# A Dynamic and Stochastic Model for Distribution of Empty Containers

by

Qin Chu

Submitted to the Department of Ocean Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1994

© Massachusetts Institute of Technology 1994. All rights reserved.

Author ........................................................................
Department of Ocean Engineering
May 5, 1994

Certified by ........................................................................
Ernest G. Frankel
Professor of Marine Systems
Thesis Supervisor

Accepted by ........................................................................
Richard C. Larson
Co-Director, Operations Research Center

Accepted by ........................................................................
A. Douglas Carmichael
Chairman, Departmental Committee on Graduate Students

Eng.
WITHDRAWN FROM
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
MIT LIBRARIES
JUN 20 1994

# A Dynamic and Stochastic Model for Distribution of Empty Containers

by

Qin Chu

Submitted to the Department of Ocean Engineering
on May 5, 1994, in partial fulfillment of the
requirements for the degree of
Master of Science in Operations Research

## Abstract

The dynamic container allocation problem arises when a carrier need to manage the positioning of a fleet of containers over time to carry goods from shippers to receivers. It involves dispatching available empty containers to meet requests by shippers and redistribution of other empty containers to other depots or ports in anticipation of future demands. In this thesis, we define the container allocation problem as a general dynamic vehicle problem with leasing options, and study the basic structures and main characteristics of the problem. We then model the container allocation process by dynamic networks and propose dynamic deterministic and stochastic models for container allocation. This provides a general modeling framework for this class of problems. The scope of the models includes both land side component and sea side component of the container distribution process of a shipping company. Both types of models capture the time and space dependency of the allocation process, and the stochastic models also incorporate the uncertainty of future demands and supplies explicitly in the optimization process. The mathematical formulation of the stochastic model is a two stage stochastic program with recourse problem, and the stochastic quasigradient method, in particular, the stochastic linearization method is employed to to obtain approximate solutions to the model. Implementation issues of the stochastic linearization method are discussed, which include the choices of step directions and stepsize. Numerical experiments are conducted for both deterministic and stochastic models under a rolling horizon simulation procedure on randomly generated testing problems to evaluate the models. Performance measures are accumulated over the simulation procedure, and the results show that the stochastic model gives better solutions than the deterministic model for most of testing problems.

Thesis Supervisor: Ernest G. Frankel
Title: Professor of Marine Systems

# Acknowledgments

I would like to thank my advisor Ernest G. Frankel for his guidance through out the project that resulted in this thesis.

Funding for this thesis, provided by the Neptune Orient Lines of Singapore, is gratefully acknowledged.

I would like to extend my special thanks to Professor Warren B. Powell and Professor Raymond K.M. Cheung for their valuable advice.

I would like to thank my friends, especially Spyridon A. Maragos, Ronald Chu and Chengyany Fei for their support in many ways.

Finally, a very special thanks to my wife Changjing Xu for accompanying me through these meaningful and lovely years and years to come.

# List of Tables

# List of Figures

# Contents

# Chapter 1

# Introduction

As the use of containerized transportation practices increases, the multimodal transportation becomes more popular. Many freight transportation companys, such as railroads and maritime transportation companys, normally own a large fleet of containers and insure their multimodal transportation from the initial shipper to the final receivers.

In shipping market, a liner operator (a type of shipping companies) usually provides fixed and regular service. The operator first designs the service pattern and then releases them in the form of ship schedules to shippers. Shippers will select service suitable to them and sign a contract with the liner operator through a cargo broker. A large liner company may own more than fifty container ships, each carrying thousands of containers. A company's service may cover several trade regions, for example, the trans-Pacific trade between East coast of North America and Far East, and the trans-Atlantic trade between West coast of North America and West Europe. Within a trade region, the number of principal container ports and inland depots visible to the operator ranges from thirty to more than a hundred depending on the size of hinderland of the trade region. The liner operator usually signs contracts with railroad and trucking companies, who are responsible for the move of containers on land. The planning of routes is a long term strategic decisions. Once they are determined, they remain fixed for a period of time. At the daily operations, the main concern is how to allocate empty containers to meet customers' demands for empty containers needed for transportation at the fixed routes determined in strategic planning. In this thesis, we focus on how to construct a decision support model for the empty container allocation process.

The dynamic container allocation problem arises when the transportation company must manage the repositioning of a fleet of containers over time. Shippers might request for empty containers to fill with loads of freight so as to be transported to their destinations. The problem the transportation company faces is to ensure that both empty and loaded containers are in the right place at the right time, given the forecast of shippers' demand. If their own containers are not enough to satisfy the shippers' demand, particularly in high demand seasons, the shipping company has to lease empty containers from leasing companies. Therefore, how many containers, when and where to lease containers on a short-term basis, and how long the shipping company keeps this leased containers become another set of important decisions. Substantial savings may be expected by making "good" container distribution decisions and leasing decisions. For example, a major European shipping company operated over 300,000 land container movements in 1986 with an estimated total distribution and transportation cost of some US $50,000,000, 40% of these were empty movements (Crainic, Gendreau and Dejax [1993]).

