Stochastic Transportation-Inventory Network Design Problem

Shu Jia, Chung-Piaw Teo and Zuo-Jun Max Shen

Abstract— **In this paper, we study the stochastic transportation-inventory network design problem involving one supplier and multiple retailers. Each retailer faces some uncertain demand. Due to this uncertainty, some amount of safety stock must be maintained to achieve suitable service levels. However, risk-pooling benefits may be achieved by allowing some retailers to serve as distribution centers (and therefore inventory storage locations) for other retailers. The problem is to determine which retailers should serve as distribution centers and how to allocate the other retailers to the distribution centers. Shen et al. (2000) and Daskin et al. (2001) formulated this problem as a set-covering integer-programming model. The pricing subproblem that arises from the column generation algorithm gives rise to a new class of submodular function minimization problem. They only provided efficient algorithms for two special cases, and assort to ellipsoid method to solve the general pricing problem, which run in $O(n^7 log(n))$ time, where n is the number of retailers. In this paper, we show that by exploiting the special structures of the pricing problem, we can solve it in $O(n^2 \log n)$ time. Our approach implicitly utilizes the fact that the set of all lines in 2-D plane has low VC-dimension. Computational results show that moderate size transportation-inventory network design problem can be solved efficiently via this approach.**

I. Introduction

Managing inventory has become a major challenge for companies as they simultaneously try to reduce costs and improve service levels in today's increasing competitive market. Managing inventory consists of two critical tasks. First, we must determine the optimum number and location of distribution centers. Second, we must determine the amount of inventory to maintain at each of the distribution centers. Often these tasks are undertaken separately, resulting in a degree of suboptimization.

We study the design of a stochastic distribution network in which a single supplier ships products to a set of distribution centers (DCs). Each DC serves a pool of retailers with uncertain customer demand. The number and locations of DCs are not given a priori. They are chosen from the set of retailers. After being chosen as a DC, this retailer-based DC is served directly by the supplier and distributes products to some of the other retailers. The central issue in the stochastic distribution network design problem is how many and which retailers should be selected to be the DCs, how to assign the other retailers to the DCs, and how to manage the inventory at each distribution center.

The model we presented above was motivated by a study at a Chicago-based blood bank conducted by Shen et al. (2000) and Daskin et. al. (2001). "The blood bank supplied roughly 30 hospitals in the greater Chicago area. Its focus was on the production and distribution of platelets, the most expensive and most perishable of all blood products. If a unit of platelets is not used within 5 days of the time it is produced from whole blood, it must be destroyed. The demand for platelets is highly variable as they are needed in only a limited number of medical contexts. When they are used, however, multiple units are often needed. The hospitals supplied by the blood bank collectively owned the blood bank and set prices. As a result they could return a unit of platelets up to the time it becomes outdated and not be charged for it. Thus, there was little incentive to manage inventories in an efficient manner. Many of the larger hospitals ordered almost twice the number of platelet units that they used each year resulting in the need to destroy thousands of units of this expensive blood product. Other hospitals ordered almost all of their needed platelets on a STAT or emergency basis. The blood bank often had to ship the units to these hospitals using a taxi or express courier at significant expense to the network system. Clearly an improved system was needed. The idea was to establish regional centers at which platelets would be stored at a collection of nearby hospitals. By storing platelets at regional centers (located at a subset of the hospitals) and distributing platelets to nearby hospitals on an as-needed or daily basis, three objective were likely to be achievable. First, and the most importantly, we could use risk-pooling principles to reduce the necessary safety stock needed to protect against shortages. Second, the cost of emergency shipments could be reduced since platelets would be stored closer to each of the hospitals. Finally, the training for inventory managers in an improved network system would be simpler and more cost-effective since fewer individuals would be involved as the inventory would be maintained at a small number of regional distribution centers instead of being maintained at each individual hospital."

II. Literature Review

We need to consider both the strategic location costs and the operational inventory and transportation related costs in the integrated supply chain network design problem. Traditionally, the inventory theory research (see for example, Graves et. al (1993), Nahmias (1997), and Zipkin (1997) for a review) treat the number and locations of DCs as given. Actually, a lot of research in this field focus on single location or two echelon system with a fixed sin-

Shu Jia, High Performance Computation for Engineered Systems, Singapore-MIT Alliance.

Chung-Piaw Teo, Department of Decision Sciences, National University of Singapore.

Zuo-Jun Max Shen, Department of Industrial and Systems Engineering, University of Florida.

gle warehouse serving a number of retailers. The goal of this stream of research is to develop and evaluate inventory policies so as to minimize the inventory related costs while meet some service level standards. On the other hand, the literature on facility location theory (see for example, Daskin (1995), Mirchandani and Francis (1990), Drezner (1995), and Geoffrion and Power (1995) for a review) usually ignore or simplify inventory related costs while focus on the fixed facility location costs and transportation costs. The objective of location analysis is to decide the number and location of the DCs and the DC-retailer assignments.

Eppen (1979) showed that significant inventory-cost savings can be achieved by grouping retailers, and thus capitalizing on the so-called "risk pooling effects." Location issue is therefore an important factor in the overall performance of inventory system. Barahona and Jensen (1996) studies a version of the distribution network design problem for computer spare parts. Their model takes into account the costs of building the DCs and maintaining inventories at the various locations. To make the model tractable, they imposed very restrictive assumptions on the inventory costs. Teo et. al. (2001) studied the impact on inventory costs with consolidation of distribution centers. They design an algorithm that solves for a distribution system with the total fixed facility location cost and inventory costs within $\sqrt{2}$ of the optimum. They ignore the transportation costs in their model. Erlebacher and Meller (2000) formulate a joint location-inventory model with highly non-linear integer objective functions. Continuous approximation and some other heuristics are used to solve the problem. For a 600-node problem, it took 117 hours on a Sun Ultra Sparcstation.

