examined, it is possible that some savings could be made, but this is not the focus of our paper. If the MST transshipment mechanism is used with asymmetric costs or the Multi-path transshipment mechanism is used, then the complexity of the algorithm is $O(L^2 2^L T^{L+1} + L^3 2^L)$ because of the preprocessing step. Otherwise, the complexity of the algorithm after preprocessing dominates the complexity of the preprocessing step and the complexity of the overall algorithm is $O(L^2 2^L T^{L+1})$.
- Both the optimal and the heuristic algorithms can be implemented without modification.

7.2. *General concave replenishment and transshipment costs*

The analysis thus far also applies if we allow the replenishment costs to be any concave functions (instead of fixed plus linear) and if we allow the variable transshipment costs to be any concave functions (instead of linear). To modify the proof of Lemma 1 we must substitute the word marginal for variable. In the proof of Lemma 2 we need to note that when transshipments are combined the total variable transshipment cost may be reduced due to concavity. The proof of Theorem 1 is unchanged.

7.3. *Constant supplier leadtimes*

Since our model is deterministic we can easily incorporate constant supplier leadtimes into our model. This can be done by solving the problem as presented in this paper and shifting the orders back in time. We note that this extension is particularly appropriate to our model since its motivation, as given in the Introduction, is based on clustered locations and a distant supplier.

8. Conclusions

We have analyzed a multi-location supply chain in which transshipments are allowed. In a dynamic deterministic demand environment, the benefit associated with transshipments is the saving of fixed and possibly variable replenishment costs. The cost of transshipments, on the other hand, is associated with transferring the stock among locations, and is modeled in this paper by both fixed and variable components. Several transshipment mechanisms are considered to model the fixed component of the transshipment costs, each being associated with a different way in which the stock may be transported among locations. We have developed, for this problem, both an optimal and a heuristic algorithm. They are based on a structure which we found to be satisfied by an optimal solution. The heuristic is required for larger instances. Through an extensive computational study we have demonstrated that our heuristic performs very well.

Acknowledgements

The authors gratefully acknowledge the work of Ilan Estrugo who both coded and ran the algorithms described in this paper. We would also like to thank Paula S. Herer for her help in the statistical analysis contained herein.

References

Archibald, T.W., Sassen, S.A.E. and Thomas, L.C. (1997) An optimal policy for a two depot inventory problem with stock transfer. *Management Science*, **43**, 173–183.

Arkin, E., Joneja, D. and Roundy, R. (1989) Computational complexity of uncapacitated multi-echelon production planning problems. *Operations Research Letters*, **8**, 61–66.

Bensoussan, A., Proth, J.M. and Queyranne, M. (1991) A planning horizon algorithm for deterministic inventory management with piecewise linear concave costs. *Naval Research Logistics*, **38**, 729–742.

Denardo, E.V. (1982) *Dynamic Programming Models and Applications*, Prentice Hall, Englewood Cliffs, NJ.

Federgruen, A. and Tzur, M. (1994a) The joint replenishment problem with time-varying parameters: efficient, asymptotic and ϵ-optimal solutions. *Operations Research*, **42**(6), 1067–1086.

Federgruen, A. and Tzur, M. (1994b) Minimal forecast horizons and a new planning procedure for the general dynamic lot sizing model: nervousness revisited. *Operations Research*, **42**(3), 456–468.

Federgruen, A. and Tzur, M. (1999) Time-partitioning heuristics: application to one warehouse, multiitem, multiretailer lot-sizing problems. *Naval Research Logistics*, **46**, 463–486.

Gondran, M. and Minoux, M. (1989) *Graphs and Algorithms*, Wiley, New York, NY.

Harris, F. (1915) *Operations and Cost*, A.W. Shaw Co., Chicago, IL, pp. 48–53.

Held, M. and Karp, R.M. (1962) A dynamic programming approach to sequencing problems. *SIAM Journal on Applied Mathematics*, **10**, 196–210.

Herer, Y.T. and Rashit, A. (1999) Lateral stock transshipments in a two-location inventory system with fixed and joint replenishment costs. *Naval Research Logistics*, **46**, 525–547.

Herer, Y.T. and Tzur, M. (2001) The dynamic transshipment problem. *Naval Research Logistics*, **48**, 386–408.

Joneja, D. (1990) The joint replenishment problem: new heuristics and worst case performance bounds. *Operations Research*, **38**, 711–723.

Robinson, L.W. (1990) Optimal and approximate policies in multiperiod, multilocation inventory models with transshipments. *Operations Research*, **38**, 278–295.

Rudi, N., Kapur, S. and Pyke, D.F. (1998) A transshipment game. Working paper, The Wharton School, Philadelphia, PA.

Tagaras, G. (1999) Pooling in multi-location periodic inventory distribution systems. *OMEGA-International Journal of Management Science*, **27**, 39–59.

Wagner, H.M. and Whitin, T.M. (1958) Dynamic version of the economic lot size model. *Management Science*, **5**, 89–96.

Biographies

Yale T. Herer, B.S. (1986), M.S. (1990), Ph.D. (1990), Cornell University, Department of Operations Research and Industrial Engineering. Yale joined the Faculty of Industrial Engineering and Management at the Technion—Israel Institute of Technology in 1990 immediately after the completion of his graduate studies and, except for a few year hiatus at the Department of Industrial Engineering at Tel Aviv University, he has been there ever since. He has worked for several industrial concerns, both as a consultant and as an advisor to project groups. In 1996 Yale received the *IIE Transactions* best paper award. Yale is a member of the Institute for Operations Research and Management Sciences (INFORMS), Institute of Industrial Engineers (IIE), and the Operations Research Society of Israel (ORSIS). His research interests include supply chain management, especially when integrated with transshipments. He is also interested in production control and production system design.

Michal Tzur is a Senior Lecturer in the Department of Industrial Engineering at Tel Aviv University, Israel. Michal joined Tel Aviv University in 1994 after spending 3 years at the Operations and Information Management department at the Wharton School at the University of Pennsylvania. She received her B.A. from Tel Aviv University and her M. Phil. and Ph.D. from Columbia University. Michal is a member of IIE, INFORMS and POMS. Her research interests are in the areas of supply chain management, multi-echelon inventory management, operations scheduling and production planning.

Contributed by the Location Department