Production/Distribution Planning Problems with Piece-wise Linear Concave Costs*

Lap Mui Ann Chan†
Ana Muriel‡
David Simchi-Levi§

Submitted September 1998; Revised June 4, 2000

Abstract

In this paper we analyze the problem faced by companies that rely on LTL (Less than TruckLoad) carriers for the distribution of products across their supply chain. Typically, these carriers offer volume discounts to encourage larger shipments and, as a result, the transportation charges borne by the shipper are often piecewise linear and concave. In this case, the timing and routing of shipments need to be coordinated so as to minimize system-wide costs, including production, inventory, transportation and shortage costs, by taking advantage of economies of scale offered by the carriers. We assume that the size and due date of shipments required over the planning horizon are known and deterministic. We model this problem using a set-partitioning approach and characterize structural properties of the resulting formulation. These properties are used to identify cases in which the linear programming relaxation of the set-partitioning formulation is tight, and to suggest an efficient algorithm. An extensive computational study illustrates the effectiveness of the algorithm on a set of test problems.

Keywords: Integrating Production, Inventory and Transportation, Logistics, Linear Programming Based Algorithms, Less-Than-Truck-Load (LTL).

1 Introduction

In recent years many companies have recognized that important cost savings and improved service levels can be achieved by effectively integrating production plans, inventory control and transportation policies throughout their supply chains. The focus in this paper is on a planning model integrating decisions across the supply chain for companies that rely on third party carriers.

---

*Research supported in part by ONR Contracts N00014-90-J-1649 and N00014-95-1-0232 and NSF Contracts DDM-9322828 and DMI-9732795.

†School of Management, University of Toronto, Toronto, Ontario

‡Mechanical and Industrial Engineering Dept., University of Massachusetts, Amherst, MA

§Dept. of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL

The model described in this paper is motivated in part by the great development and growth of many competing transportation modes, mainly as a consequence of deregulation of the transportation industry. This has led to a significant decrease in transportation costs charged by third party distributors and therefore there is an ever growing number of companies that rely on third parties for the transportation of their goods.

In the retail, grocery and electronic industries, the most frequently used mode of transportation is the LTL (Less than TruckLoad) mode, which is attractive when shipment sizes are considerably less than truck capacity. Indeed, in these industries, most items are of small size and the number of products required by a specific retail outlet or a grocery store at any given time is fairly small. As a result, LTL carriers offer volume discounts to encourage demand for larger, more profitable shipments. In fact, in many cases, the shipping cost associated with a specific origin-destination pair is a *piece-wise linear concave* function of the quantity shipped. Similarly, production costs can often be approximated by piece-wise linear and concave functions in the quantity produced, e.g, setup plus linear manufacturing costs. These economies of scale motivate the *shipper* to coordinate the production, routing and timing of shipments over the transportation network to minimize system-wide costs.

In this paper, we analyze a deterministic model in which the planner has forecast demand for the next few time periods, say ten to twelve weeks. Of course, forecast demand is not enough to determine an effective inventory policy; uncertainty in demand also needs to be incorporated in the analysis. In practice, this is typically done by decomposing the problem into two parts: The first is identifying an inventory policy that balances holding and fixed cost assuming forecast demand over a given planning horizon, see Stenger (1994). The second is determining safety stock levels and incorporating these in the inventory level that should be maintained at the beginning of each period. Indeed, this is precisely the approach used in a number of Decision Support tools we are familiar with. Thus, the model analyzed in this paper helps optimize inventory decisions associated with the first part of the decomposition approach used in practice.

The objective of the shipper is to find an inventory policy and a routing strategy so as to minimize total cost and satisfy all the demands. Backlogging of demands may be allowed, incurring a known penalty cost which is a function of the length of the shortage period and the level of shortage. In this case, four different costs must be balanced to obtain an overall optimal policy:

production costs, LTL shipping charges, holding costs incurred when carrying inventory at some facility and penalty costs for delayed deliveries.

As discussed in Section 2, this tactical problem can be formulated as a concave cost multicommodity network flow problem. Unfortunately, most of the literature on network flows is devoted to the analysis of minimum-cost network flow problems for which the cost is a linear function of the amount shipped on an arc, see Ahuja, Magnanti and Orlin (1993). In practice, however, situations in which there is a set-up charge, or a discount due to economies of scale give rise to concave cost functions. In this case, an exhaustive search of all extreme points would provide an optimal flow, since a concave function achieves its minimum at an extreme point of the convex feasible region. However, such an approach is impractical for all but the simplest of problems. This, of course, is not surprising since a special case of the problem, the point-to-point connection problem in which the cost of using an edge is simply a fixed charge independent of the quantity shipped, is NP-Complete, see Li, McCormick and Simchi-Levi (1992). Consequently, the exact algorithms that have been developed are either valid only for networks with special structures or run in exponential time in the general case.

For instance, Zangwill (1968) is one of the first authors to analyze the minimum-concave-cost problem. He presents an algorithm with complexity $O(an^d)$, for acyclic networks with a single source (or a single destination), a arcs, n nodes, and $d+1$ destinations (or sources in the single destination case). This algorithm can also be applied to the multicommodity case, again with either a single source or a single destination, since the problem can be reduced to a single-commodity network flow problem. For the general single-commodity minimum-concave-cost problem, Erickson, Monma and Veinott (1987) give a dynamic-programming procedure, called the send-and-split method. The algorithm runs in polynomial time for planar networks in which all demand nodes lie in a bounded number of faces. When the underlying network enjoys the strong-series-parallel property, Ward (1995) develops a polynomial time algorithm to solve the multicommodity network flow problem with aggregate concave cost. This appears to be the first algorithm to solve the problem in polynomial time.

While all algorithms mentioned above are exact and share a dynamic programming approach, Falk and Soland (1969) and Soland (1971) present branch and bound heuristics based on approximations of the concave functions by linear ones. Gallo and Sodini (1979) find local optimality conditions for the concave-cost multicommodity network flow problem on uncapacitated networks, and propose a vertex following algorithm to determine the local minima. Yaged (1971) proposes a different method to find local optima; in this case, the point satisfying the Kuhn-Tucker conditions is found by a successive-approximation, fixed-point algorithm. The quality of the local optimum can be improved by using stronger optimality conditions and a greedy-type algorithm, see Minoux (1989) for a survey of these results. Balakrishnan and Graves (1989) consider a multicommodity network flow problem, very similar to the one analyzed in the current paper, in which the arc costs are piece-wise linear concave functions. They develop a composite algorithm that combines good lower bounds and effective heuristic solutions, based on solving the Lagrangian relaxation of a specific formulation of the problem.

Models integrating inventory control policies and vehicle-routing strategies have also been analyzed extensively in the literature. See Bramel and Simchi-Levi (1997) for a review. These models are quite different from the model analyzed in the current paper due to the structure of the transportation cost and the fact that most of them assume that the shipper operates its own fleet of vehicles.

Related models analyzing the distribution problem from the carriers point of view are discussed in Farvolden, Powell and Lustig (1993) and Farvolden and Powell (1994). The first paper develops a fast algorithm for solving large-scale linear programming multicommodity network flow problems with capacity constraints. The second suggests a heuristic strategy for the problem of determining the number of vehicles the carrier should use in different links of the service network.

Finally, we must point out that the multicommodity network flow problem with piece-wise linear concave costs generalizes the fixed-charge network design problems that arise in various applications in telecommunications, transportation, logistics and production planning (Magnanti and Wong (1984), Balakrishnan, Magnanti and Mirchandani (1997), Balakrishnan et al. (1991), Gavish (1991), Minoux (1989)). These models have been extensively studied, especially in the telecommunications literature in the context of the network loading problem. In this case, capacitated facilities are to be installed on edges of a telecommunication network to support prescribed point-to-point demand flow, see for instance Stoer and Dahl (1994) or Bienstock et al. (1998). For a review, we refer the reader to Gendron, Crainic and Frangioni (1999). A common approach used to solve these network design problems is Lagrangian relaxation, together with dual ascent, subgradient optimization and/or bundle methods to optimize the Lagrangian dual. Crainic, Frangioni and Gendron (1999) report on the performance of different relaxations and dual optimization methods.

This paper is organized as follows. In Section 2 the problem is conveniently modeled to incorporate the time dimension into the model. Section 3 presents a set-partitioning formulation of the problem, which is found to have surprising properties. These properties are used in Section 4 to develop an efficient algorithm, and to show that the linear programming relaxation of the set-partitioning formulation is tight in certain special cases. Computational results, demonstrating the performance of the algorithm on a set of test problems, are reported in Section 5. Finally, the last section points out the generality of the model considered in the paper and extends the work to some other important applications. For simplicity, we consider only the pure distribution problem in most of our analysis, not including production decisions. This is done without loss of generality, as we describe in Section 2, since production decisions can be easily included in our model. However, we do test the efficiency of the algorithm on instances in which production, inventory and distribution decisions needs to be made.

2 The LTL Shipper Model

Consider a generic transportation network, $G = (N, A)$, with a set of nodes N representing the suppliers, warehouses and customers. Customer demands for the next T periods are assumed to be deterministic and each of them is considered as a separate **commodity**, characterized by its origin, destination, size and the time period when it is demanded. Our problem is to plan production and route shipments over time so as to satisfy these demands while minimizing the total production, shipping, inventory and penalty costs.