Finally, Shen (2000), Shen et. al. (2000), and Daskin et al. (2001) studied the model presented in the next section. They were able to solve the pricing problem efficiently (in time $O(nlog(n))$ for two special cases: when the variance of the demand is proportional to the mean (as in the poisson demand case) or when the demand is deterministic, using the column generation framework. Although they prove that the general pricing problem is a submodular function minimization problem, which can be solved in polynomial time (for example, Grotschel et. al. (1981), Schrijver (1999), and Iwata et. al. (1999)). Preliminary computational evidence shows that these algorithms ($O(n^7 \log n)$) are still not computationally efficient. We propose a much faster algorithm ($O(n^2 \log n)$) to generate columns for the general case, using ideas from Chakravarty et al. (1985) and computational geometry.

Using an advance incremental algorithm for enumeration of vertices over a zonotope (see Onn and Schulman (2001), where the column generation problem in this paper is a special case), the running time complexity can be further slashed to $O(n^2)$. However, the reduction in running time comes at the expense of more complicated data structure to implement the incremental algorithm. Instead, we show that the variable fixing idea proposed by Daskin et al. (2001) in their lagrangian approach can be extended to the column generation framework. Using this approach, we are able to slashed the running time by a factor of 10 and solve a moderate size transportation-inventory network design problem (up to 120 retailers) in under 1 minute.

III. Model Formulation

For ease of exposition and for the completeness of this section, we introduce the model proposed by Shen et al. (2000) and Daskin et al. (2001) in this section. The readers may want to refer to the original papers for detailed derivation of the model.

Given a set I of retailers, there is a fixed cost f_j of locating a DC at retailer j. Two different types of inventories are kept at each DC: the working inventory, which is determined by the inventory ordering policy adopted, and the safety stock, which is kept at each DC to protect against the possibilities of running out of stocks during replenishment leadtime. We assume each DC orders inventory from the supplier using an economic order quantity model (EOQ). To determine the optimal reorder interval and the order quantity at each DC is not a straight forward task, since the corresponding order frequency and order quantity are determined by the mean demand served by the DC which is a function of the assignment of retailers to the DC. Other cost terms include the transportation costs from each DC to the retailers it serves, which are also dependent on the decisions of retailer assignments. The objective of this network design problem is the following: for a given set I of retailers, each facing independent uncertain demand, decide how many distribution centers to set up, where to locate them, which retailers to assign to each distribution center, how often to reorder at the distribution centers and what level of safety stock to maintain, to minimize the total facility location, shipment, working inventory and safety stock inventory costs.

Following another assumption made in Shen et. al. (2000), we assume that the non-DC retailers maintain only a minimal amount of inventory, and we therefore ignore this inventory in the model below.

To model this problem, we define the following notation:

Inputs and Parameters

- μ_i: mean (yearly) demand at retailer i, for each $i \in I$
- σ_i^2: variance of (daily) demand at retailer i, for each $i \in I$
- f_j: fixed (annual) cost of locating a regional distribution center at retailer j, for each $j \in I$
- d_{ij}: cost per unit to ship from retailer j to retailer i, for each $i \in I$ and $j \in I$
- α: desired percentage of retailers orders satisfied (fill rate)
- β: weight factor associated with the shipment cost
- θ: weight factor associated with the inventory cost
- z_α: standard normal deviate such that $P(z \leq z_\alpha) = \alpha$
- h: inventory holding cost per unit of product per year
- F_j: fixed cost of placing an order at distribution center j, for each $j \in I$
- L: lead time in days

• g_j: fixed shipment cost from external supplier to distribution center j
• a_j: per unit shipment cost from external supplier to distribution center j

Note that to simplify the notation, we have assumed that all lead times are equal.

Decision Variables

• $X_j = 1$, if retailer j is selected as a distribution center location, and 0, otherwise, for each $j \in I$
• $Y_{ij} = 1$, if retailer i is in served by a distribution center based at retailer j, and 0, otherwise, for each $i \in I$ and each $j \in I$

The model can now be formulated as follows:

$$\begin{aligned} \min \quad & \sum_{j\in I}\Bigg(f_j X_j + [\textstyle\sum_{i\in I}(\beta\mu_i d_{ij} + \beta a_j \mu_i) Y_{ij}] \\ & + \sqrt{2\theta h(F_j+\beta g_j)}\sqrt{\textstyle\sum_{i\in I}\mu_i Y_{ij}} \\ & + \theta h z_\alpha \sqrt{L}\sqrt{\textstyle\sum_{i\in I}\sigma_i^2 Y_{ij}}\Bigg) \\ & \equiv \sum_{j\in I}\Bigg(f_j X_j + (\textstyle\sum_{i\in I}\hat{d}_{ij}Y_{ij}) \\ & + K_j\sqrt{\textstyle\sum_{i\in I}\mu_i Y_{ij}} + q\sqrt{\textstyle\sum_{i\in I}\sigma_i^2 Y_{ij}}\Bigg) \end{aligned} \tag{1}$$

$$\sum_{j\in I} Y_{ij} = 1, \qquad \text{for each } i \in I \tag{2}$$

$$Y_{ij} - X_j \le 0, \qquad \text{for each } i, j \in I \tag{3}$$

$$Y_{ij} \in \{0,1\}, \qquad \text{for each } i, j \in I \tag{4}$$

$$X_j \in \{0,1\}, \qquad \text{for each } j \in I \tag{5}$$

where

$$\begin{aligned} \hat{d}_{ij} &= \beta\mu_i(d_{ij}+a_j), \\ K_j &= \sqrt{2\theta h(F_j+\beta g_j)}, \\ q &= \theta h z_\alpha \sqrt{L}. \end{aligned}$$

The objective function minimizes the weighted sum of the following four cost components:
• The fixed cost of locating facilities, given by the term $\sum_j f_j X_j$.
• The annual shipment cost from the distribution centers to the non-DC retailers, given by the term $\beta\Bigg(\sum_{i\in I}(\mu_i d_{ij} + a_j\mu_i)Y_{ij}\Bigg)$.
• The expected working inventory cost, given the solution to the EOQ equation with ordering cost $F_j + \beta g_j$, holding cost θh and demand $\sum_{i\in I}\mu_i Y_{ij}$.
• The annual safety stock cost, given by $\theta h z_\alpha \sqrt{L}\sqrt{\sum_{i\in I}\sigma_i^2 Y_{ij}}$.