Dejax and Crainic [1987] noted that relatively little effort was directed toward developing models specifically at container transportation problems. The models proposed by Dejax and Crainic are designed for allocating empty containers in a landside distribution only. However, most of the task for allocating empty containers is carried out through ships. In this thesis, we proposed a model for allocating containers over sea and land.

## 1.1 Problem Context

This research is motivated by the analysis of fleet management problems encounter by a typical international shipping company. This company operates a large fleet of containers over a large marine and land network. To facilitate the understanding of the allocation model, this section is dedicated to a brief overview of the container allocation process.

The task of a shipping company is to carry goods from shippers to receivers over its transportation network. The transportation network consists of numerous coastal depots (i.e. ports) and inland depots. Ships carry containers, which generally comes in two sizes and about 15 different types, from one port to another. Arriving ships carry both loaded containers with imported goods and empty containers returning from previous exports. Loaded containers usually need to be transported to their destinations, i.e. the receivers' sites which may be located somewhere inland on the continent. They are moved by rail, trucks, barges and mixed modes. This latter option usually consists of an initial rail shipment to an inland terminal, followed by a final movement by truck. The empty containers that are imported may either held at the port or be immediately dispatched whenever they are needed for subsequent export. The requests for empty containers are made by shippers who need to export the goods. To meet these requests (demands), the company should provide enough empty containers at depots close to the shippers' sites at the time required by the shippers. Then the shippers can pick up those empty containers and move them to their own places to fill with loads of freight. After that, the loaded containers are returned to a nearby depot (usually the one from where the shippers pick up the empty containers). The loaded containers are transported to their destinations (usually one of the company's depot) by the shipping company. The receivers will then come to pick up the loaded containers, move them to their own sites and unload the containers. Once the containers are unloaded, the receivers are supposed to return the empty containers to a nearby depot within a certain period of time so that the shipping company can use these empty containers again to carry receivers' loads .

The demand for movements among various locations is often imbalances, and this implies the need for redistribution of empty containers over the companys transportation service network from locations at which they have become idle to locations at which they can be reused. Container dispatching decisions are made at discrete time intervals, say, at the beginning of each day.

The other aspect of the problem is how to make leasing decisions when the total supply of empty containers from the shipping company's container fleet (which consists of owned containers and long-term leased containers) is not enough to meet shippers' demand. A shipping company usually leases containers on short-term basis from leasing companys. Therefore a shipping company should decide when, where, and how many short-term containers to lease. Consequently, a shipping company should also decide when and where to return those short-term containers to the leasing companys. For example the shipping company signs contracts with several leasing companys, in which leasing conditions are specified. Each leasing company has a set of storage depots where the leased containers can be picked and returned (the so called "drop-off") to the leasing company. The costs of leasing a container consists of a pick-up cost, daily rental, and a drop-off cost.

This problem has arised in other mode of transportation as well. In trucking industry, the problem is how to allocate available motor carriers to move loads from origins to destinations. In railroads, it need to be decided how to allocate the rail cars to customers' demands given schedules of trains and capacities of each train.

## 1.2 Model Scope and Objectives

Our research focuses on modelling the container allocation and leasing process over the landside routes as well as sea routes. It is designed to be used in a real-time environment for making dispatch and leasing decisions while anticipating the downstream impacts of decisions made now. We assume that the forecasts of future demands and supplies are given and also the costs coefficients of moving containers. The locations of depots, ports and the route configuration are fixed during a planning horizon. The assignments of regions to depots or ports are also determined before hand.

The objective of the model is to obtain "good" decisions in terms of certain objective function, for example, minimizing total costs or maximizing total profit, on the following aspects:

1. allocating empty containers to meet the current demand of shippers,
2. redistributing remaining empty containers to storage depots of ports in anticipation of future demand,
3. where, when and how many containers does the shipping company need to lease containers when it experiences a container shortage.
4. where, when and how many containers does the shipping company should drop-off.

## 1.3 Outline of the Thesis

Chapter two presents the problem definition of container allocation, the literature review on vehicle allocation problems and model formulations. Chapter three is devoted to solution methods for the container allocation models. Chapter four presents the experimental design and experimental results and Chapter five gives the conclusions and recommendations.

# Chapter 2

# Model Formulation of Container Allocation Problem

## 2.1 Problem Definition

The term "vehicle" in the dynamic vehicle allocation problem can be a motor carrier, a trailer, a railroad car, or a container. So the container allocation problem is a type of dynamic vehicle allocation (DVA) problems. This problem was initially proposed by Powell [1984] for allocating truck motors to carry loads . However, for the DVA problems arising in shipping and railroads companies, there is usually a leasing option available. When the supplies of vehicles fall short at a depot, there may be a choice of leasing vehicles from vehicle leasing companies. In the DVA problem which is initially proposed by Powell, Sheffi and Thiriez [1984], the leasing options is not included in the problem definition.