A standard technique to efficiently incorporate the time dimension into the model, see for instance, Farvolden, Powell and Lustig (1993), is to construct the following **expanded network**. Let $\tau_1, \tau_2, \ldots, \tau_T$ be an enumeration of the relevant time periods of the model. In the original network, G, each node i is replaced by a set of nodes $i_1, i_2, \ldots, i_T$. We connect node i_u with node j_v if and only if $\tau_v - \tau_u$ is exactly the time it takes to travel from i to j. Thus, arc $i_u \to j_v$ represents freight being carried from i to j starting at time τ_u and ending at time τ_v. We call such arcs *shipping links*. In order to account for penalties associated with delayed shipments, a new node is created for each commodity and serves as its ultimate sink. For a given commodity, a link

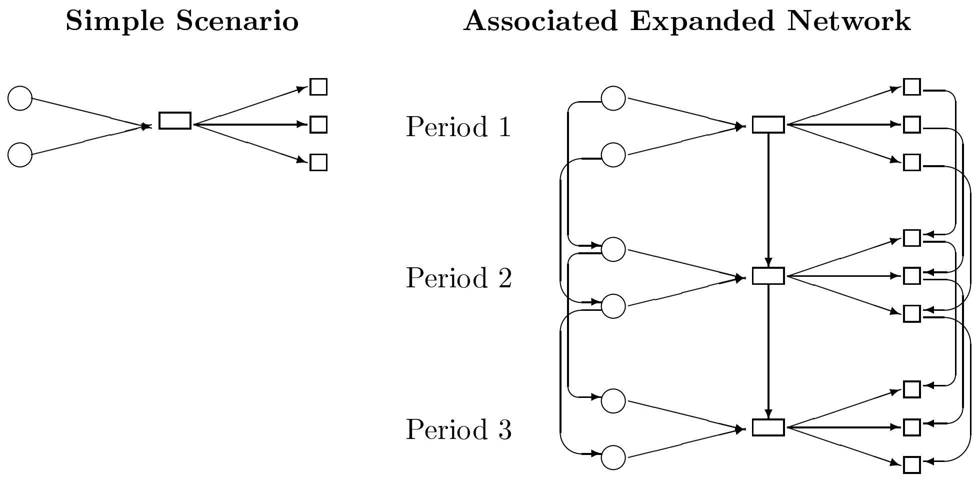

Figure 1: Example of expanded network

between a node representing its associated retailer at a specific time period, and its corresponding sink node, represents the penalty cost of delivering a specific shipment in that time period, and is called *penalty link*. Finally, we add links (i_l, i_{l+1}) for $l = 1, 2, \ldots, T-1$, referred to as *inventory links*. Let $G_T = (V, E)$ be the **expanded network**. Figure 1 illustrates the expanded network for a simple scenario where the shipping and inventory costs have to be balanced over a time horizon of just three periods and shortages are not allowed. For simplicity, we assume that travel times are zero.

Observe that, using the expanded network, the shipper problem can be formulated as a concave-cost multicommodity network flow problem. Production decisions can be easily incorporated into this model, as long as the production cost is a piece-wise linear and concave function of the amount produced. For this purpose, in the expanded network, each production facility at a specific time is represented by two nodes connected by a single link whose cost represents the concave (e.g., set-up plus linear) manufacturing costs. This link is not different from the shipping links in our original model and, consequently, we can restrict the discussion, without loss of generality, to the pure distribution problem.

3 A Set-Partitioning Approach

Let $\mathcal{K} = \{1, 2, \ldots, K\}$ be the index set of all commodities, or different demands with fixed origin and destination, and let w_k, $k = 1, 2, \ldots, K$, be their corresponding size. For instance, commodity $k = 1$ may correspond to a demand of $w_1 = 100$ units that needs to be shipped from a certain supplier to a certain retailer and must arrive by a particular period of time or incur appropriate delay penalties. Let the set of all possible paths for commodity k be P_k and let c_{pk} be the sum of inventory and penalty costs incurred when commodity k is shipped along path p. Observe that the shipping cost associated with a path will depend on the total quantity of all commodities being sent along each of its shipping links and, consequently, it can't be added to the path cost a priori. Thus, each shipping edge, whose cost must be globally computed, needs to be considered separately. Let the set of all shipping edges be SE and for each edge $e \in SE$, let z_e be the total sum of weight of the commodities traveling on that edge.

We assume that the cost of a shipping edge e, $e \in SE$, of the expanded network $G_T(V, E)$, is $F_e(z_e)$, a **piece-wise linear and concave cost function** which is non-decreasing in the total quantity, z_e, of the commodities sharing edge e. As presented in Balakrishnan and Graves (1989), this special cost structure allows for a formulation of the problem as a mixed integer linear program. For this purpose, the piece-wise linear concave functions are modeled as follows. Let R be the number of different slopes in the cost function, which we assume, without loss of generality, is the same for all edges to avoid cumbersome notation. Let M_e^{r-1}, M_e^r, $r = 1, \ldots, R$, denote the lower and upper limits, resp., on the interval of quantities corresponding to the rth slope of the cost function associated with edge e. Note that $M_e^0 = 0$ and M_e^R can be set to the total quantity of all commodities that may use arc e. We associate with each of these intervals, say r, a variable cost per unit, denoted by α_e^r, equal to the slope of the corresponding line segment, and a fixed cost, f_e^r, defined as the y-intercept of the linear prolongation of that segment. See Figure 2 for a graphical representation. Observe that the cost incurred by any quantity on a certain range is the sum of its associated fixed cost plus the cost of sending all units at its corresponding linear cost. That is, we can express the arc flow cost function, $F_e(z_e)$, as

$$F_e(z_e) = f_e^r + \alpha_e^r z_e,$$

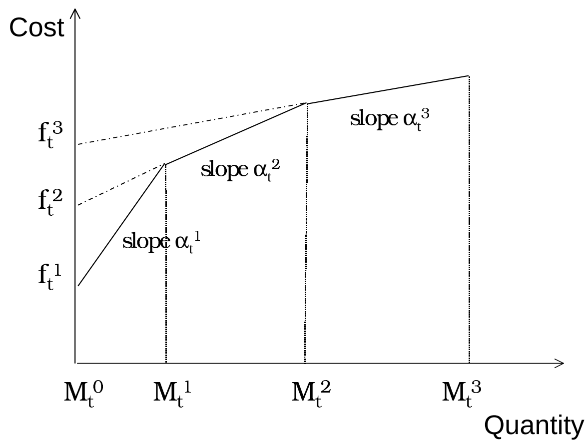

Figure 2: Piece-wise linear and concave cost structure

if $z_e \in (M_e^{r-1}, M_e^r]$.

Clearly,

Property 3.1 *The concavity and monotonicity of the function F_e implies that,*

1. $\alpha_e^1 > \alpha_e^2 > \ldots > \alpha_e^R \geq 0$,
2. $0 \leq f_e^1 < f_e^2 < \ldots < f_e^R$,
3. $F_e(z_e) = \min_{r=1,\ldots,R} \left\{ f_e^r + \alpha_e^r z_e \right\}$. *The minimum is achieved at a unique index s, unless $z_e = M_e^s$, in which case the two consecutive indexes s and $s+1$ lead to the same minimum cost.*

We are now ready to introduce an integer linear programming formulation of the Shipper Problem for this special cost structure. Recall that z_e denotes the total flow on edge e and let z_{ek} be the quantity of commodity k that is shipped along that edge. For all $e \in SE$ and $r = 1, \ldots, R$ define variables,

$$x_e^r = \begin{cases} 1, & \text{if } z_e \in (M_e^{r-1}, M_e^r], \\ 0, & \text{otherwise,} \end{cases}$$

and, in addition, for every k, $k \in \mathcal{K}$, let

$$z_{ek}^r = \begin{cases} z_{ek}, & \text{if } z_e \in (M_e^{r-1}, M_e^r], \\ 0, & \text{otherwise.} \end{cases}$$

In the following, the first set is referred to as *interval* variables and the second as *quantity* variables.

In order to relate these edge flows to path flows we define, for each $e \in SE$ and $p \in \bigcup_{k=1}^{K} P_k$,

$$\delta_p^e = \begin{cases} 1, & \text{if shipping link } e \text{ is in path } p, \\ 0, & \text{otherwise.} \end{cases}$$

Finally, let variables

$$y_{pk} = \begin{cases} 1, & \text{if commodity } k \text{ follows path } p \text{ in the optimal solution} \\ 0, & \text{otherwise,} \end{cases}$$

for each $k \in \mathcal{K}$ and $p \in P_k$. These variables are referred to as *path flow* variables. Observe that defining these variables as binary implies that for every commodity k only one of the variables y_{pk} takes a positive value. This reflects a common business constraint that requires each commodity, that is, items originated at the same source and destined to the same sink in the expanded network, to be shipped along a single path.

In the *Set-Partitioning* formulation of the LTL Shipper Problem, the objective is to select a minimum cost set of feasible paths. Thus, we formulate the LTL shipper problem for piece-wise linear concave edge costs as the following mixed integer linear program, which we denote by Problem P.

$$\text{Problem } P: \quad Min \quad \sum_{k=1}^{K} \sum_{p \in P_k} y_{pk} c_{pk} + \sum_{e \in SE} \sum_{r=1}^{R} \left[f_e^r x_e^r + \alpha_e^r (\sum_{k=1}^{K} z_{ek}^r) \right]$$

$$s.t.$$

$$\sum_{p\in P_k} y_{pk} = 1, \quad \forall k = 1,2,\ldots,K, \tag{1}$$

$$\sum_{p\in P_k} \delta_p^e y_{pk} w_k = \sum_{r=1}^{R} z_{ek}^r, \quad \forall e \in SE, k = 1,\ldots,K, \tag{2}$$

$$z_{ek}^r \le w_k x_e^r \quad \forall e, r, k, \tag{3}$$

$$\sum_{k=1}^{K} z_{ek}^r \le M_e^r x_e^r, \quad \forall e \in SE, r = 1,\ldots,R, \tag{4}$$

$$\sum_{k=1}^{K} z_{ek}^r \ge M_e^{r-1} x_e^r, \quad \forall e \in SE, r = 1,\ldots,R, \tag{5}$$

$$\sum_{r=1}^{R} x_e^r \le 1 \quad \forall e \in SE, \tag{6}$$

$$y_{pk} \in \{0,1\}, \quad \forall k = 1,2,\ldots,K, \ and \ p \in P_k, \tag{7}$$

$$x_e^r \in \{0,1\}, \quad \forall e \in SE, \ and \ r = 1,2,\ldots,R, \tag{8}$$

$$z_{ek}^r \ge 0, \quad \forall e \in SE, \ \forall k = 1,2,\ldots,K, \ and \ r = 1,2,\ldots,R.$$

In this formulation, constraints (1) ensure that exactly one path is selected for each commodity and constraints (2) set the total flow on an edge e to be equal to the total flow of all the paths that use that edge. Constraints (3)–(6) are used to model the piece-wise linear concave function. Constraints (3) specify that if some commodity k is shipped on edge e using cost index r, the associated interval variable, x_e^r, must be 1. Constraints (4) and (5) make sure that if cost index r is used on edge e, then the total flow on that edge must fall in its associated interval, $[M_e^{r-1}, M_e^r]$. Finally, constraints (6) indicate that at most one cost range can be selected for each edge.

Constraints (3) are not required for a correct integer programming formulation of the problem. We include them because they improve significantly the performance of the linear programming relaxation of Problem P.

Observe that this formulation is equivalent to one in which the problem is described using only interval and quantity variables, x_e^r and z_{ek}^r, and flow balance constraints. The resulting formulation contains far less variables and is thus much easier to solve in practice. In fact, it is the one we use in our computational work. However, the set-partitioning formulation described proves to be useful in the derivation of structural properties, and will hence be used throughout the paper.