The first two terms are structurally identical to those of the uncapacitated facility model. The last two terms are related to inventory costs, which are non-linear in the assignment variables. The constraints of the model are identical to those of the uncapacitated facility location problem, thus the problem we are studying is more difficult than the standard uncapacitated facility location problem, which is already a notorious NP-hard problem. We next formulate our decision problem as a set-covering model, and we present a column generation based approach to solve this model.

Note that every feasible solution to our decision problem consists of a partition of the set I of retailers into nonempty subsets, $R_1, R_2, \ldots, R_n$, together with one designated retailer for each of these n sets.

Let $\mathcal{R}$ be the collection of all nonempty subsets of the set I. For each set $R \in \mathcal{R}$, and each member $j \in R$, let $c_{R,j}$ be the total cost associated with set R with j as the DC. That is,

$$c_{R,j} = f_j + \sum_{i\in R}\hat{d}_{ij} + K_j\sqrt{\sum_{i\in R}\mu_i} + q\sqrt{\sum_{i\in R}\sigma_i^2}$$

By switching DC location from j to another retailer $i \in R$, $i \neq j$, we may get a different total cost associated with set R. Now we define c_R to be the lowest cost of having one distribution center serve exactly the set R. That is, $c_R = \min_{j\in R} c_{R,j}$.

Note that here we assume a distribution center always serves itself. This may not be the case for the optimal solution to our decision problem. See Shen et. al. (2000) for an example in which some retailer is chosen as a DC location in the optimal solution but does not serve itself. However, the method we propose can be modified in a straightforward way to allow for this possibility.

Let $z_{R,j} = 1$ if retailer j is used to serve the set of retailers in R. Note that R has to contain j by our assumption. The set covering model for the network design problem can now be formulated as:

$$\begin{array}{cc} \min & \sum_{R\in\mathcal{R}}\sum_{j\in R} c_{R,j} Z_{R,j} \\ \text{subject to} & \sum_{R\in\mathcal{R}: i\in R}\Bigg(\sum_{j\in R} Z_{R,j}\Bigg) \ge 1, \quad \forall i \in I, \\ & Z_{R,j} \in \{0,1\}, \quad \forall R \in \mathcal{R}. \end{array}$$

Note that in the above formulation, we can group the variable $\sum_{j\in R} Z_{R,j}$ together and replaced by Z_R, and replace the cost term $\sum_{j\in R} c_{R,j}Z_{R,j}$ by $c_R z_R$ to obtain a simplified model. The new set-covering model has one variable for each set $R \in \mathcal{R}$: $Z_R = 1$ if set R is served together in the solution, and 0, otherwise, for each $R \in \mathcal{R}$.

Now the model, which we will call $\mathcal{M}_{\mathcal{R}}$, can be expressed as follows:

$$\begin{array}{cc} \min & \sum_{R\in\mathcal{R}} c_R Z_R \\ \text{subject to} & \sum_{R\in\mathcal{R}: i\in R} Z_R \ge 1, \quad \forall i \in I, \\ & Z_R \in \{0,1\}, \quad \forall R \in \mathcal{R}. \end{array}$$

We begin each iteration by solving the linear relaxation of the above set-covering model, obtaining an optimal solution $\bar{Z}_R$, $R \in \mathcal{R}$, and the corresponding optimal dual solution $\bar{\pi}_i$, $i \in I$.

We want to know, for each column R, whether the reduced cost

$$c_R - \sum_{i\in R} \bar{\pi}_i \geq 0,$$

is non-negative for each $R \in \mathcal{R}$. If the answer is yes, then $\bar{Z}$ is an optimal solution to $\mathcal{M}_{\mathcal{R}}$. If, on the other hand, a set R with negative reduced cost is found, then R is added to $\mathcal{R}'$, and the next iteration begins.

Finding $R \subset \mathcal{R}$ with negative reduced cost, or proving that no such R exists, is called the *pricing problem.*

Thus, the pricing problem reduces to finding R_j^*, for each $j \in I$. To find R_j^* we must solve the following integer programming problem, $\mathcal{P}_j$:

$$\begin{array}{rc}
\min & f_j + \sum_{i\in I}(\hat{d}_{ij} - \bar{\pi}_i)Y_{ij} \\
 & +K_j\sqrt{\sum_{i\in I}\mu_i Y_{ij}} + q\sqrt{\sum_{i\in I}\sigma_i^2 Y_{ij}} \\
\text{subject to} & Y_{ij} \in \{0,1\}, \forall i \in I \\
 & Y_{j,j} = 1.
\end{array}$$

Given an optimal solution Y^* to $\mathcal{P}_j$, the set R_j^* is then the set $\{i \in I : Y_{ij}^* = 1\}$.

IV. The Pricing Problem

In this section we propose an algorithm to solve the pricing problem $\mathcal{P}_j$. To simplify the notation, we define

$$\begin{array}{rcl}
a_i & := & \hat{d}_{ij} - \bar{\pi}_i, \\
b_i & := & K_j^2 \mu_i, \\
c_i & := & q^2\sigma_i^2, \\
z_i & := & Y_{ij},
\end{array}$$

for each $i \in I$. Note that f_j does not depend on Y_{ij} and hence can be ignored for discussion here. We now have the following problem $\mathcal{P}_j$, for designated distribution center $j \in I$:

$$\begin{array}{rc}
\min & \sum_{i\in I} a_i z_i + \sqrt{\sum_{i\in I} b_i z_i} + \sqrt{\sum_{i\in I} c_i z_i} \\
\text{subject to} & z_i \in \{0,1\}, \forall i \in I \\
 & z_j = 1.
\end{array}$$

For each $j \in I$, define set function g_j on $E_j \equiv I\backslash\{j\}$ as follows. For each $S \subseteq E_j$,

$$g_j(S) \equiv a_j + \sum_{i\in S} a_i + \sqrt{b_j + \sum_{i\in S} b_i} + \sqrt{c_j + \sum_{i\in S} c_i}. \quad (6)$$

A. Solving the Pricing Problem

Lemma 1: Given a retailer $j \in I$, and associated minimum-reduced-cost set $R_j^* \subset I$. For every $i \in R_j^*\backslash\{j\}$, $a_i < 0$.