Here we define the vehicle allocation problems for a more general case in the sense that a leasing option is available to the carrier. The definition of DVA problem can be written as follows : a carrier operates a fleet of vehicles, at any decision point of time, a set of vehicles are available for dispatching at some locations on the transportation system where the carrier operates. At the same time, there are some other locations in the transportation system where there are demands for vehicles. Note that the vehicle in the DVA problem can be a motor carrier, a container or a rail car. Besides the supply of vehicles owned by the carrier, there are supplies of vehicles provided by leasing firms with a certain rental fee at a set of locations on transportation system. The objective of the DVA problem is to find the "best" allocation and leasing decisions in terms of minimum total costs , or maximum profit over the long run.

The problem possesses the following properties. First, dispatching or allocation decisions have to be made before the future supplies and demands reveal their random outcomes i.e. we only have probabilistic knowledge about future demands (and supplies) of vehicles. Second, a certain length of planning horizon needs to be chosen to consider the downstream effects of container allocation decision made at current time period. Long planning horizons will obviously increase the size of the problem, whereas short planning horizons may not fully reflect the downstream effects. Third, on each time period, we run the model, but only the allocation decisions on the first period are implemented.

For the dynamic vehicle allocation problems, a possible list of problem input data may consist of the followings:

1. The supplies of vehicles at each depot on each time period.

2. The demands of vehicles at each depot on each time period to carry goods.

3. The costs of relocating empty vehicles and the revenues obtained by carrying loads.

4. The available traffic routes and capacities on each of those routes over the transportation system of the carrier.

5. The cost of leasing a vehicle and the maximum number of vehicles that can be leased at each locations.

Given those data, the objective of DVA is to find a best strategy to allocate the supplies of vehicles, and to determine the number of vehicles that should be leased to satisfy customer demands. The definition of the DVA problem given above can be considered as a more general case comparing to the one proposed initially by Powell , Sheffi and Thiriez [1984]. In the next section, we present a brief review on the DVA problems arising in different modes of transportation.

## 2.2 Literature Review on Vehicle Allocation Problems

The vehicle allocation problem was originally formulated as a transportation problem by considering current supplies of and demands for railroad cars, and ignoring downstream impacts of the decisions made at current time period [see Misra, 1972]. White and Bomberault [1969] and White [1972] indicated how the problem might be approached as a dynamic problem by setting up a network where each node represented a particular region at particular point in time. Future demands for vehicles between regions would be forecasted over a specified planning horizon.

Recognizing the frequently large uncertainty associated with the demand forecasts, Powell , Sheffi and Thiriez [1984] and Powell [1986] formulated the DVA problem as a nonlinear network by representing forecasts as random variables with known probability distributions. The simple and null recourse strategies were employed in above two papers respectively to solve the underlying stochastic model for the DVA problem. When the demand between two depots falls short of the number of trucks assigned to carry those loads, the simple recourse strategy assumes that the excess vehicles move empty anyway, while the null recourse strategy assumes that the excess vehicles are held at the origin depot rather than being dispatched empty. These two assumptions result in quite large errors in the final solutions of the these two models in some cases. Powell [1987] reformulated an alternative model in which he combined the ideas developed for the multi-period deterministic DVA with an approximate algorithm for two-stage stochastic DVA. Crainic, Gendreau and Dejax [1993] proposed formulations of a dynamic stochastic model for container allocation designed for the landside operations of a shipping company by using dynamic network modelling approach. Their model includes all type of containers, exchanges among different type of containers, all major "identified" customers, etc. There is no numerical experiments presented in the paper.

In all of above (DVA) models, the leasing option is not included in the modelling context. The unsatisfied demand due to shortage of container supply is simply assumed to be lost or rolled over to the next day. For the container models proposed by Crainic, Gendreau and Dejax [1993], they include the container allocation process on the land side operation of a shipping company only. The allocation strategies on the sea component of the service network is determined by another model using "balancing flows" among different regions of the company's network. These "balancing flows" are obtained by a separate model instead of by the same container allocation model. In their paper, possible approaches of solving the stochastic allocation model is discussed briefly, there is no numerical results presented. A shipping company usually needs to make container dispatching decisions and leasing decisions over its whole service network consisting of both sea routes and land routes. The allocation task of containers over the sea routes is a important component of the whole process. Determining allocation strategies by separate models will most likely result in suboptimal solutions.

In this thesis, we propose container allocation models which include both container dispatching and leasing decisions over the landside routes as well as sea routes. For the purpose of comparison, we present two different types of models: one type is dynamic deterministic models, and the other type is dynamic stochastic models. Our approach to the problem is to use dynamic networks to model the container allocation process over the transportation network consisting of coastal ports and inland depots. In the remaining of the thesis, we use the term "depot" for a inland depot as well as a coastal port unless otherwise specified. This approach of using network to model the allocation process was initially introduced in White and Bomberault [1969] and White [1972].