Let Z^* be the optimal solution to Problem P. Due to the complexity of solving this integer program, our objective is to find a robust and efficient heuristic algorithm. For that purpose, we study the performance of the following relaxations of Problem P:

- *Problem* P_{R_y}: A first relaxation in which constraint (7) is replaced by $y_{pk} \geq 0$ for every p and k. We let Z_{R_y} be its optimal solution. Note that problem P_{R_y} is equivalent to allowing commodities to be split among several routes.
- *Problem* P_{R_x}: A second relaxation where the integrality of the interval variables, x_e^r, is dropped. We let Z_{R_x} be its optimal solution.
- *Problem* P_{LP}: The linear programming relaxation in which all integrality constraints are relaxed. Let Z_{LP} be its optimal solution.

In what follows we analyze the properties of these different models. We first show that the integrality of either path-flow or interval variables, but not both, can be dropped while keeping the same optimal solution. That is,

Property 3.2 $Z^* = Z^{R_y} = Z^{R_x}$

Proof. Refer to Muriel (1997) for details. ∎

In any case, the integrality requirements, even if only kept for a set of variables, make the problem computationally intractable. Therefore, we focus on the study of the linear programming relaxation, Problem P_{LP}. Our objectives are (i) to simplify the formulation so that the linear program relaxation for large-size instances can be solved efficiently, and (ii) to find structural properties of the linear programming relaxation that help in generating a good integer solution from the solution to the linear program. This is the subject of the next two subsections.

3.1 Equivalent Linear Programming Relaxation

Constraints (4)-(6) in Problem P are not required for a correct mixed-integer programming formulation of the problem, as a direct consequence of Property 3.1 (3). In this section, we prove that these constraints are also redundant in the linear programming relaxation of problem P, Problem P_{LP}. This implies that the formulation of the problem can be reduced, while preserving the tightness of its linear programming relaxation.

Let Problem P_{LP}^R be the linear program obtained from Problem P by relaxing the integrality constraints and constraints (4)-(6). That is,

$$\text{Problem } P_{LP}^R: \quad Min \quad \sum_{k=1}^{K}\sum_{p\in P_k} y_{pk}c_{pk} + \sum_{e\in SE}\sum_{r=1}^{R}\Big[f_e^r x_e^r + \alpha_e^r(\sum_{k=1}^{K} z_{ek}^r)\Big]$$

$$s.t.$$

$$(1)-(3)$$

$$y_{pk} \geq 0, \quad \forall k = 1,2,\ldots,K, \ and \ p \in P_k,$$

$$x_e^r \geq 0, \quad \forall e \in SE, \ and \ r = 1,2,\ldots,R,$$

$$z_{ek}^r \geq 0, \quad \forall e \in SE, \ \forall k = 1,2,\ldots,K,$$

$$and \ r = 1,2,\ldots,R.$$

As stated in Property 3.1 (1), the higher the index r, the smaller the associated linear cost α_e^r. Therefore, to minimize cost, as much as possible of each commodity must be sent on intervals with higher index. In the terms of our model this implies that,

Observation 3.3 *For every edge e and range s, $s = 1,2,\ldots,R$, an optimal solution (y,x,z) to Problem P_{LP}^R must satisfy:*

$$\sum_{r=s}^{R} z_{ek}^r = min\{w_k \sum_{r=s}^{R} x_e^r, z_{ek}\};$$

that is,

$$z_{ek}^s = min\{w_k x_e^s, (z_{ek} - w_k \sum_{r=s+1}^{R} x_e^r)^+\},$$

where $z_{ek} = \sum_{r=1}^{R} z_{ek}^r$.

Also, for a solution (y,x,z) to be optimal, the interval variables x must take the smallest possible feasible values. Thus,

Observation 3.4 *For each edge e and range r, $r = 1,\ldots,R$, with associated fixed cost $f_e^r > 0$,*

$$x_e^r = max\{z_{ek}^r/w_k : \ k = 1,\ldots,K\}.$$

That is, there must exist a commodity $k(r)$ such that

$$z^r_{ek(r)} = w_{k(r)} x^r_e.$$

If $f^1_e = 0$ then all values

$$x^1_e \geq max\{z^1_{ek}/w_k : \; k = 1, \ldots, K\}$$

are feasible and lead to the same objective cost.

This is true since otherwise, if $f^r_e > 0$, by reducing x^r_e we would obtain a cheaper feasible solution.

Using these observations we can show the following lemma.

Lemma 3.5 *The optimal solution value to Problem P^R_{LP} is equal to the optimal solution value to the linear programming relaxation of Problem P, Problem P_{LP}.*

Proof. For details, please refer to the on-line Appendix or to Muriel (1997). ∎

In addition, the proof of Lemma 3.5 shows that,

Corollary 3.6 *Given an optimal solution (y, x, z) to problem P^R_{LP}, the cost associated with a shipping edge e with positive fractional flow can be written as*

$$\sum_{r=1}^{R} F_e\left(\frac{\sum_{k=1}^{K} z^r_{ek}}{x^r_e}\right) x^r_e.$$

3.2 The Fixed-Flow Linear Subproblem

In this section, we find structural properties of the relaxed problem, Problem P^R_{LP}. These properties provide the foundation needed for the development of an efficient algorithm to transform the solution to the linear program P^R_{LP} into an integer solution to Problem P. As we shall see, they are also useful in the analysis of its performance. These issues are discussed in the subsequent sections.

In order to analyze the relaxed problem, we start by fixing the fractional path flows and studying the behavior of the resulting linear program. Let $y = (y_{pk})$ be the vector of path flows in a feasible solution to the relaxed linear program, Problem P^R_{LP}. What is the cost that the linear program associates with this solution? What are the values of the corresponding interval and quantity variables, x^r_e and z^r_{ek}?

Observe that, given the vector of path flows y, the amount of each commodity sent on each edge is known and, thus, Problem P_{LP}^R can be decomposed into multiple subproblems, one for every edge. Each subproblem determines the cost that the linear program associates with the corresponding edge flow. We refer to the subproblem associated with edge e as the Fixed-Flow Subproblem on edge e, or Problem FF_y^e.

Let the proportion of commodity k shipped along edge e be

$$\gamma_{ek} = \sum_{p \in P_k} \delta_p^e y_{pk}.$$

Using equation (2), the equality $\sum_{r=1}^R z_{ek}^r = w_k \gamma_{ek}$ must clearly hold; that is, the sum of all the flows of commodity k on the different cost intervals on edge e must be equal to the total quantity, $w_k \gamma_{ek}$, of commodity k that is shipped on that edge.

For each edge e, the total shipping cost on e, as well as the corresponding variables z_{ek}^r and x_e^r, can be obtained by solving the Fixed Flow Subproblem on edge e:

$$\text{Problem } FF_y^e: \qquad Min \quad \sum_{r=1}^R [f_e^r x_e^r + \alpha_e^r \sum_{k=1}^K z_{ek}^r]$$

$$s.t.$$

$$z_{ek}^r \le w_k x_e^r \quad \forall k = 1, \ldots, K, \; and \; r = 1, \ldots, R, \tag{9}$$

$$\sum_{r=1}^R z_{ek}^r = w_k \gamma_{ek}, \quad \forall k = 1, \ldots, K, \tag{10}$$

$$z_{ek}^r \ge 0, \quad \forall k = 1, \ldots, K, \; and \; r = 1, \ldots, R,$$

$$x_e^r \ge 0, \quad \forall r = 1, \ldots, R.$$

Let $C_e^*(y) \equiv C_e^*(\gamma_{e1}, \ldots, \gamma_{eK})$ be the optimal solution to the Fixed-Flow Subproblem on edge e for a given vector of path flows y, or, equivalently, for given corresponding proportions $\gamma_{e1}, \ldots, \gamma_{eK}$ of the commodities on that edge.

The following lemma determines the solution to the subproblem, answering the questions posed above.

Lemma 3.7 *For any given edge $e \in SE$, let the proportion γ_{ek} of commodity k to be shipped on edge e be known and fixed, for $k = 1, 2, \ldots, K$, and let the commodities be indexed in non-decreasing* *order of their corresponding proportions, that is,*

$$\gamma_{e1} \leq \gamma_{e2} \leq \ldots \leq \gamma_{eK}.$$

Then, the optimal solution to the Fixed-Flow Subproblem on edge e is

$$C_e^*(\gamma_{e1}, \ldots, \gamma_{eK}) = \sum_{k=1}^{K} F_e(\sum_{i=k}^{K} w_i)[\gamma_{ek} - \gamma_{ek-1}], \tag{11}$$

where $\gamma_{e0} := 0$.

Proof. We assume w.l.o.g. that the proportions γ_{ek} are positive for all commodities k in the subproblem. Commodities with zero flow on edge e do not play any role in the subproblem and are, thus, removed.

Let x_e^r and z_{ek}^r, for all $r = 1, 2, \ldots, R$ and $k = 1, 2, \ldots, K$, be the components of an optimal solution vector for the subproblem. In the proof of Lemma 3.5 we show that $\sum_{k=1}^{K} z_{ek}^r / x_e^r \in [M_e^{r-1}, M_e^r]$ must hold, for all $r = 1, \ldots, R$. In addition, to avoid ambiguity, we assume that the solution satisfies

$$\sum_{k=1}^{K} z_{ek}^r / x_e^r \in (M_e^{r-1}, M_e^r],$$

for all $r = 1, 2, \ldots, R$, since if $\sum_{k=1}^{K} z_{ek}^r / x_e^r = M_e^{r-1}$ holds, a solution with the same cost can be obtained by adding x_e^r to the current x_e^{r-1} and z_{ek}^r to the current z_{ek}^{r-1}, $\forall k$.

Note that, as pointed out in Corollary 3.6, the cost associated with this solution can be written as,

$$C_e^*(\gamma_{e1}, \ldots, \gamma_{eK}) = \sum_{r=1}^{R} [f_e^r x_e^r + \alpha_e^r (\sum_{k=1}^{K} z_{ek}^r)] = \sum_{r=1}^{R} F_e(\sum_{k=1}^{K} z_{ek}^r / x_e^r) x_e^r.$$

Let S be the highest indexed cost interval, i.e., the one with lowest cost per unit, with positive x_e^S. Clearly this interval must be fully utilized (up to x_e^S), as pointed out in Observation 3.3, and therefore,

$$z_{ek}^S = min\{w_k x_e^S,\ w_k \gamma_{ek}\}. \tag{12}$$

We are now ready to show that the optimal cost to the Fixed-Flow Subproblem satisfies the following recursive equation.

$$C_e^*(\gamma_{e1}, \ldots, \gamma_{eK}) = F_e(\sum_{k=1}^{K} w_k)\gamma_{e1} + C_e^*(\gamma_{e2} - \gamma_{e1}, \ldots, \gamma_{eK} - \gamma_{e1}). \tag{13}$$

Two cases need to be considered.