Hence we may restrict our search for R_j^* to retailers in I^-, where $I^- \equiv \{i \in I\backslash\{j\} : a_i < 0\}$. We next identify a nice structural property of the set R_j^* by extending an argument in Chakravarty et al. (1985).

Let $a_S = \sum_{i\in S} a_i$, $b_S = \sum_{i\in S} b_i$ and $c_S = \sum_{i\in S} c_i$. Define a new function

$$h_j(x,y,z) := (a_j + x) + \sqrt{b_j + y} + \sqrt{c_j + z}. \quad (7)$$

Note that $h(x,y,z)$ is a separable concave function, and

$$\min\nolimits_{S\subseteq I^-} g_j(S) \;=\; \min\nolimits_{S\subseteq I^-} h_j(a_S, b_S, c_S) \quad (8)$$

Since the set of ordered pairs $\{(a_S, b_S, c_S) : S \subseteq I^-\}$ is finite, its convex hull, which will be denoted by H, is a convex polyhedron. It now follows from (6) that

$$\begin{array}{rcl}
\min_{S\subseteq I^-} g_j(S) & = & \min_{S\subseteq I^-} h_j(a_S, b_S, c_S) \\
 & \geq & \min_{(A,B,C)\in H} h_j(A,B,C).
\end{array}$$

Since the function $h_j(A,B,C)$ is concave in the variables (A,B,C), the latter minimization problem attains a minimum at an extreme point of H.

Let (A^*, B^*, C^*) be an extreme point of H. Since H is a polyhedron, it is well known that there exists a linear function f on H that attains its *unique* minimum over H at (A^*, B^*, C^*). Since f is linear, it has a representation $f(A,B,C) = \alpha A + \beta B + \gamma C$ defined by real numbers α, β and γ. The uniqueness of (A^*, B^*, C^*) as the minimizer of f over H assures that we do not have $\alpha = \beta = \gamma = 0$.

Since H is the convex hull of $\{(a_S, b_S, c_S) : S \subseteq I^-\}$,

$$\begin{array}{rcl}
\alpha A^* + \beta B^* + \gamma C^* & = & \min_{(A,B,C)\in H} \alpha A + \beta B + \gamma C \\
 & = & \min_{S\subseteq I^-} \alpha a_S + \beta b_S + \gamma c_S \\
 & = & \min_{S\subseteq I^-} \sum_{i\in S}(\alpha a_i + \beta b_i + \gamma c_i).
\end{array}$$

The set $S^* = \{i : \alpha a_i + \beta b_i + \gamma c_i < 0\}$ is clearly optimal for the last optimization problem. Hence, we conclude from the uniqueness of (A^*, B^*, C^*) as the minimizer of f over H, that $(A^*, B^*, C^*) = (a_{S^*}, b_{S^*}, c_{S^*})$. Furthermore, $R_j^* = S^*$.

Note that $S^* = \{i : \alpha a_i + \beta b_i + \gamma c_i < 0\} = \{i : \alpha + \beta\frac{b_i}{a_i} + \gamma\frac{c_i}{a_i} > 0\} = \{i : \beta x_i + \gamma y_i < \alpha\}$, where $x_i = -b_i/a_i$ and $y_i = -c_i/a_i$. Note that $x_i, y_i \geq 0$ for all i.

Although there are infinitely many choices for the parameters α, β and γ, it turns out that the number of distinct partitions obtained by varying the parameters are limited. This follows from a general result in the theory of VC-dimension. To describe this result, we need to first introduce some notations.

The *VC dimension* is defined for any set system $\mathcal{S} \subset 2^X$ on an arbitrary set X. It is the supremum of the sizes of all shattered subsets $\mathcal{A} \subset X$; here $\mathcal{A}$ is called shattered if $\mathcal{S}|_{\mathcal{A}} = 2^{\mathcal{A}}$, i.e. for any $\mathcal{B} \subset \mathcal{A}$ there exists a set $S \in \mathcal{S}$ such that $\mathcal{B} = \mathcal{A} \cap \mathcal{S}$. For example, if $\mathcal{H}$ denote the system of all closed halfplanes in the plane, then it is not difficult to check that the VC-dimension of the set system $\mathcal{H}$ is 3, since no 4 points in the plane can be shattered by using only halfplanes.

The following well-known result shows that the number of possible candidates for S^* are essentially small:

Lemma 2 (Vapnik and Chervonenkis (1971); Sauer (1972)) For any set system $\mathcal{S}$ of VC dimension at most d, we have $\left|\mathcal{S}|_X\right| \leq \Phi_d(|X|)$, where $\Phi_d(m) = \begin{pmatrix} m \\ 0 \end{pmatrix} + \begin{pmatrix} m \\ 1 \end{pmatrix} + \ldots + \begin{pmatrix} m \\ d \end{pmatrix}$.

The above lemma suggests that we need to search among at most $O(n^3)$ possible subsets to determine S^*.

B. Solving the Pricing Problem: Dual Approach

Using point-line duality, we can cast the above algorithm in a different way. To describe this dual approach, we first observed that at the optimal set S^*, the parameters α, β, γ satisfy the additional properties:

- $\alpha = 1$,
- $1/(2\sqrt{b_j + \sum_{i \in S} b_i}) \le \beta \le 1/(2\sqrt{b_j})$, and
- $1/(2\sqrt{c_j + \sum_{i \in S} c_i}) \le \gamma \le 1/(2\sqrt{c_j})$.

The above follows from gradient conditions at the optimal solution, since the concave objective function is of the form $h(x, y, z) = x + \sqrt{y} + \sqrt{z}$ and $0 \le z_i \le 1$ for all i.

We know $S^* = \{i : \alpha a_i + \beta b_i + \gamma c_i < 0\} = \{i : \alpha + \beta\frac{b_i}{a_i} + \gamma\frac{c_i}{a_i} > 0\} = \{i : \beta x_i + \gamma y_i < \alpha\}$ for some choice of α, β and γ. Furthermore, the additional properties allow us to restrict our search to finding β and γ such that $S^* = \{i : \beta x_i + \gamma y_i < 1, \beta > 0, \gamma > 0\}$. The possible choices for β and γ now lie in the positive orthant. Furthermore, the inequality $\beta x_i + \gamma y_i < 1$ denotes a half-space in this region.