## 2.3 Modelling the Allocation Process by Dynamic Networks

Our approach of modelling the container allocation process is to use dynamic networks, in which nodes represent depots or ports, arcs represent traffic routes between two depots or ports. This section describes how the decision making process of container allocation can abstracted as time-space networks. The space and time dimensions of the allocation process can be modeled by a network. Each node in the network represents a depot or port on one particular day. A depot may correspond to an individual depot in the actual network or represent a group of neighboring depots (usually within a few hours of travel time). The depots and available routes in the transportation system define a network with time and space dimensions, we call it service network. A service network with 3 depots and 4 time period planning horizon is shown in Figure 2-1.

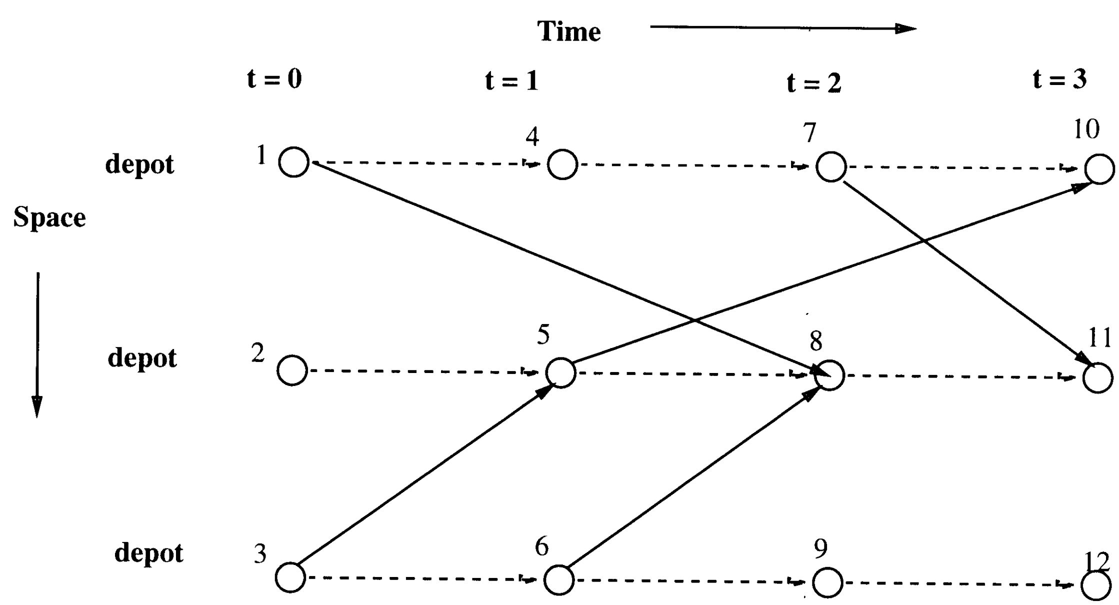


Figure 2-1: Time Space Network: Modelling Traffic Routes and Inventory Activities

Node 1 represents depot 1 at time period 1, node 2 represents depot 2 at time period 1. Node 4 represents depot 1 at time period 2. Arc (1,8) represents an available traffic route leaving depot 1 at time period 1 and arriving depot 2 at time period 3. The horizontal dotted arcs model the inventory activities at each depot.

To model the demand for empty containers at depot $i = 1$ on $t = 2$, we add an arc $(4, d_3)$ (node 4 represents the depot 1 at time period 2), and set the capacity of this arc equal to the demand for empty container at node 4. The cost of arc $(4, d_3)$ may set to equal to two values: in case (1) the value is the negative of the average net revenue earned by satisfying one unit of demand, wheras in case (2) the value is the negative of the cost of leasing an empty container at this node. In the second case, it is implicitly assumed that all demand has to be satisfied. The flow on this arc represents the portion of demand satisfied by company's own containers, while the difference between the arc capacity and the flow represents the portion of demand satisfied by leased containers. Adding those arcs to the network shown in Figure 2.1, the service network becomes the one as shown in Figure 2-2.