- *Case 1:* $x_e^S \geq \gamma_{e1}$

  In this case, using (12) we have that $z_{ek}^S = min\{w_k x_e^S,\ w_k \gamma_{ek}\} \geq w_k \gamma_{e1}$ for all commodities $k = 1, 2, \ldots, K$, and we can write the cost associated with interval S as,

$$f_e^S x_e^S + \alpha_e^S(\sum_{k=1}^{K} z_{ek}^S) = [f_e^S \gamma_{e1} + \alpha_e^S(\sum_{k=1}^{K} w_k \gamma_{e1})] + [f_e^S(x_e^S - \gamma_{e1}) + \alpha_e^S(\sum_{k=1}^{K}(z_{ek}^S - w_k\gamma_{e1}))]$$

$$\geq F_e(\sum_{k=1}^{K} w_k)\gamma_{e1} + F_e(\sum_{k=1}^{K} \frac{z_{ek}^S - w_k\gamma_{e1}}{x_e^S - \gamma_{e1}})(x_e^S - \gamma_{e1}).$$

  The last inequality is due to the fact that the concave cost function $F_e(z)$ always chooses the index r that leads to the minimum cost $f_e^r + \alpha_e^r z$. Equality must indeed hold, since our solution is optimal, and we can rewrite the optimal cost as

$$C_e^*(\gamma_{e1}, \ldots, \gamma_{eK}) = F_e(\sum_{k=1}^{K} w_k)\gamma_{e1} + F_e(\sum_{k=1}^{K} \frac{z_{ek}^S - w_k\gamma_{e1}}{x_e^S - \gamma_{e1}})(x_e^S - \gamma_{e1}) + \sum_{r=1}^{S-1} F_e(\sum_{k=1}^{K} z_{ek}^r / x_e^r) x_e^r.$$

  Notice that, ignoring the first term, we are left with a solution to the Fixed-Flow Subproblem on edge e for fractions $\gamma_{e2} - \gamma_{e1}, \ldots, \gamma_{eK} - \gamma_{e1}$ of commodities $2, 3, \ldots, K$, respectively. This solution must also be optimal for the new subproblem, due to the optimality of the original, fractional, solution. Thus,

$$C_e^*(\gamma_{e1}, \ldots, \gamma_{eK}) = F_e(\sum_{k=1}^{K} w_k)\gamma_{e1} + C_e^*(\gamma_{e2} - \gamma_{e1}, \ldots, \gamma_{eK} - \gamma_{e1}).$$

- *Case 2:* $x_e^S < \gamma_{e1}$

  We show by contradiction that this case is not possible. Consider the second highest cost interval, say S', with $x_e^{S'} > 0$. Let $\epsilon = min\{x_e^{S'}, \gamma_{e1} - x_e^S\}$ and let $(\hat{x}, \hat{z})$ be a new solution identical to the previous one, except for

$$\hat{x}_e^S = x_e^S + \epsilon; \quad \hat{z}_{ek}^S = z_{ek}^S + \epsilon w_k, \text{ for } k = 1, 2, \ldots, K,$$

$$\hat{x}_e^{S'} = x_e^{S'} - \epsilon; \quad \hat{z}_{ek}^{S'} = z_{ek}^{S'} - \epsilon w_k, \text{ for } k = 1, 2, \ldots, K.$$

Observe that the difference in cost between new and old solutions is,

$$\epsilon(f_e^S + \alpha_e^S \sum_{k=1}^{K} w_k - f_e^{S'} - \alpha_e^{S'} \sum_{k=1}^{K} w_k) < 0,$$

which contradicts the fact that the initial solution was optimal. The strict inequality holds because $\sum_{k=1}^{K} w_k \in (M_e^{S-1}, M_e^S]$.

Thus, the recursive equation (13) holds, and applying it successively, we get the desired expression for the optimal cost to the Fixed-Flow Subproblem. ∎

Intuitively, the above lemma just says that in an optimal solution to the Fixed Flow Subproblem associated with any edge e, fractions of commodities are consolidated to be shipped at the cheapest possible cost per unit. At first, a fraction γ_{e1} of all commodities $1, 2, \ldots, K$ is available. Thus, these commodities get consolidated to achieve a cost per unit of $F_e(\sum_{k=1}^{K} w_k)/\sum_{k=1}^{K} w_k$, i.e. the cost per unit associated with sending the full K commodities on that edge, and the available fraction γ_{e1} is sent incurring a cost of $\gamma_{e1} F_e(\sum_{k=1}^{K} w_k)$. At that point, none of commodity 1 is left and a fraction $(\gamma_{e2} - \gamma_{e1})$ is the maximum available simultaneously from all commodities $2, 3, \ldots, K$. Again these commodities get consolidated and that fraction, $(\gamma_{e2} - \gamma_{e1})$, from each commodity is sent at a cost $(\gamma_{e2} - \gamma_{e1}) F_e(\sum_{k=2}^{K} w_k)$. This process continues until the desired proportion of each commodity has been sent.

The analysis above also leads to closed form expressions for the values of the interval and quantity variables, x_e^r and z_{ek}^r, associated with any edge $e \in SE$ and a given vector of path flows y. Furthermore, we can show that these values are unique. For details, please refer to the online Appendix or Muriel (1997). Upon inspection of the associated interval and quantity variables, we observe that when the vector of path flows y is integer, the interval variables, x, associated with these path flows are also integer, either 0 or 1. This is in agreement with Property 3.2 above.

Using Lemma 3.7, we can easily calculate the cost that the linear program associates with any given vector y of path flows. Furthermore, we can quickly compute the impact of modifying the flow on certain paths. This is the key to the algorithm developed in the following section, which transforms an optimal fractional solution to the linear program P_{LP}^{R} into an integer solution by modifying path flows, choosing successively for each commodity the path that leads to the lowest increase in the objective of the linear program. In addition, it allows us to prove that in some cases we can transform the optimal fractional solution into an integer solution without increasing the associated costs, showing that the linear programming relaxation is tight.

4 Solution Procedure

We propose the following heuristic for solving Problem P. This algorithm builds upon the insight obtained in the previous analysis, and in particular, on Lemma 3.7. Specifically, the Lemma gives a simple expression of the cost that the relaxed problem, Problem P_{LP}^{R}, assigns to any given fractional path flows. In the algorithm, this expression is used repeatedly to identify the increase in cost incurred when modifying fractional flows.

The Linear Programming Based Heuristic:

Step 1: Solve the linear program, Problem P_{LP}^{R}, using an equivalent formulation where path flow variables are replaced by flow-balance constraints . Initialize $k = 1$.

Step 2: For each arc compute a *marginal cost* which is the increase in cost incurred in the Fixed Flow Subproblem by augmenting the fractional flow of commodity k to 1. Note that this is easy to compute using Lemma 3.7.

Step 3: Determine a path for commodity k by finding the minimum cost path on the expanded network with edge costs equal to the marginal costs.

Step 4: Update the flows and the costs on each link (again employing Lemma 3.7) to account for commodity k being sent along that path.

Step 5: Let $k = k + 1$ and repeat steps (2)–(5) until $k = K + 1$.

Next, we study the performance of this heuristic. Obviously, its effectiveness will strongly depend on the tightness of the linear programming relaxation of Problem P. For this reason, we study the difference between integer and fractional solutions to Problem P for several special cases and in the worst case. We show that,

Theorem 4.1 *In the following cases:*

1. *Single period, multiple suppliers, multiple retailers, two warehouses,*
2. *Two periods, single supplier, multiple retailers, single warehouse,*
3. *Two periods, multiple supplier, multiple retailers, single warehouse using a cross-docking strategy,*
4. *Multiple periods, single supplier, single retailer, single warehouse that uses a cross-docking strategy,*

the solution to the linear programming relaxation of problem P is the optimal solution to the shipper problem. That is,

$$Z^* = Z^{\mathrm{LP}}.$$

Furthermore, in the first three cases, all extreme point solutions to the linear program are integer.

Proof. Please refer to the online Appendix or Muriel (1997) for details. ∎

This theorem demonstrates the exceptional performance of the linear programming relaxation, and consequently of the heuristic, in some special cases. A natural question at this point is whether these results can be generalized. The answer is no in general. To show this, we construct examples with a single supplier and a single warehouse for which

$$\frac{Z^*}{Z^{\mathrm{LP}}} \to \infty,$$

as the number of retailers and time periods increases.

Lemma 4.2 *The linear programming relaxation of Problem P can be arbitrarily weak, even for a single-supplier, single-warehouse, multi-retailer case in which demand for the retailers is constant over time.*

Proof. Please refer to the online Appendix or Muriel (1997) for details. ∎

In such cases the performance of the linear programming based algorithm can be very poor as well. We would like to point out that the instances in which the heuristic solution is found to be arbitrarily bad are characterized by the unrealistic assumption that the shipping cost between two facilities is a pure fixed charge, regardless of quantity shipped, in some periods, while in others that shipping cost is linear (with no fixed charges).

The following section reports the practical performance of the algorithm on a set of randomly generated instances.

5 Computational Results

The computational tests carried out are divided into three categories:

1. Single-period layered networks,
2. General networks,
3. Multi-period single-warehouse distribution problems:
   - Pure distribution instances.
   - Production/distribution instances.

The first two categories are of special interest because they allow us to compare our results with those reported by Balakrishnan and Graves (1989). The third set of problems models practical situations in which each of the retailers is assigned to a single warehouse and production and transportation costs have to be balanced with inventory costs over time.

In the three categories the tests were run on a Sun SPARC20 and CPLEX was used to solve the linear program. During our computational work, we observed that the dual simplex method is more efficient than the primal simplex method in solving these highly degenerate problems, an observation also made by Melkote (1996). This is usually the case for programs with variable upper bound constraints, such as our constraints $z^r_{ek} \leq w_k x^r_e$. We should also point out that most of the CPU time reported in our tests is used in solving the linear program. Thus, to enhance the computational performance of our algorithm and increase the size of the problems that it is capable of handling, future research focused on efficiently solving the linear program is needed. In these tests, however, we focused on evaluating the quality of the integer solutions provided by the heuristic and the tightness of the linear programming relaxation.

We now discuss each class of problems and the effectiveness of our algorithm.