For each pair of i, j, we solve the equation

$$\begin{cases} \beta x_i + \gamma y_i = 1 \\ \beta x_j + \gamma y_j = 1 \end{cases}$$

This gives rise to solution $(\beta_{ij}, \gamma_{ij})$. We can discard the solution if any of the β_{ij}, γ_{ij} is non-negative.

In the following we use a three retailer example to illustrate how to get the optimal set S^* (cf. Figure 1).

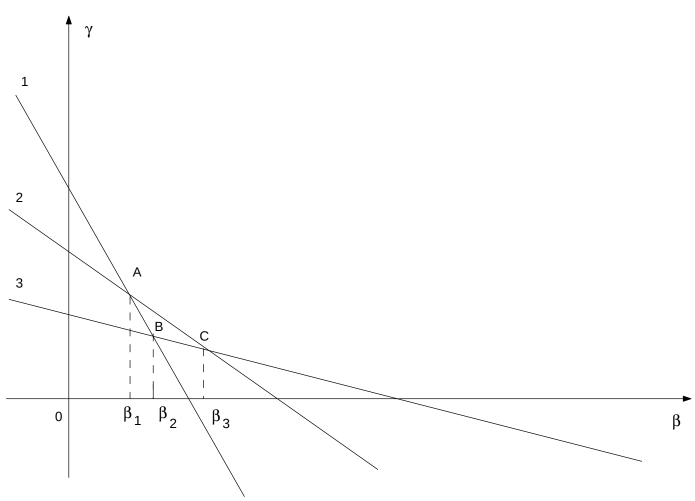

Fig. 1. Illustration of the dual algorithm

First sort all the intersection points according to the value of the β coordinates. For ease of exposition, we relabelled the points $(\beta_{ij}, \gamma_{ij})$ as (β_k, γ_k) so that $\beta_k \le \beta_{k+1}$ for all k, $k = 1, 2, \cdots, m$, and $m \le n^2$.

Note that when $\beta \in [0, \beta_1)$, the changes of possible candidates for S^* as γ varies follow an obvious pattern. In the above example, the possible candidates are $\{1, 2, 3\}$, $\{1, 2\}$, and $\{1\}$ (as γ increases). Similarly, the possible candidates for S^* are $\{2, 1, 3\}$,$\{2, 1\}$, and $\{2\}$ when $\beta \in [\beta_1, \beta_2)$; the possible candidates for S^* are $\{2, 3, 1\}$,$\{2, 3\}$ and $\{2\}$ when $\beta \in [\beta_2, \beta_3)$; the possible candidates for S^* when $\beta \in [\beta_3, \infty)$ are $\{3, 2\}$ and $\{3\}$ respectively.

More formally, the algorithm can be described as follows:

0 Given points (β_k, γ_k), $k = 1, 2, \ldots, m$, with $\beta_1 \le \beta_2 \le \ldots \le \beta_m$.
1 For each k in 0, 1, 2, …, m (define $\beta_0 = 0$),
a For each line i, let $\Gamma_i \equiv \frac{1 - \beta_k x_i}{y_i}$. i.e., when the β value is set at β_k.
b Sort the lines in non-decreasing value of Γ_i. WLOG, let $k_1 \le \ldots \le k_n$ denote the ordering of the lines when the β value is set at β_k.
c The candidate solutions are $\{k_j, k_{j+1}, \ldots, k_n\}$, for each j in $1, 2, \ldots, n$, provided $\Gamma_{k_j} \ge 0$.

Theorem 1: The problem $\min_{S \subset I^-} g_j(S)$ can be solved in $O(n^2 \log n)$ time.

V. Computational Results

In this section, we summarize our computational experience with the algorithms outlined in the previous section.

A. Stochastic Network Design Problem

In this subsection we report the results of solving the network design problem using column generation method. The algorithm for the general network distribution problem is coded in C++, and the linear programming problem is solved using CPLEX LP Solver.

We generate all the instances of the problem by varying the number of retailers and the values of β and γ. The mean demands μ_i and σ_i^2 are randomly generated in $[100, 1600]$ for all $i \in I$. Holding cost is 1, $z_\alpha = 1.96$ (97.5% service level), $a_i = 5$, $g_i = 10$, $F_i = 10$ for all $i \in I$. Our goal is to find ranges of values for β and θ that resulted in instances that varied in solution difficulty as well as the fraction of retailers used as distribution centers in the solution.

For each of the instances, we first solve the linear programming relaxation of the set-covering model via column generation. The initial set of columns include all singletons. The column labelled “No. of Columns Generated” indicates the total number of columns added during this phase. The resulting final optimal objective value is denoted by Z_{LP}. In most of all instances generated, the corresponding optimal solutions are integral. We denote by Z_H the best upper bound we obtained. In the case where the linear-programming relaxation solution is not integral, Z_H is obtained by applying an integer-programming solver to the final master problem. The column labelled ”No. of DCs Opened” indicates the number of sets R with value $Z_R^* = 1$ in the optimal linear programming solution. Tables 3 and 4 highlight the results of our computational study.

INPUT		OUTPUT			
β	θ	DCO	CT	NCG	z^H/z_{LP}
0.001	0.1	4	66	2793	1
0.002	0.1	7	29	879	1
0.003	0.1	8	24	792	1.001
0.004	0.1	10	16	483	1
0.005	0.1	13	13	382	1
0.001	0.1	4	66	2793	1
0.002	0.2	6	32	966	1.001
0.005	0.5	8	22	685	1
0.005	0.1	13	13	382	1
0.005	1	9	22	644	1.004
0.005	5	5	49	1963	1

Table 1: Computational results for 40 retailers instance.