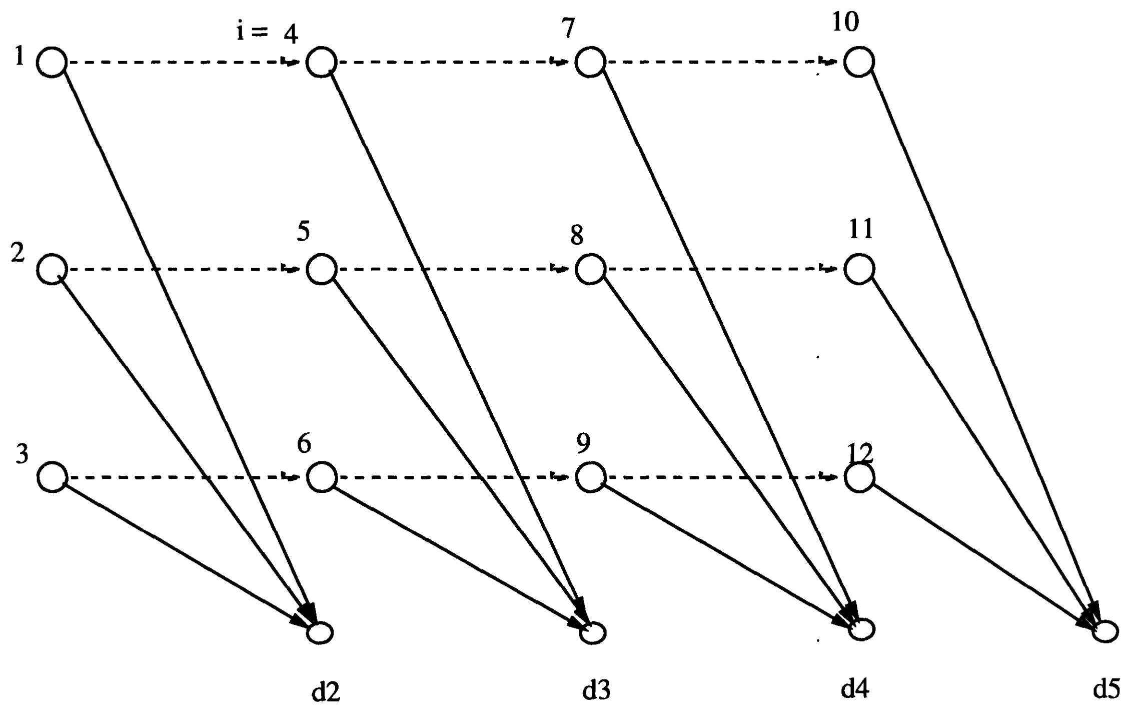


Figure 2-2: Time Space Network: Modelling the Demands for Empty Containers

The leasing cost of a container consists of "on-hire" cost, daily rental and "off-hire" cost. The "on-hire" cost is a fixed amount cost that the carrier should pay to the leasing company when the carrier picks up the container, while the "off-hire" cost is the amount paid when the carrier returns the container The "off-hire" cost also includes the cost spent for the maintanence of this container to ensure the returning conditions of this leased container meet certain returning "standards" set by the leasing company. Apart from these, there is also a rental cost which depends on the length of leasing. To model those three components of the total leasing cost, we add a source node for each time period and a set of arcs from each source node to all nodes representing the depots at next time period (see Figure 2-1). The costs on those arcs equal to the total leasing costs of a container over an average length of those trips originating from the head node of the corresponding arcs. The supply at each source for each time period are set to equal to the maximum number of containers that can be leased at that particular node during the corresponding time period. Since there is usually a delay between the time of requests for leasing containers and the time when the leasing company send over the empty containers to the shipping company. We assume for simplicity that the delay is one time period at all depots in the network. Thus if the shipping company sends a request to the leasing firm on how many containers needed at time $t$, then those containers will arrive at the shipping company's depot for use at time period $t+1$. Figure 2-3 shows the those leasing arcs. Combining Figure 2-1, Figure 2-2 and Figure 2-3, we obtain a complete network that models the container allocation process, and it is shown in Figure 2-4. It is clear from Figure 2-4 that this network is a dynamic network.