5.1 Single-period Layered Networks

Balakrishnan and Graves (1989) present exceptional computational results for single-period layered networks. In these instances, commodities flow from the manufacturing facilities to distribution centers, where they are consolidated with other shipments. These shipments are then sent to a number of warehouses, where they are split and shipped to their final destinations. Thus, every commodity must go through two layers of intermediate points: **consolidation points**, also referred to as distribution centers, and **breakbulk points**, or warehouses.

To test the performance of our algorithm and to compare it with that of Balakrishnan and Graves (1989), we generated instances of the layered networks following the details given in their paper. In this computational work, five different problem classes, referred to as LTL1 - LTL5, are considered.

Table 1 shows the sizes of the different classes of problems. For each of these classes, the first column (B&G) of Table 2 presents the average ratio between the upper bounds generated by the heuristic proposed by Balakrishnan and Graves and a lower bound on the optimal solution, over 5 randomly generated instances. The numbers are taken from their paper. We do not include, though, their average CPU times because the machines they use are completely different than ours and, in addition, they do not report total computational time for the entire algorithm.

While testing our algorithm, we generated 10 random instances of each of the problem classes, following exactly the process specified in Balakrishnan and Graves (1989). In all of them, our algorithm **finds the optimal integer solution**; furthermore, the solution to the linear program in the first step of our algorithm is integer, providing the optimal solution to the problem. We also report the average CPU time for our algorithm for each problem class. These computational results are reported in Table 2.

Of course, since in all the previous instances the linear program provided the optimal integer solution, the performance of our procedure has not really been tested. In the following subsections we present computational results for problem classes in which the solution to the linear program is

Number of Nodes	Problem Class				
	LTL1	LTL2	LTL3	LTL4	LTL5
SOURCE	4	5	6	8	10
CONSOLIDN	5	10	12	15	20
BREAKBULK	5	10	12	15	20
DESTN	4	5	6	8	10
Arcs	42-47	131-141	190-207	309-312	358-372
Commodities	10	20	30	50	60

Table 1: Test problems generated as in Balakrishnan and Graves (1989)

	B&G	Set-Partitioning	
Problem Class	LB/UB Percentage	LP/Heuristic Percentage	Avge. CPU Time in seconds
LTL1	99.8	100	1.04
LTL2	100	100	7.94
LTL3	99.6	100	20.74
LTL4	99.1	100	55.72
LTL5	99.5	100	100.48

Table 2: Computational results for layered networks. Balakrishnan and Graves results (B&G) versus our Set-Partitioning results.

not always integer.

5.2 General Networks

In this subsection, we report on the performance of our algorithm on general networks generated exactly as they are generated by Balakrishnan and Graves (1989). These results together with those of Balakrishnan and Graves are reported in Table 3. In this category, Balakrishnan and Graves consider five different problem classes, referred to as GEN1,…, GEN5, and generate five random instances for each of them. We, in turn, solve ten different randomly generated instances for each of the problem classes. Again, we do not include their average CPU times due to the reasons mentioned above.

Problem Class	Size			B&G	Set-Partitioning	
	No. of Nodes	No. of Arcs	No. of Comm.	LB/UB Percentage	LP/Heuristic Percentage	Avge. CPU Time in seconds
GEN1	10	47–54	10	99.9	100	2.18
GEN2	15	109–136	20	98.7	99.53	24.04
GEN3	20	196–235	30	98.4	99.88	139.83
GEN4	30	364–428	50	96.2	98.59	1313.06
GEN5	40	340–370	60	98.5	99.98	159.57

Table 3: Computational results for general networks. Balakrishnan and Graves results (B&G) versus our Set-Partitioning results.

5.3 Multi-Period Single-Warehouse Distribution Problems

Here we consider a single-warehouse model where a set of suppliers replenishes inventory of a number of retailers over time. We test two different types of instances: pure distribution instances in which the routing and timing of shipments are to be determined, and production/distribution instances in which the production schedule is also integrated with the transportation and inventory decisions.

5.3.1 Pure Distribution Instances

We assume that shortages are not allowed and analyze three different strategies:

1. *Classical Inventory/Distribution Strategy:* Material flows always from the suppliers through a single warehouse where it can be held as inventory.

2. *Crossdocking Strategy:* All material flows through the warehouse where shipments are reallocated and immediately sent to the retailers.

3. *A Distribution Strategy that Allows for Direct Shipments:* Items may be sent either through the warehouse or directly to the retailer. The warehouse may keep inventory.

For each strategy, we analyze different situations where the number of suppliers is either 1, 2 or 5, the number of retailers is 10, 12 or 20 and the number of periods is 8 or 12. For each combination of the number of suppliers, retailers and periods presented in Table 6, 10 instances are generated. The retailers and suppliers are randomly located on a 1000×1000 grid, while the

Type of arc	α_e^1	α_e^2	α_e^3	SETUP
Supplier-Whs.	0.15	0.105	0.084	25
Whs.-Retailer	0.25	0.20	0.16	10

Table 4: Linear and set-up costs used for all the test problems.

warehouse is randomly assigned to the 400×400 subgrid at the center. Demand is generated for each retailer-supplier pair at each time period, except for the cases with 5 suppliers in which each of these pairs has an associated demand with probability 1/3. These demands are generated from a uniform distribution on the integers in the interval $[0, 100)$.

All suppliers and retailers are linked to the warehouse and the distance associated is the corresponding Euclidean distance between the nodes of the grid. In the case of *a Distribution Strategy that Allows for Direct Shipments*, shipping edges from each of the suppliers to each of the retailers are added. The holding costs per unit of inventory are different at the warehouses and retailer facilities and are presented in Table 5. All holding costs at the suppliers are set to zero. Two shipping-cost functions, representing cost per item per unit distance, are considered: The first is assigned to shipments from the suppliers to the warehouse. The second is incurred by the material flowing from the warehouse to the retailers. The cost function (dollars per mile per unit) associated with direct shipments is equal to that of shipments from the warehouse to a retailer. Both functions have an initial setup cost for using the link and three different linear rates depending on the quantity shipped, see Table 4. However, the ranges to which those linear costs correspond are different for the different Problem classes. This is done so that, in an optimal solution, shipments are consolidated and thus the concave cost function plays an important role in the analysis. These ranges and the corresponding problem classes are presented in Table 5.

Observe, see Table 6, that in most of the instances tested, the linear program is tight and it provides the optimal integer solution. Only in three out of the 150 instances generated, the solution to the linear program is not integer and, in such cases, our algorithm finds a solution which is within 0.8% from the optimal fractional solution.

Problem Class	Inventory Cost		Supplier-Whs. Cost		Whs. - Retailer Cost	
	Warehouse	Retailer	Range 1	Range 2	Range 1	Range 2
I1	5	10	800	1500	200	400
I2					300	600
I3					300	600
I4	10	20	1000	2000	150	300
I5					200	400
I6					200	400
C1	10	20	800	1500	200	400
C2					300	600
C3					300	600
C4	10	20	1000	2000	150	300
C5					200	400
C6					200	400
D1	10	20	500	1000	150	300
D2					200	400
D3					200	400

Table 5: Inventory costs and different ranges for the different test problems.

5.3.2 Production/Distribution Instances

This section demonstrates the effectiveness of the algorithm when applied to production/distribution systems, i.e., systems in which one needs to coordinate production planning, inventory control and transportation strategies over time. For that purpose, we consider the same set of problems, I1 – I3, as in the *Classical Inventory/Distribution Strategy* described in the previous section and add production decisions at each of the supplier sites. This is incorporated into the model as explained in Section 2.

We consider a fixed setup cost for producing at any period plus a certain cost per unit. The setup cost is varied in the set $\{50, 100, 500, 1000\}$ and the linear production cost is set to 1. Inventory holding rate at the supplier site (after production) is set to half of that at the warehouse. For the sixty different instances generated, the linear programming relaxation gave an **integer** solution every time.

STRATEGY	Problem Class	Number of Suppliers	Number of Stores	Number of Periods	$LP/Heuristic$ Percentage	CPU Time in seconds
Classical Inventory/ Distribution Strategy	I1	1	10	12	100	65.21
	I2	2			100	187.37
	I3	5			100	163.23
	I4	1	20	8	99.946	83.5
	I5	2			100	210.51
	I6	5			99.953	200.68
Crossdocking Strategy	C1	1	10	12	100	60.0
	C2	2			100	174.13
	C3	5			100	159.06
	C4	1	20	8	100	79.73
	C5	2			100	202.83
	C6	5			100	186.0
Direct Shipments Allowed	D1	1	12	8	100	51.23
	D2	2			100	165.83
	D3	5			99.921	117.27

Table 6: Computational results for a single warehouse.

6 Conclusions

In this paper we consider the problem of integrating production planning, inventory policies and transportation strategies so as to take advantage of economies of scale. We characterize the structure of the optimal solution to the linear programming relaxation of a set-partitioning formulation of this problem. This allows us to identify instances for which the relaxation is tight and leads to the development of a new algorithm. Computational results show the algorithm to be effective.

It is important to point out that while the algorithm performs very well on the set of test problems, we have been able to construct contrived instances in which the relative deviation between optimal integer and fractional solutions to the set-partitioning formulation is arbitrarily large. For such cases, the algorithm generates poor solutions in general.

Finally, while the framework considered in this paper is that of a production/distribution problem in supply-chain management, the formulation presented is fairly general. Indeed, various other important problems, such as network design and facility location, are amenable to a similar approach. Testing the performance of our algorithm for these special cases is the subject of future research.

7 References

Ahuja, Magnanti and Orlin (1993), *Network Flows: Theory, Algorithms and Applications*. Prentice Hall. Englewood Cliffs, New Jersey.

Balakrishnan, A. and S. Graves (1989), A Composite Algorithm for a Concave-Cost Network Flow Problem. *Networks* **19**, pp. 175–202.

Balakrishnan, A., T. L. Magnanti and P. Mirchandani (1997), Network Design. Chapter 18 in *Annotated Bibliographies in Combinatorial Optimization*, M. Del'Amico, F. Maffioli and S. Martello (Eds.), John Wiley & Sons, pp. 311–334.

Balakrishnan, A., T. L. Magnanti, A. Shulman, and R. T. Wong (1991), Models for Planning Capacity Expansion in Local Access Telecommunication Networks, *Annals of Operations Research* **33**, pp. 239–284.

Bienstock, D., S. Chopra and O. Günlük (1998), Minimum Cost Capacity Installation for Multicommodity Network Flows. *Mathematical Programming* Series B, **81**, No. 3-1, pp. 177–199.

Bramel, J. and D. Simchi-Levi (1997), The Logic of Logistics: Theory, Algorithms, and Applications for Logistics Management. Springer Series in Operations Research, Springer-Verlag, New York.