INPUT		OUTPUT			
β	θ	DCO	CT	NCG	z^H/z_{LP}
0.001	0.1	6	448	10943	1
0.002	0.1	8	233	6318	1
0.003	0.1	12	105	2607	1
0.004	0.1	21	54	1102	1
0.005	0.1	24	44	538	1
0.001	0.1	6	448	10943	1
0.002	0.2	13	101	2210	1
0.005	0.5	18	61	1324	1
0.005	0.1	24	44	538	1
0.005	1	10	117	3147	1
0.005	5	7	378	9112	1

Table 2: Computational results for 80 retailers instance.

INPUT		OUTPUT			
β	θ	DCO	CT	NCG	z^H/z_{LP}
0.0001	0.01	10	826	22143	1
0.0002	0.01	15	563	11766	1
0.0003	0.01	24	248	5104	1
0.0004	0.01	28	124	2138	1
0.0005	0.01	33	68	1187	1
0.0001	0.01	10	826	22143	1
0.0002	0.02	23	249	5145	1
0.0005	0.05	29	119	2015	1
0.0005	0.01	33	68	1187	1
0.0005	0.1	15	571	12026	1
0.0005	0.5	11	717	18522	1.002

Table 3: Computational results for 120 retailers instance.

DCO: No. of DCs opened;CT: CPU Time (in seconds);NCG: No. of columns Generated.

VI. Variable Fixing

In the straight forward implementation of the above algorithm, we need to solve, for each retailer, a related submodular function minimization problem where the retailer is assumed to be the DC. This slows down the column generation routine considerably. We show next how information on the primal and dual solution can be used to "fix" variables, so that we can determine whether a retailer will be a DC candidate in an optimal solution early in the column generation routine.

Recall that the set covering model we are trying to solve is of the form:

$$\begin{array}{ll} \min & \sum_{R\in\mathcal{R}}\sum_{j\in R} c_{R,j}Z_{R,j} \\ \text{subject to} & \sum_{R\in\mathcal{R}:i\in R}\left(\sum_{j\in R} Z_{R,j}\right)\geq 1, \quad \forall i\in I, \\ & Z_{R,j}\in\{0,1\}, \quad \forall R\in\mathcal{R}. \end{array}$$

At each stage of the column generation routine, we have:

- A set of dual prices $\{\pi_j\}$.
- A set of primal feasible (fractional) solution $z_{S,j}$. Note that the variable $z_{S,j}$ is only defined for S, j with $j \in S$, since only retailers within the set S can be the designated DC to serve S.
- After solving the pricing problem (one for each retailer), we obtain the reduced cost $r_j \equiv \min_{S:j\in S}\left(c_{S,j} - \sum_{k\in S}\lambda_k\right)$. Note that some of the r_j's may be nonnegative.

Let Z_{IP} and Z_{LP} denote the optimal integral and fractional solution to the set covering problem.

Claim 1: $\sum_{j:r_j\leq 0} r_j + \sum_j \pi_j$ is a lowerbound to Z_{LP}. Hence it is a lowerbound to Z_{IP} too.

Let j^* be a retailer such that $r^*_j > 0$. Let UB be an upperbound for Z_{IP}.

Claim 2: If $\sum_{j:r_j\leq 0} r_j + \sum_j \pi_j + r_{j^*} > UB$, then retailer j^* will never be used as a DC in the optimal solution to the (integral) set covering problem.

Note that once we determine that the retailer j^* will never be used as a DC in the optimal solution, then we do not need to solve the pricing problem corresponding to j^* anymore in the rest of the column generation procedure. In fact, all columns arising from using j^* as DC (generated previously) can also be deleted from the LP. This is the key advantage of the variable fixing method.

The variable fixing method depends largely on the quality of the upperbound UB. If $Z_{LP} = Z_{IP}$, then the solution $\sum_{S,j} c_{S,j}z_{S,j}$ generated at each stage of the column generation routine will be an upperbound to Z_{IP}. Unfortunately this is not true for all instances. As in Daskin et al. (2001), we generate an upperbound for the IP by generating a feasible solution in the following way:

- Let z^* be the optimal LP solution obtained by solving the problem using a partial set of columns.
- Order the retailers according to non-decreasing value of demand.
- Starting from the first retailer (say i) on the list, if for some S and j, $i \in S$ and $z^*_{S,j} = 1$, then retailer i is served by DC j. Otherwise, there exists S, T, both containing i, and j, k, such that $z^*_{S,j} > 0$, $z^*_{T,k} > 0$. We serve i using the DC that will lead to the least total cost, and remove retailer i from the list.
- Repeat the previous step until the list is empty.

In this way, we can generate a feasible solution to the distribution network design problem. This solution will be used as a bound to perform variable fixing in the column generation routine.

A. Stochastic Network Design with variable fixing

The column labelled "No. of DCs OUT" indicates the number of retailers ruled out from being possible DCs in the optimal linear programming solution by variable fixing technique. The parameters we use for the instances below are the same as what we use for the previous subsection. Tables 5 and 6 highlight the results of our computational study.

By applying variable fixing technique, we are able to cut down the computational time dramatically. The average CPU time is only about 9% of the CPU time without variable fixing technique. The savings range from about 83% to 94%. It is especially effective for those difficult instances that require the most CPU times before. For example, for $\beta = 0.001$, $\theta = 0.1$ in the 80-retailer case, we are able to solve the problem in about 20 seconds after applying variable fixing technique, which used to take 448 seconds.

INPUT		OUTPUT				
β	θ	DCO	DCT	CT	NCG	z^H/z_{LP}
0.001	0.1	4	35	4.12	132	1
0.002	0.1	7	31	2.58	87	1
0.003	0.1	8	31	2.01	74	1.001
0.004	0.1	10	29	1.64	59	1
0.005	0.1	13	27	0.82	39	1
0.001	0.1	4	35	4.12	132	1
0.002	0.2	6	33	3.47	106	1.001
0.005	0.5	8	32	1.92	70	1
0.005	0.1	13	27	0.82	39	1
0.005	1	9	29	1.80	66	1.003
0.005	5	5	34	3.89	117	1

Table 4: Computational results for 40 retailers instance.