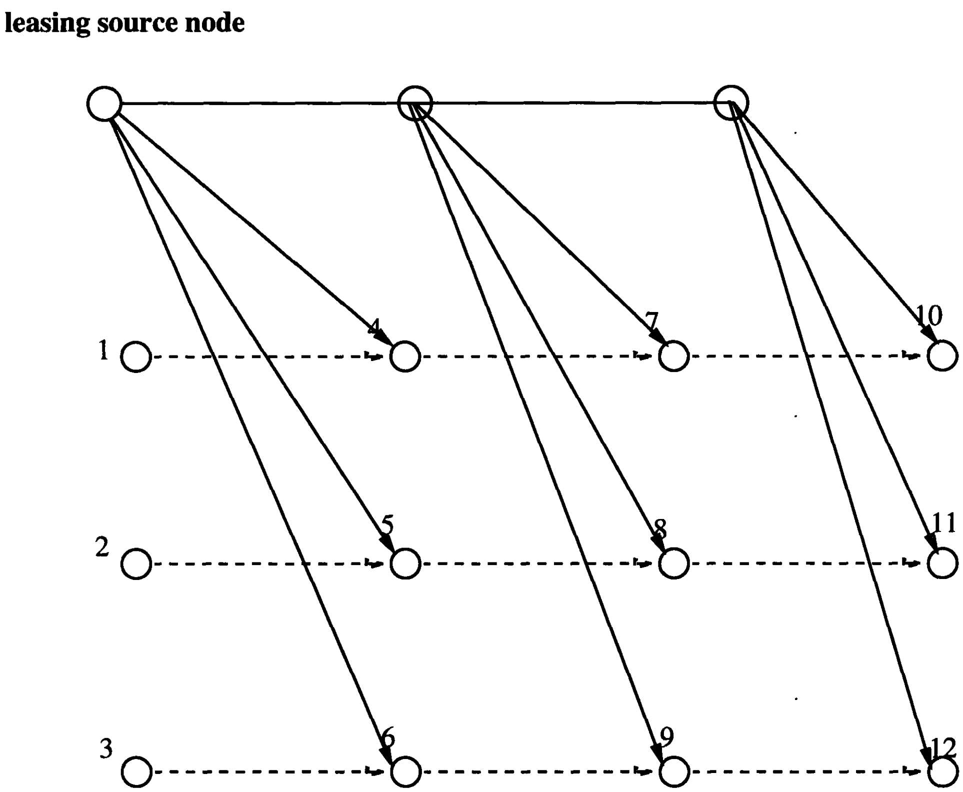


Figure 2-3: Time Space Network: Modelling the Leasing Options

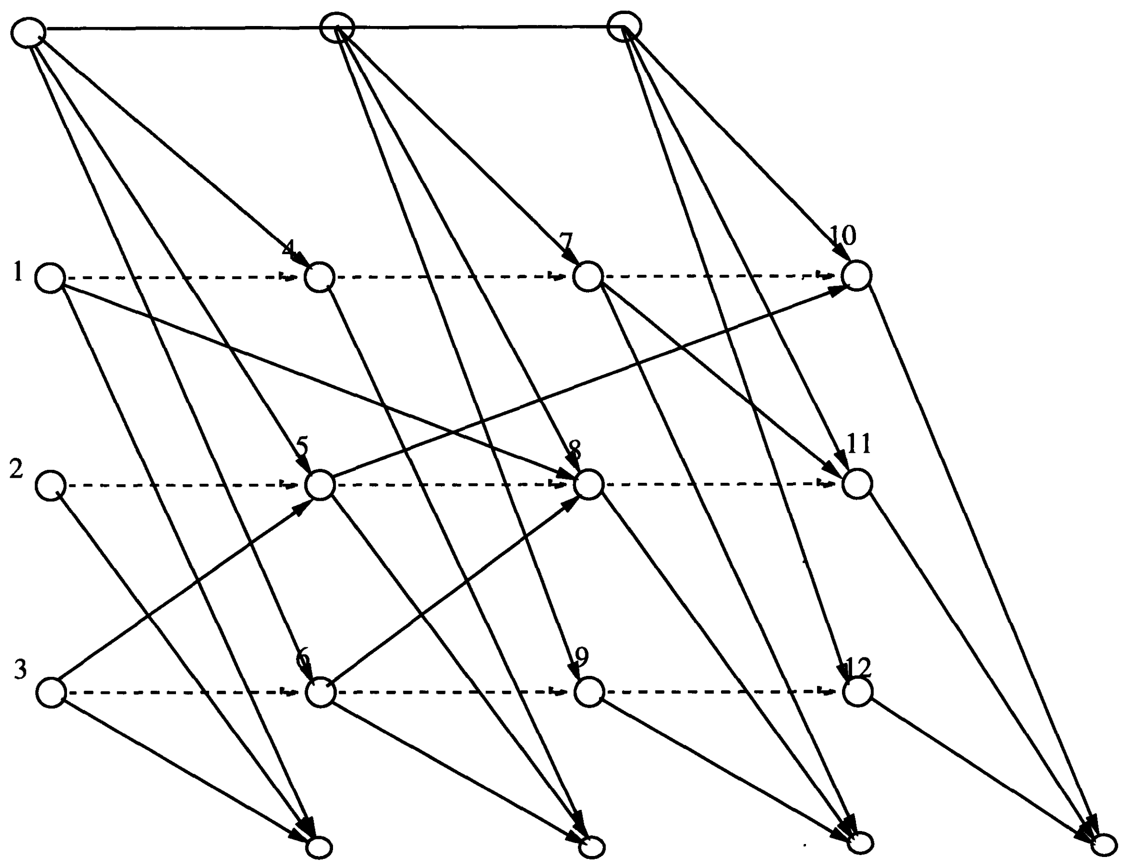


Figure 2-4: Time Space Network: Modelling the Container Allocation Process

Suppose that a planning horizon of $N$ time periods is chosen, decisions are made at discrete time intervals, say, at the beginning of each time period. Let $t_1, t_2, ..., t_N$ denote the times at which decisions are made during a planning horizon. Decisions made at time $t_1$ (current time period), will affect the container supplies at other depots at later times. To find "good" dispatching decisions at time $t_1$, we must consider their effect on container supplies and demands in later time periods and the expected decisions that would be made at time $t_2, ..., t_N$. This reflects the dynamic nature of the container allocation problem.