Crainic, T. G., A. Frangioni and B. Gendron (1999), Bundle-Based Relaxation Methods for Multicommodity Capacitated Fixed Charge Network Design Problems. To appear in Discrete Applied Mathematics.

Erickson, R, C. Monma and A. Veinott, Jr. (1987), Send-and-Split Method for Minimum-Concave-Cost Network Flow. *Mathematics of Operations Research* **12**, pp. 634–664.

Falk, J. and R. Soland (1969), An Algorithm for Separable Nonconvex Programming Problems. *Management Science* **18**, pp. B378–B387.

Farvolden, J. M., W. B. Powell and I. L. Lustig (1993), A Primal Partitioning Solution for Multicommodity Network Flow Problem. *Operations Research* **41**, pp. 669–693.

Farvolden, J. M. and W. B. Powell (1994), Subgradient Methods for the Service Network Design Problem. *Transportation Science* **28**, pp. 256–272.

Gavish, B. (1991), Topological Design of Telecommunications Networks-Local Access Design Methods. *Annals of Operations Research* **33**, pp. 17–71.

Gallo, G. and C. Sodini (1979), Concave Cost Minimization on Networks. *European Journal of Operations Research* **3**, pp. 239–249.

Gendron, B., T. G. Crainic and A. Frangioni (1999), Multicommodity Capacitated Network Design. *Telecommunications Network Planning*, B. Sansó and P. Soriano (Eds.), Kluwer Academic Publishers, pp. 1–19.

Li, C. L, S. T. McCormick and D. Simchi-Levi (1992), The Point-to-Point Delivery and Connection Problems: Complexity and Algorithms. *Discrete and Applied Mathematics* **36**, pp. 267–292. North Holland.

Magnanti, T. L. and R. T. Wong (1984), Network Design and Transportation Planning: Models and Algorithms. *Transportation Science* **18**, pp. 1–55.

Melkote, S. (1996), Integrated Models of Facility Location and Network Design. Ph.D. Thesis, Northwestern University.

Minoux M. (1989), Network Synthesis and Optimum Network Design Problems: Models, Solution Methods and Applications. *Networks* **19**, pp. 313–360.

Muriel, A. (1997), On the effectiveness of the Set-Partioning Formulation for Large-Scale Machine Scheduling and Supply-Chain Management Problems. Ph.D. Thesis, Northwestern University.

Soland, R (1971), Optimal Facility Location with Concave Costs. *Operations Research* **22**, pp. 373–382.

Stenger, A. J., (1994), Distribution Resource Planning. The Logistics Handbook, Robeson and Copacino (eds.), Free Press.

Stoer M. and G. Dahl (1994), A Polyhedral Approach to Multicommodity Survivable Network Design. *Numerische Mathematik* **68**, pp. 149–167.

Ward, J. A. (1995), Minimum-Aggregate-Concave-Cost Multicommodity Flows in Strong-Series-Parallel Networks. Work in paper, Hewlett-Packard Laboratories.

Yaged, B. (1971), Minimum Cost Routing for Dynamic Network Models. *Networks* **1**, pp. 139–172.

Zangwill, W. I. (1968), Minimum Concave Cost Flows in Certain Networks. *Management Science* **14**, pp. 429–450.

A Properties of the Set-Partitioning Formulation

A.1 Proof of Lemma 3.5

It is sufficient to show that there always exists an optimal solution to Problem P_{LP}^R which is feasible for the linear programming relaxation of Problem P.

First we show that any optimal solution to Problem P_{LP}^R must satisfy inequalities (4) and (5) The proof is by contradiction. Assume that there exists an optimal solution (y, x, z) of Problem P_{LP}^R such that for certain $e \in SE$ and $r \in \{1, 2, \ldots, R\}$, either (4) or (5) is not satisfied. Note that x_e^r must be positive since, otherwise, the variables z_{ek}^r would also be zero to satisfy constraint (3) and the above constraints, i.e., constraints (4) and (5), would be trivially satisfied, in contradiction with our assumption. Therefore, there must exist $x_e^r > 0$ and z_{ek}^r, $k = 1, 2, \ldots, K$, such that

$$\frac{\sum_{k=1}^K z_{ek}^r}{x_e^r} \notin [M_e^{r-1}, M_e^r].$$

Let s denote an index such that $\frac{\sum_{k=1}^K z_{ek}^r}{x_e^r} \in [M_e^{s-1}, M_e^s]$. We can consider w.l.o.g. that $M_e^R = \infty$ to ensure that such interval always exists. Then, using Property 3.1 (3) above,

$$f_e^r x_e^r + \alpha_e^r (\sum_{k=1}^K z_{ek}^r) = [f_e^r + \alpha_e^r \frac{\sum_{k=1}^K z_{ek}^r}{x_e^r}] x_e^r > [f_e^s + \alpha_e^s \frac{\sum_{k=1}^K z_{ek}^r}{x_e^r}] x_e^r = F_e(\frac{\sum_{k=1}^K z_{ek}^r}{x_e^r}) x_e^r.$$

Thus, a new solution $(y, \hat{x}, \hat{z})$, identical to the original one except that

$$\hat{x}_e^r = 0 \text{ and } \hat{z}_{ek}^r = 0 \quad \forall k \in \mathcal{K},$$

$$\hat{x}_e^s = x_e^r + x_e^s, \text{ and } \hat{z}_{ek}^s = z_{ek}^r + z_{ek}^s \quad \forall k \in \mathcal{K},$$

will have a better cost. This contradicts the optimality of our original solution.

It remains to show that constraint (6) is also redundant. For that purpose, consider an optimal solution (y, x, z) to problem P_{LP}^R. If $f_e^1 = 0$ for some edge e, we choose the solution (y, x, z) such that the corresponding interval variables are set to their smallest possible value, $x_e^1 = max\{z_{ek}^1 / w_k : k = 1, \ldots, K\}$. We show by contradiction that this solution must satisfy $\sum_r x_e^r \leq 1$, for all shipping edges, e.

Assume that for a certain shipping link e the corresponding $x_e^1, x_e^2, \ldots, x_e^R$ satisfy $\sum_{r=1}^{R} x_e^r > 1$, and let m be the smallest index such that $x_e^m > 0$. Observation 3.4 together with the way in which we select the solution if $f_e^1 = 0$, implies that there exist $l \geq 1$ commodities, k_i, $i = 1, \ldots, l$, such that $z_{ek_i}^m = w_{k_i} x_e^m$.

For those commodities, in view of Observation 3.3 and the fact that $z_{ek_i}^m = w_{k_i} x_e^m > 0$, for all $i = 1, 2, \ldots, l$, we have,

$$z_{ek_i} > \sum_{r=m+1}^{R} z_{ek_i}^r = w_{k_i} \sum_{r=m+1}^{R} x_e^r.$$

Thus,

$$z_{ek_i} = \sum_{r=m}^{R} z_{ek_i}^r = w_{k_i} \sum_{r=m}^{R} x_e^r > w_{k_i},$$

since m is the smallest index such that $x_e^m > 0$ and $\sum_{r=1}^{R} x_e^r > 1$.

Observe that the quantity being shipped, z_{ek_i}, is larger than the demand, w_{k_i}. However, the total flow of commodity k_i on the paths that use edge e, $\sum_{p \in P_{k_i}} \delta_p^e y_{pk_i} w_{k_i}$, cannot exceed w_{k_i} due to constraints (1). Therefore, constraints (2) are not satisfied for commodities k_i, $i = 1, 2, \ldots, l$. This contradicts the feasibility of the original solution and completes the proof.

A.2 Uniqueness of the LP Solution

For each range r, $r = 1, 2, \ldots, R$, let $\mathcal{K}(r) = \{k \in \mathcal{K} \text{ such that } \sum_{i=k}^{K} w_i \in (M_e^{r-1}, M_e^r]\}$. The set $\mathcal{K}(r)$, for $r = 1, 2, \ldots, R$, can be written as,

$$\mathcal{K}(r) = \{k \in \mathcal{K} \text{ such that } \sum_{i=k}^{K} w_i \in (M_e^{r-1}, M_e^r]\} = \{k_r, k_r + 1, \ldots, k_r + n_r - 1\},$$

where k_r is the smallest index among all commodities in the set $\mathcal{K}(r)$ and n_r is the number of commodities in that set. Observe that the higher the index of the commodity, k, the lower the index r of the set $\mathcal{K}(r)$ it belongs to.

Lemma A.1 *Consider a given set of proportions of the commodities to be shipped on edge e, and let the commodities be indexed such that $\gamma_{e1} \leq \gamma_{e2} \leq \ldots \leq \gamma_{eK}$. The associated interval and* *quantity variables optimal for the Fixed Flow Subproblem can be written as,*

$$x_e^r = \sum_{k \in \mathcal{K}(r)} (\gamma_{ek} - \gamma_{ek-1}), \;\; \textit{where } \gamma_{e0} := 0,$$

$$z_{ek}^r = \begin{cases} 0 & \textit{if } k \in \bigcup_{s>r} \mathcal{K}(s), \\ w_k \sum_{l=k_r}^{k} (\gamma_{el} - \gamma_{el-1}) & \textit{if } k \in \mathcal{K}(r), \\ w_k x_e^r & \textit{otherwise.} \end{cases}$$

This solution is unique if the initial fixed cost $f_e^1 > 0$ and $\sum_{i=k}^{K} w_i \neq M_e^r$ for all $r = 1, 2, \ldots, R-1$ and $k \in \mathcal{K}(r)$.

Note that when $f_e^1 = 0$, any feasible value of x_e^1 leads to the same objective cost. Also, if $\sum_{i=k}^{K} w_i = M_e^r$ for a certain commodity k and cost interval r, then an optimal solution different than the one presented in the Lemma can be obtained. This is done by adding commodity k, or any fraction of it, to set $\mathcal{K}(r+1)$ instead of $\mathcal{K}(r)$. Thus, the solution is not unique in such cases.

Proof. The value of x_e^r is easily derived from Lemma 3.7. It is enough to substitute, for every $k \in \mathcal{K}(r)$, the quantity $F_e(\sum_{i=k}^{K} w_i)$ in equation (11) by the function $f_e^r + \alpha_e^r \sum_{i=k}^{K} w_i$. To derive the value for z_{ek}^r, we use Observation 3.3 which shows that $z_{ek}^r = min\{w_k x_e^r, (w_k \gamma_{ek} - \sum_{s=r+1}^{R} w_k x_e^s)^+\}$, for $k = 1, 2, \ldots, K$. Substituting x_e^s, for $s = r, \ldots, R$, we obtain the value of z_{ek}^r.