INPUT		OUTPUT				
β	θ	DCO	DCT	CT	NCG	z^H/z_{LP}
0.001	0.1	6	74	20.73	382	1
0.002	0.1	8	72	12.86	224	1
0.003	0.1	12	66	10.46	171	1
0.004	0.1	21	58	6.86	101	1
0.005	0.1	24	56	4.83	73	1
0.001	0.1	6	74	20.73	382	1
0.002	0.2	13	66	9.54	160	1
0.005	0.5	18	62	8.02	129	1
0.005	0.1	24	56	4.83	73	1
0.005	1	10	69	11.24	188	1
0.005	5	7	73	17.63	314	1

Table 5:Computational results for 80 retailers instance.

INPUT		OUTPUT				
β	θ	DCO	DCT	CT	NCG	z^H/z_{LP}
0.0001	0.01	10	109	58.37	817	1
0.0002	0.01	15	105	38.06	493	1
0.0003	0.01	24	94	25.65	302	1
0.0004	0.01	28	91	18.71	191	1
0.0005	0.01	33	87	11.54	103	1
0.0001	0.01	10	109	58.37	817	1
0.0002	0.02	23	96	26.46	318	1
0.0005	0.05	29	90	17.85	175	1
0.0005	0.01	33	87	11.54	103	1
0.0005	0.1	15	104	39.11	502	1
0.0005	0.5	11	109	50.33	718	1.003

Table 6: Computational results for 120 retailers instance. DCO: No. of DCs opened;DCT: No. of Dcs out;CT: CPU Time (in seconds); NCG: No. of columns Generated.

VII. Extension and Generalization

A. More General Pricing Problems

Recall that the pricing problem we solved has the following formulation:

$$g_j(S) \equiv u_1(x) + f_1(u_2(y)) + f_2(u_3(z)) \quad (9)$$
$$= (a_j + x) + \sqrt{b_j + y} + \sqrt{c_j + z} \quad (10)$$

If the management decides to use some inventory model other than EOQ, as long as the corresponding pricing problem can be written as (10), then the approaches detailed above can be applied to solve the problem. For example, suppose they decided to use the (Q, r) policy, where an order of size Q is placed whenever the inventory level reaches (or goes below) r. If we specify that at DC j, the probability of not stocking out in the leadtime is, say α. Let S denote the set of retailers served by j which excludes j itself, let

$$D := \textstyle\sum_{i\in S} \mu_i + \mu_j,$$
$$\sigma := \sqrt{\textstyle\sum_{i\in S} \sigma_i^2 + \sigma_j}$$

It is well known that the optimal

$$Q = EOQ = \sqrt{2F_j D/h},$$

and the value of r can be calculated by

$$F(r) = \alpha,$$

where F is the distribution function of leadtime demand at DC j. If F follows Normal distribution with mean D and standard deviation of σ, then

$$r = D + z\sigma,$$

where z is a constant depends on α. For example, if $\alpha = 98\%$, then $z = 2.05$.

The total setup and inventory holding costs at DC j can be calculated as:

$$\begin{aligned} TC_j(S) &= F_j D/Q + h(Q/2 + r - D) \\ &= \sqrt{2F_j hD} + h\sigma \\ &= \sqrt{2F_j h(\textstyle\sum_{i\in S} \mu_i + \mu_j)} + h\sqrt{\textstyle\sum_{i\in S} \sigma_i^2 + \sigma_j}. \end{aligned}$$

Since $TC_j(S)$ is structurally identical to the corresponding inventory costs terms in (10), we can apply our algorithm to solve the corresponding pricing problem.

B. Distance Constraint

In the distribution network design problem, it is quite common to impose additional constraints on the collection of retailers a DC can serve. For example, a typical geographical constraint stipulates that the designated DC and the retailer cannot be too far apart. To enforce this constraint in our approach is easy, since we can set the distance function d_{ij} to a huge number if retailer j cannot act as the DC for retailer i (or vice versa).

C. Capacity Constraint

Another common constraint states that a DC cannot handle too many retailers (say not more than k retailers can be served by a single DC), due to capacity or other technical limitations. In this paper, we describe how our technique can be extended to handle the additional constraint of the type: $\sum_i Y_{i,j} \leq k+1$, for some fixed k.

In this case, the column generation phase reduces to solving a problem of the type:

$$\begin{array}{cc} \min & \sum_{i\in I} a_i z_i + \sqrt{\sum_{i\in I} b_i z_i} + \sqrt{\sum_{i\in I} c_i z_i} \\ \text{subject to} & z_i \in \{0,1\}, \forall i \in I \\ & z_j = 1. \\ & \sum_i z_i \leq k+1. \end{array}$$

Since the objective function is concave and separable, we can use the same argument to reduce the problem to a parametric version:

$$\begin{array}{cc} (PK): \quad \min_{S\subseteq I^-} & \sum_{i\in S}(\alpha a_i + \beta b_i + \gamma c_i) z_i \\ \text{subject to} & z_i \in \{0,1\}, \forall i \in I^- \\ & \sum_{i\in I^-} z_i \leq k. \end{array}$$

The candidate solution for the column generation phase comes out as the solution to the above linear discrete optimization problem for some choice of α, β and γ.

Let $b(\alpha,\beta,\gamma)$ denote the value of the kth smallest entry in the set

$$\{\alpha a_i + \beta b_i + \gamma c_i : i \in I^-\}.$$

It is clear that if $z_i = 1$ in an optimal solution to problem (PK), then clearly $\alpha a_i + \beta b_i + \gamma c_i \leq b(\alpha,\beta,\gamma)$ since $\alpha a_i + \beta b_i + \gamma c_i$ cannot be bigger than the k-th smallest value. Furthermore, we need $\alpha a_i + \beta b_i + \gamma c_i < 0$, otherwise we would have $z_i = 0$ in the optimal solution. Conversely, it is easy to see that $z_i = 1$ in the optimal solution if the point i satisfies both inequalities.

The inequality

$$\alpha a_i + \beta b_i + \gamma c_i < \min(b(\alpha,\beta,\gamma), 0)$$

determines a halfplane in 3D, and at most k out of possible $n-1$ points in the set

$$S \equiv \{(a_i, b_i, c_i) : i \in I^-\}$$

lies in this halfplane.