## 2.4 Dynamic Deterministic Models

For all deterministic models presented in this section, it is assumed that the model inputs are known with certainty. Two types of models are presented here, the first type involves only the allocation process of empty containers. It is assumed that all shippers' demand for empty containers must be satisfied. We can use a simulation model to track the movements of the loaded containers and to forecast the future supplies of empty containers at each depot. The second type involves both empty and loaded containers. When loaded movements are included, we need to determine the travel time of a loaded container from its origin to its destination, which is actually the turnarround time for a loaded container. In real world, the average turnaround time for a container is about 10 to 40 days depending on the distance from its origin to its destination, which is usually longer than the average turnaround time for vehicles in other modes of transportation. Also once a loaded container is sent to the receiver, it is receiver's decision when he will return the container to the shipping company although there may be penalties for keeping a container too long. The length of time the receiver will keep the container differ from place to place and from time to time. Therefore, the travel time of a loaded container (i.e. the turnaround time), which consists of following three components:

1. time since the shipper picks up the empty container at its origin depot until the time the container is sent back to shipping company's depot after it is loaded with goods,

2. the time for carrying the loaded container from its origin to its destination depot,

3. the time since its arrival at the its destination depot until its receiver get the container, unload the cargo and return it empty back to the shipping company.

is not known with certainty. It is difficult to model this uncertain travel time using dynamic networks. One approximation is to use an average travel time which may result large errors in the solution if the standard deviation of the travel time is relatively large. We will discuss the allocation models with both empty and loaded flow in next section. In this section, we focus on the model with empty flow only, and present a mathematical formulation of this model and solution methods to this model.

For the model with empty container flow only, we only consider the empty container movements of the allocation process. A separate simulation model can be used to track the loaded movements at each location and their current status, e.g. in shipper's location or in receiver's location or at a destination port to be sent to their receivers. This simulation model takes the allocation decisions as model inputs and forecasts the number of containers that will be returned to each depot or port by the receivers based on the historical data on the time the receivers at each location return the containers. We then can use these containers to satisfy new demands by using the historical data on the time the receivers at each location return the containers. We will not discuss this simulation in details since it is beyond the scope of this thesis. Here we simply assume that we already have the forecast of future supplies of empty container at each locations in our network. The objective of our container allocation model is to get the following outputs :

1. decisions of allocating empty containers to meet the demands for containers at the current time period.

2. decisions of redistributing remaining empty containers to storage depots of ports in anticipation of future demands.

3. where, when and how many does the firm need to lease on short term basis. container if supplies is in short.

To consider the downstream effects of decisions made at current time period, we need to choose a certain length of time as the planning horizon of the model. Here we assume that the planning horizon is $N$ time periods. In the following, a mathematical formulation of a deterministic model is presented.

### 2.4.1 Deterministic Models with Only Empty Flow

Before the presentation of our model, we define the following variables. Let

$\mathcal{N}$ = the set of depots in the system. The number of nodes at each time period equal to $|\mathcal{N}|$.

$\tau_{ij}$ = travel time from depot $i$ to depot $j$.

$a_j$ = average total cost for each container leased on depot $j$ (over an average length of leasing period).

$r_j$ = negative of average net revenue earned by satisfying a container load demand at depot $j$. When all the demands have to be satisfied, then we use $-a_j$ instead of $r_j$ to model this special case.

$c_{ij}$ = cost of moving an empty container form depot $i$ to depot $j$.

$\bar{\xi}_j(t)$ = market demand for empty containers at depot $j$ in period $t$.

$\bar{\eta}_j(t)$ = external supply for empty containers at depot $j$ in period $t$.

$\bar{\eta_{v}}_j(t)$ = available limit on number of containers that can be requested for leasing at depot $j$ at time $t$.

$\bar{u}_{ij}(t)$ = remaining capacity for empty containers on traffic route form depot $i$ to depot $j$ at time $t$.

$z_j(0)$ = number of empty containers leased in previous planning horizon at depot $j$.

$z_{j\beta}(0)$ = leasing quota carried over from previous planning horizon at depot $j$.

**Decision Variables :**

$x_{ij}(t)$ = empty container allocated from depot $i$ to depot $j$ at time $t$.

$y_j(t)$ = demand satisfied at depot $j$ at time $t$.

$z_j(t)$ = the request of number of empty containers the shipping company send to the leasing company for leasing at depot $j$ at time $t$, they will arrive at shipping company's depot for use at next time period.

$$\min_{x(t),y(t),z(t)} \sum_{t=1}^{N} \left\{ c^T x(t) + r^T y(t) + a^T z(t) \right\} \quad (2.1)$$

subject to

$$
\begin{aligned}
\sum_{k\in\mathcal{N}} x_{jk}(1) + y_j(1) &= \bar{\eta}_j(1) + z_j(0) \quad \forall j \in \mathcal{N} \\
z_j(1) + z_{j\beta}(1) &= \bar{\eta_{vj}}(1) + z_{j\beta}(0) \quad \forall j \in \mathcal{N} \\
0 \le x_{ij}(1) \le \bar{u}_{ij}(1) & \qquad \forall i \in \mathcal{N} \quad \forall j \in \mathcal{N} \\
0 \le y_j(1) \le \bar{\xi}_j(1) & \qquad \forall j \in \mathcal{N}
\end{aligned}
\tag{2.2}
$$

and for $t = 2, \ldots, N$ :

$$
\begin{aligned}
\sum_{k\in\mathcal{N}} x_{jk}(t) + y_j(t) - \sum_{(i,j)\in\mathcal{A}(t)} x_{ij}(t-\tau_{ij}) &= \bar{\eta}_j(t) + z_j(t-1) \quad \forall j \in \mathcal{N} \\
z_j(t) + z_{j\beta}(t) &= \bar{\eta_{vj}}(t) + z_{j\beta}(t-1) \quad \forall j \in \mathcal{N} \\
0 \le x_{ij}(t) \le \bar{u_{ij}}(t) & \qquad \forall i \in \mathcal{N} \quad \forall j \in \mathcal{N} \\
0 \le y_j(t) \le \bar{\xi}_j(t) & \qquad \forall j \in \mathcal{N}
\end{aligned}
$$