It remains to show that the solution is unique under the above conditions. Assume that $f_e^1 > 0$ and $\sum_{i=k}^{K} w_i \neq M_e^r$ for all $k \in \mathcal{K}$ and $r = 1, 2, \ldots, R-1$. We show by contradiction that the solution to the Fixed Flow Subproblem presented in Lemma A.1 is unique. For that purpose, let $(\bar{x}, \bar{z})$ be an optimal solution different than that in the Lemma and let S be the highest index interval whose associated variables do not satisfy the expressions in the Lemma. That is,

$$\bar{x}_e^S \neq \sum_{k \in \mathcal{K}(S)} (\gamma_{ek} - \gamma_{ek-1}), \;\; \text{and } \bar{x}_e^l = \sum_{k \in \mathcal{K}(l)} (\gamma_{ek} - \gamma_{ek-1}), \;\; \text{for all } l > S.$$

Two cases need to be considered.

- *Case 1:* $\bar{x}_e^S < \sum_{k \in \mathcal{K}(S)} (\gamma_{ek} - \gamma_{ek-1})$:

  Let u be the index of the commodity such that: If $\bar{x}_e^S < \gamma_{ek_S} - \gamma_{ek_S - 1}$ then $u = k_S$. Otherwise, u satisfies $\sum_{k=k_S}^{u-1} (\gamma_{ek} - \gamma_{ek-1}) \leq \bar{x}_e^S < \sum_{k=k_S}^{u} (\gamma_{ek} - \gamma_{ek-1})$. Observe that this, together with the definition of x_e^l for $l > S$, implies that $\gamma_{eu-1} \leq \sum_{l \geq S} \bar{x}_e^l < \gamma_{eu}$. Let r be the highest index interval smaller than S and with $\bar{x}_e^r > 0$, which must clearly exist since there are still remaining proportions to be shipped. This, of course, implies that $S > 1$.

Define ϵ as

$$\epsilon = \sum_{k=k_S}^{u} (\gamma_{ek} - \gamma_{ek-1}) - \bar{x}_e^S,$$

and note that $\epsilon > 0$.

Consider a new feasible solution $(\hat{x}, \hat{z})$ identical to $(\bar{x}, \bar{z})$, except for

$$\hat{x}_e^S = \bar{x}_e^S + min\{\epsilon, \bar{x}_e^r\} \text{ and } \hat{z}_{ek}^S = \bar{z}_{ek}^S + min\{\epsilon w_k, \bar{z}_{ek}^r\} \text{ for } k = 1, 2, \ldots, K;$$

$$\hat{x}_e^r = \bar{x}_e^r - min\{\epsilon, \bar{x}_e^r\} \text{ and } \hat{z}_{ek}^r = \bar{z}_{ek}^r - min\{\epsilon w_k, \bar{z}_{ek}^r\} \text{ for } k = 1, 2, \ldots, K.$$

Using Observation 3.3 and the fact that $\gamma_{eu-1} \leq \sum_{l \geq S} x_e^l < \gamma_{eu}$, we have,

$$\bar{z}_{ek}^r = min\{\bar{x}_e^r w_k, (w_k \gamma_{ek} - \sum_{l \geq S} w_k \bar{x}_e^l)^+\} = \begin{cases} 0 & \text{if } k < u, \\ w_k min\{\bar{x}_e^r, \gamma_{ek} - \sum_{l \geq S} \bar{x}_e^l\} & \text{if } k \geq u \end{cases}$$

Note that, the definitions of ϵ and x_e^l for $l > S$ imply that, for all $k \geq u$,

$$\gamma_{ek} - \sum_{l \geq S} \bar{x}_e^l \geq \gamma_{eu} - \sum_{l \geq S} \bar{x}_e^l = \sum_{v=k_s}^{u} (\gamma_{ev} - \gamma_{ev-1}) - x_e^S = \epsilon.$$

Therefore,

$$min\{\epsilon w_k, \bar{z}_{ek}^r\} = min\{\epsilon w_k, \bar{x}_e^r w_k, (w_k \gamma_k - \sum_{l \geq S} w_k \bar{x}_e^l)^+\} = \begin{cases} 0 & \text{if } k < u, \\ w_k min\{\epsilon, \bar{x}_e^r\} & \text{if } k \geq u \end{cases}$$

The difference in cost between new and old solutions can, thus, be written as,

$$min\{\epsilon, \bar{x}_e^r\}[f_e^S + \alpha_e^S \sum_{k=u}^{K} w_k - (f_e^r - \alpha_e^r \sum_{k=u}^{K} w_k)] < 0,$$

which contradicts the optimality of the original solution, $(\bar{x}, \bar{z})$. The strict inequality is true because $u \in \mathcal{K}(S)$ and, therefore, S is the unique cheapest interval for the quantity $\sum_{k=u}^{K} w_k$.

- *Case 2:* $\bar{x}_e^S > \sum_{k\in\mathcal{K}(S)}(\gamma_{ek} - \gamma_{ek-1})$:

  If $S = 1$, that is, $\bar{x}_e^l = \sum_{k\in\mathcal{K}(l)}(\gamma_{ek} - \gamma_{ek-1})$ for all $l > 1$, and $\mathcal{K}(S) = \{k_1, k_1+1, \ldots, K\}$, then for any commodity k, $k = 1, 2, \ldots, K$, we have,

$$z_{ek}^1 = min\{w_k \bar{x}_e^1, (w_k\gamma_{ek} - \sum_{l>S} w_k \bar{x}_e^l)^+\} < w_k \bar{x}_e^1.$$

  This is true since,

$$(\gamma_{ek} - \sum_{l>1} \bar{x}_e^l)^+ = (\gamma_{ek} - \gamma_{ek_1-1})^+ = \begin{cases} 0 & \text{if } k < k_1, \\ \gamma_{ek} - \gamma_{ek_1-1} & \text{if } k \geq k_1, \end{cases}$$

  and, for all $k \geq k_1$,

$$\gamma_{ek} - \gamma_{ek_1-1} \leq \gamma_{eK} - \gamma_{ek_1-1} = \sum_{k\in\mathcal{K}(1)} (\gamma_{ek} - \gamma_{ek-1}) < \bar{x}_e^1.$$

  Therefore, $\bar{x}_e^1$ can be reduced to $\bar{x}_e^1 = \max_k\{z_{ek}^1/w_k\}$ to obtain a feasible solution with a better associated cost, since $f_e^1 > 0$, in contradiction with the optimality of $(\bar{x}, \bar{z})$.

  Assume now that $S > 1$ and let u be the index of a commodity such that: If $x_e^S \geq \sum_{k=k_S}^K(\gamma_{ek} - \gamma_{ek-1})$ then $u = K$. Otherwise, u satisfies

$$\sum_{k=k_S}^{u} (\gamma_{ek} - \gamma_{ek-1}) \leq \bar{x}_e^S < \sum_{k=k_S}^{u+1} (\gamma_{ek} - \gamma_{ek-1}).$$

  Clearly, $u \geq k_S + n_S - 1$ since $\bar{x}_e^S > \sum_{k\in\mathcal{K}(S)}(\gamma_{ek} - \gamma_{ek-1})$.

  Let $\epsilon = \bar{x}_e^S - \sum_{k=k_S}^u(\gamma_{ek} - \gamma_{ek-1})$. Observe that in this case the quantity variables are given by,

$$\bar{z}_{ek}^S = \begin{cases} 0 & \text{if } k < k_S, \\ w_k \sum_{l=k_S}^k(\gamma_{el} - \gamma_{el-1}) & \text{if } k_S \leq k \leq u, \\ w_k \bar{x}_e^S & \text{otherwise.} \end{cases}$$

We can split the cost associated with interval S, $f_e^S \bar{x}_e^S + \alpha_e^S \sum_{k=1}^{K} \bar{z}_{ek}^S$, in the following way:

$$\begin{aligned}
&\sum_{k\in\mathcal{K}(S)} [f_e^S + \alpha_e^S \sum_{i=k}^{K} w_i](\gamma_{ek} - \gamma_{ek-1}) + \\
&\sum_{k=k_S+n_S}^{u} [f_e^S + \alpha_e^S \sum_{i=k}^{K} w_i](\gamma_{ek} - \gamma_{ek-1}) + [f_e^S + \alpha_e^S \sum_{i=u+1}^{K} w_i]\epsilon \\
&> \sum_{k\in\mathcal{K}(S)} F_e(\sum_{i=k}^{K} w_i)(\gamma_{ek} - \gamma_{ek-1}) + \\
&\sum_{k=k_S+n_S}^{u} F_e(\sum_{i=k}^{K} w_i)(\gamma_{ek} - \gamma_{ek-1}) + F_e(\sum_{i=u+1}^{K} w_i)\epsilon.
\end{aligned}$$

In the cases of $u = k_s + n_s - 1$ or $u = K$, we consider the summations with upper limit smaller than the lower limit to be 0.

If $u \geq k_S + n_S$ or $u < K$, the strict inequality comes from the fact that for $k \geq k_S + n_S$ $\sum_{i=k}^{K} w_k < M_e^{S-1}$. A new solution with better cost could then be obtained by using for each of the quantities above the intervals corresponding to minimum cost. If $K = k_S + n_S - 1$ and $u = K$, observe that $\epsilon > 0$ and the strict inequality also holds because $f_e^S > 0$. A better solution could be obtained by simply reducing x_e^S by ϵ. In any case, we arrive to a contradiction with the optimality of the original solution. ∎

B On the difference between integer and fractional solutions

In this section we analyze the difference between integer and fractional optimal solutions to Problem P for various instances of the shipping problem. We start by describing, in Sections B.1 and **??**, various distribution problems in which the linear programming relaxation is tight, thus showing that the proposed algorithm always finds the optimal solution in such cases. This is followed by Section **??** which illustrates, by constructing an example, that the relative error between fractional and integer solutions can be arbitrarily high.

B.1 Single-Period Two-Warehouse Model

Consider an instance of the *Shipper Problem*, Problem P, with a single period and where each of the shipments, from a set of suppliers to a number of retailers, must flow through one out of two existing warehouses.

Theorem B.1 *For the single-period, multiple-supplier, multiple-retailer, two-warehouse model, we have*

$$Z^* = Z^{\mathrm{LP}}.$$

Proof. Observe that, in this case, there are only two possible paths a commodity may take: either being shipped through warehouse 1 or through warehouse 2. Let y be an optimal vector of path flows for the linear programming relaxation of Problem P and let y_{ik}, $i = 1, 2$, $k \in \mathcal{K}$, denote the corresponding flow of commodity k on the path that uses warehouse i.