Hence the number of candidate solutions depends on the number of $\leq k$-set. Here, a $\leq k$-set is the intersection of S and a halfplane containing at most k points. Clarkson and Shor (1989) showed that the number of such solutions in 3D is bounded above by $O(nk^2)$.

VIII. Conclusion

In this paper, we have outlined a formulation of a stochastic transportation-inventory network design model. The model determines how many and where to locate regional distribution centers and how to assign retailers to the distribution centers to minimize the total system costs, which include distribution center location costs, inventory costs at the distribution centers, and the transportation costs within this two echelon supply chain.

The model was originally proposed in Shen et. al. (2000). They were able to solve efficiently only two special cases of the general model. We proposed two different algorithms, primal approach and dual approach, to solve the general model. Although both algorithms have a worst case running time of $O(n^2 logn)$, computational results suggest that the dual approach is indeed much faster than the primal approach. We then applied the variable fixing technique, which helped us to reduce the CPU time dramatically.

We would like to emphasize the importance of being able to solve the general supply chain design problem. The two cases considered in Shen et. al. (2000) require that the demand be either deterministic or $\frac{\sigma_i^2}{\mu_i} = \gamma$ for every retailer. However, in a lot of real life situations, the demand processes can be very different from retailer to retailer, and the ratio of demand variance to mean demand are not the same for different retailers. Supply chain network design problems under such conditions are the ones that the management is most concerned with, and our model can be applied successfully in the decision making process for problems of this kind.

A number of extensions were also discussed in the paper. First, we show that our solution techniques can be easily extended to more general pricing problems, which may come from more general inventory management models, or more general transportation cost structures. Second, we show that the results in this paper can be generalized to handle models with distance or some capacity constraints.

We propose two important related future research directions. First, we believe that the primal and dual approaches are still applicable if we use the (Q,r) model without service level constraints. That is, we wish to show that the optimal cost of a (Q,r) policy is concave in the average demand. Second, we hope to consider the cases with multiple items as well as more general capacity constraints. Finally, we want to work on the network design models with more realistic transportation cost structures.

References

[1] Barahona, F. and D. Jensen (1998), "Plant Location with Minimal Inventory", *Mathematical Programming* **83,** pp. 101-111.

[2] Chakravarty, A. K. , Orlin, J. B., and U. G. Rothblum (1985), "Consecutive Optimizers for a Partitioning Problem with Applications to Optimal Inventory Groupings for Joint Replenishment", *Operations Research* **33,** pp. 820-834.

[3] Clarkson, K. L. and P. W. Shor (1989), "Applications of random sampling in computational geometry. II", Discrete Comput. Geom. 4, no. 5, 387–421

[4] Daskin, M. S. (1995), "Network and Discrete Location: Models, Algorithms, and Applications" Wiley-Interscience, New York.

[5] Daskin, M. S., Collette R. Coullard and Z. J. Max Shen (2001), "An Inventory-Location Model: Formulation, Solution Algorithm and Computational Results", Preprint.

[6] Drezner, Z., editor (1995), "Facility Location: A Survey of Applications and Methods", Springer, New York.

[7] Eppen, G. (1979), "Effects on Centralization on Expected Costs in a Multi-Location Newsboy Problem", *Management Science* **83,** No. 5, pp. 498-501.

[8] Erlenkotter, S. J. and R. D. Meller (2000), "The Interaction of Location and Inventory in Designing Distribution Systems", *IIE Transactions*, **32**, pp. 155-166.

[9] Geoffrion, A. M. and R. Power (1995), "Twenty Years of Strategic Distribution System Design: an evolutionary Perspective", *Interfaces*, **25**(5), pp. 105-127.

[10] Grötschel, M., L. Lovasz, and A. Schrijver (1981), "The Ellipsoid Method and Its Consequences in Combinatorial Optimization", *Combinatorica* **1,** pp. 169-197.

[11] Graves, S. C., A. H. G. Rinnooy Kan and P. H. Zipkin (1993), " Logistics of Production and Inventory", Elsevier Science Publishers, Amsterdam, The Netherlands.

[12] Iwata, S., L. Fleischer, and S. Fujishige (1999), "A Combinatorial, Strongly Polynomial-Time Algorithm for Minimizing Submodular Functions", JACM.

[13] Macalalag, E. and M. Sniedovich (1996), "Generalized Linear Programming and Sensitivity Analysis Techniques", *Naval Research Logistics*, **43**, pp. 397-413.

[14] Mirchandani, P. B. and R. L. Francis (1990), "Discrete Location Theory", John Wiley & Sons.

[15] Nahmias, S. (1997), "Production and Operations Management", Third Edition, Irwin, Chicago.

[16] Onn S. and L. Schulman (2001), "The Vecotr Partition Problems for Convex Objective Functions", *Mathematics Of Operations Research*, **26**, pp. 583-590.

[17] Rado, R. (1942), "A Theorem on Independence Relations", *The Quarterly Journal of Mathematics Oxford* **213**, pp. 83-89.

[18] Sauer, N (1972), "On the Density of Families of Sets", *Journal of Combinatorial Theory Series A*, **1**, pp. 145-1472.

[19] Schrijver, A. (1999), "A Combinatorial Algorithm Minimizing Submodular Functions in Strongly Polynomial Time", Combinatorica.

[20] Shen, Z. J. Max (2000), "Approximation Algorithms for Various Supply Chain Problems", Ph.D Thesis, Dept. of Industrial Engineering and Management Sciences, Northwestern University.

[21] Shen, Z. J. Max, Collette Coullard, and Mark S. Daskin(2000), "A Joint Location-Inventory Model", to appear in *Transportation Science*.

[22] Teo, C. P., J. H. Ou and K. H. Goh (2001), "Impact on Inventory Costs with Consolidation of Distribution Centers", *IIE Transactions* **33**, Issue 2, pp. 99-110.

[23] Vapnik, V. N. and A. Ya. Chervonenkis (1971), "On the Uniform Convergence of Relative Frequencies of Events to Their Probabilities", *Theory Probab. Appl.*, **16** pp. 264-280.

[24] Zipkin, P. H. (1997), "Foundations of Inventory Management", Irwin, Burr Ridge, IL.