We decompose the solution vector y into at most $2K$ integer vectors of path flows, such that y is a convex combination of these new vectors. For this purpose we use the following procedure.

1. Initialize $t = 1$ and a vector of remaining flow R^t to $R^1_{ik} = y_{ik}$, for all $i = 1, 2$ and $k \in \mathcal{K}$.

2. Let $\beta_t = \min\{R^t_{ik} > 0 : \quad k \in \mathcal{K}, i = 1, 2\}$ and denote by i^* the warehouse (or any of the two if there is a tie) for which the minimum is achieved. Let $\mathcal{K}_{i^*} = \{k \in \mathcal{K} : \quad R^t_{i^*k} > 0\}$.

3. Construct an integer solution vector y^t consisting of sending all commodities in $\mathcal{K}_{i^*}$ through warehouse i^* and the rest through the other warehouse.That is,

   $$y^t_{ik} = \begin{cases} 1 & \text{if } i = i^* \text{ and } k \in \mathcal{K}_{i^*}, \text{ or } i \neq i^* \text{ and } k \in \mathcal{K}\backslash\mathcal{K}_{i^*} \\ 0 & \text{otherwise.} \end{cases}$$

   Let Z_t be the cost associated with this solution. y^t is the tth integer solution vector and, in the decomposition, exactly a fraction β_t of each commodity is shipped using the paths in this solution.

4. Update the vector of remaining path flows:

   $$R^{t+1} = R^t - \beta_t y^t,$$

   and let $t = t + 1$.

5. Repeat the procedure until all remaining path flows are zero.

Observe that $y = \sum_t \beta_t y^t$, $\sum_t \beta_t = 1$, and using Lemma 3.7 we can show that $Z^{\mathrm{LP}} = \sum_t \beta_t Z_t$ since:

1. For all $k \in \mathcal{K}$, $y_{1k} + y_{2k} = 1$ must hold. Thus, at each step t, the remaining flows must satisfy $R^t_{1k} + R^t_{2k} = 1 - \sum_{u=1}^{t-1} \beta_u$.

2. In each iteration of our procedure, if a commodity k is sent on an edge e, all the commodities $k' \in \mathcal{K}$ with $\gamma_{ek'} \geq \gamma_{ek}$ (proportions corresponding to the original solution to the linear program) are also shipped. In particular, solution y^t consists on shipping all the commodities with positive remaining flow, i.e. $R^t_{i^*k} > 0$, through warehouse i^* and all commodities with $R^t_{\bar{i}^*k} = 1 - \sum_{u=1}^{t-1} \beta_u$ through the other warehouse. Then, given an edge e and commodities indexed in the order of the original proportions, $\gamma_{e1} \leq \gamma_{e2} \leq \ldots \leq \gamma_{eK}$, the proportions corresponding to solution y^t can be written as

$$\gamma^t_{ek^t_e} = \gamma^t_{ek^t_e+1} = \ldots = \gamma^t_{eK} = 1,$$

   where k^t_e is the lowest index of the commodities with positive flow on edge e in the current solution y^t.

3. Given an edge e and using Lemma 3.7 together with the previous observation, we have

$$\sum_t \beta_t C^*_e(y^t) = \sum_t \beta_t F_e(\sum_{k=k^t_e}^{K} w_k) = \sum_{k=1}^{K} F_e(\sum_{i=k}^{K} w_i)[\gamma_{ek} - \gamma_{ek-1}] = C^*_e(y),$$

   where the notation is that presented in the previous observation. Note that each $(\gamma_{ek} - \gamma_{ek-1})$ must be either equal to β_t for a certain t or the sum of several β_t's.

Thus, all the integer solutions y^t must be optimal in the linear programming relaxation of Problem P and the lemma follows. ∎

Observe that in the proof of the theorem we show that any optimal fractional vector of path flows can be written as a convex combination of integer path flows. In addition, using Lemma A.1 and the way the integer solutions y^t are constructed, we can show that any fractional optimal vector (y, x, z) can be written as a convex combination of integer vectors. Thus,

Lemma B.2 *In the model described in Theorem B.1, all extreme point solutions to the linear programming relaxation of Problem P are integer.*

Proof.

Consider an optimal fractional solution (y, x, z) and the corresponding integer path flows y^t constructed as in the proof of Theorem B.1. In what follows, using the characterization in Lemma A.1, we construct associated integer interval and quantity variables (x^t, z^t), such that $x = \sum_t \beta_t x^t$ and $z = \sum_t \beta_t z^t$.

Given an edge e, index the commodities in a non-decreasing order of the proportions: $\gamma_{e1} \leq \gamma_{e2} \leq \ldots \leq \gamma_{eK}$. Recall that, since the ordering of the commodities has not been altered when constructing the integer solutions, at any step t in the procedure, the proportions corresponding to solution y^t can be written as

$$\gamma^t_{ek^t_e} = \gamma^t_{ek^t_e+1} = \ldots = \gamma^t_{eK} = 1,$$

where k^t_e is the lowest index of the commodities with positive flow on edge e in the current solution y^t.

If the solutions (x^t, z^t) associated with the path flows y^t are unique, i.e., for all edges $e \in SE$, $f^1_e > 0$ and $\sum_{i=k^t_e}^{K} w_i \neq M^r_e$ for $r = 1, 2, \ldots, R-1$, then the associated interval variables must be

$$(x^t)^r_e = \begin{cases} 1 & \text{if } \sum_{i=k^t_e}^{K} w_i \in (M^{r-1}_e, M^r_e], \\ 0 & \text{otherwise.} \end{cases}$$

The quantity variables z are also uniquely determined by observation 3.3,

$$(z^t)^r_{ek} = \begin{cases} w_k & \text{if } k \geq k^t_e \text{ and } \sum_{i=k^t_e}^{K} w_i \in (M^{r-1}_e, M^r_e], \\ 0 & \text{otherwise.} \end{cases}$$

Thus, making use of the expression of the interval variables x^r_e presented in Lemma A.1, we have for each given edge e and range $r = 1, 2, \ldots, R$,

$$\sum_t \beta_t (x^t)^r_e = \sum_{k^t_e \in \mathcal{K}(r)} \beta_t = \sum_{k \in \mathcal{K}(r)} (\gamma_{ek} - \gamma_{ek-1}) = x^r_e.$$

Similarly, using the expression of the quantity variables in term of the interval variables presented in Lemma A.1 it is easy to see that $z = \sum_t \beta_t z^t$, which concludes our proof in this case.

If the solutions (x^t, z^t) are not unique, we pick the values that match the corresponding variables in the original solution (y, x, z). For that purpose, we need to consider two cases:

- *Case 1:* $f_e^1 = 0$

  In this case, any value $x_e^1 \geq \sum_{k\in\mathcal{K}(1)}(\gamma_{ek} - \gamma_{ek-1})$ is feasible and leads to the same objective cost. We can write x_e^1 as a convex combination of $\bar{x}_e^1 := \sum_{k\in\mathcal{K}(1)}(\gamma_{ek} - \gamma_{ek-1})$ and some integer number greater than x_e^1, which we call $\bar{\bar{x}}_e^1$. Consider corresponding solutions $(y, \bar{x}, z)$ and $(y, \bar{\bar{x}}, z)$, identical to (y, x, z) in all components except x_e^1. Clearly, (y, x, z) is a convex combination of $(y, \bar{x}, z)$ and $(y, \bar{\bar{x}}, z)$ and it is enough to show that these two solutions are a convex combination of integer solutions. For that purpose, when decomposing these solutions, we assign values as follows. For the former,

  $$(\bar{x}^t)_e^1 = \begin{cases} 1 & \text{if } \sum_{i=k_e^t}^K w_i \in (M_e^0, M_e^1], \\ 0 & \text{otherwise.} \end{cases}$$

  For the latter, we just take $(\bar{\bar{x}}^t)_e^1 = \bar{\bar{x}}_e^1$, for all t.

- *Case 2:* For a certain step t and range r, $r = 1, 2, \ldots, R-1$, $\sum_{i=k_e^t}^K w_i = M_e^r$.

  That is, the quantity $\sum_{i=k_e^t}^K w_i$ is exactly at a breakpoint of the cost function. Thus, in the original fractional solution, commodity k_e^t is assigned either to $\mathcal{K}(r)$ or to $\mathcal{K}(r+1)$ or some fraction to each. In terms of the variables this means that any values of x_e^r and x_e^{r+1} such that $\sum_{k\in\mathcal{K}(r)\setminus\{k_e^t\}}(\gamma_{ek} - \gamma_{ek-1}) \leq x_e^r \leq \sum_{k\in\mathcal{K}(r)}(\gamma_{ek} - \gamma_{ek-1})$ and $x_e^{r+1} = \sum_{k\in\mathcal{K}(r)\cup\mathcal{K}(r+1)}(\gamma_{ek} - \gamma_{ek-1}) - x_e^r$, provide an optimal solution.

  We only consider the extreme cases:

  1. $x_e^r = \sum_{k\in\mathcal{K}(r)}(\gamma_{ek} - \gamma_{ek-1})$, in which case we set $(x^t)_e^r = 1$ and $(x^t)_e^{r+1} = 0$
  2. $x_e^r = \sum_{k\in\mathcal{K}(r)\setminus\{k_e^t\}}(\gamma_{ek} - \gamma_{ek-1})$, in which case we set $(x^t)_e^r = 0$ and $(x^t)_e^{r+1} = 1$

  For other values of x_e^r, the solution can be obtained as a convex combination of these extreme-case solutions.

Once these values are set, it is easy to see that the argument presented for unique solutions follows in these cases as well. ∎

Observe that when the vector of path flows y is integer, all the proportions $\gamma_{ek} \in \{0, 1\}$, for $e \in SE$, $k \in \mathcal{K}$. In that case, the above lemma implies that the interval variables, x, associated with these path flows are also integer, either 0 or 1. This is in agreement with Property 3.2 above.

A natural question is whether or not these results can be generalized to the case with more than two warehouses. Unfortunately, one can construct an example with three warehouses for which the ratio $Z^*/Z^{\mathrm{LP}} = 16/15$, while Z^*/Z^{LP} approaches $3/2$ as the number of warehouses increases, see Muriel (1997). If, however, the distribution system includes distribution centers supplying warehouses which serve the retailers, that is, a four-stage system (suppliers, distribution centers, warehouses and retailers), then the gap between Z^* and Z^{LP} can be larger. Indeed, Muriel (1997) presents an example for which Z^*/Z^{LP} approaches two as the number of facilities increases.