The above mathematical formulation is a minimum cost flow problem, and we can solve it very efficiently by using classical network optimization techniques such as network simplex algorithm. As a result, we get optimal flows $x(t)$, $y(t)$, $z(t)$, for $t = 1, 2, \ldots, N$. However, we only implemented the decisions at time period 1, i.e. $x(1)$, $y(1)$ and $z(1)$. When $t = 2$, we update the model inputs and run the model again with $N$ period planning horizon. The model is implemented according to a rolling horizon procedures.

### 2.4.2 Other Types of the Deterministic Models

**Extension to multi-type model:** The model described above can be extended to a multi-type model. In a fleet of containers, there are two major sizes of containers, i.e. 20 foot containers and 40 foot containers. For each size, there several different types of containers. In above model, only one type of containers is considered. It is implicitly assumed that allocation decisions for each type of containers can be made separately. However, since different types and sizes of containers share the same capacities on sea or rail routes. In this case, we need to design a multi-type model in which interactions of allocation decisions for each type of containers and the route capacity constraints are considered. The resulting formulation of the model are a multi-commodity minimum cost formulation with bundle constraints. Dantzig-Wolfe decomposition and other type of multicommodity flow algorithms can be employed to solve the model. In this thesis we will not consider the details of this model.

**Extension to Models with Empty and Loaded Flows** One shortcoming of the model described in section 2.2, we cannot determine when and where we should return the short-term leased containers. One approach of approximation is to use an average trip cost for a leased container at each depot (which equals to sum of "pick-up" cost, daily rental multiplying average trip time and "drop-off" cost) as a cost on each leasing arc. When a leased container returned by a receiver, we should first decide whether we return this container or use it for another trip before add it the empty container supply of that depot.

Another approach to obtain the solutions of where, when and how many does the firm should drop-off those containers is to construct a model which involve both empty and loaded containers. This model is built on a service network and a set of leasing networks, which are described as follows:

**Service Network**

The service network is the same as the one described for the previous model except there is one more set of loaded arcs. Each loaded arc starts with the same node as each empty arc, but it heads to the final destination depot and its travel time equals to actual plus the unstuffing time. The demand here are the actual number of loaded containers needed to be carried between each pair of origin and destination and it is modeled using the capacity of the corresponding loaded arc. The imbalance on each node are external empty container supply at that depot. For this model, if we impose the capacity constraint on a traffic route, then there will be a bundle constraint among the loaded flow arcs and the empty flow arc originating the same node.

**Leasing Networks**

Suppose the shipping company has signed contracts with $K$ leasing companies and each company has its own leasing prices (consisting of on-hire cost, rental and off-hire cost) and maximum number of containers at its leasing depots. Assume that each leasing company has its depots just next to every depot of the shipping company. This assumption is just for the simplicity of presentation, if there is no leasing depot next to one of the company's depot we can just delete this depot from the leasing network. Therefore, similar to the service network of the shipping company, the leasing network corresponds to $k^{th}$ leasing company has also time and space dimensions. However, the leasing network has two more nodes and two sets of arcs. One is the source node, the other is the sink node. The set of arcs between the source node and all other nodes represent "on-hire" activities and the set of arcs between the sink node and all other node represent the "off-hire" activities. The costs on those arcs represent on-hire or off-hire costs and capacities on those arcs represents maximum number of containers we can on-hire or off-hire in that particular node.

The objective of this model is to maximize total profit, i.e. revenue minus the sum of empty movement cost and leasing cost. The resulting formulation of the model is a multi-commodity minimum cost formulation with bundle constraints. Dantzig-Wolfe decomposition or other multicommodity flow algorithms can be employed to solve the model.

Extension to multi-type model is more complicated than that of the model which involves empty container flow only because there bundle constraints with both flows of different container types and empty, load flows of same container type.

The major drawback of the model is that the travel time of loaded container is not known with certainty. In the model, we use average travel time as an approximation and this may result errors since the variance of travel time may be large relative to the average travel time. The other shortcoming is that it takes more computation time to solve the multi-type model. In this thesis, we will not exploit further details about this model.

## 2.5 Dynamic Stochastic Models

In deterministic models, we assume that demands or supplies of empty containers in future periods are known with certainty or can be represented by their mean values. In practice, there are uncertainties associated with the forecasts of future demands and supplies. Usually only a portion of customers book their empty containers one or two weeks in advance, some may just pop up one or two days before they need the empty containers. In trucking industry, typically, 60% of the loads called in for pickup is on the same day, implying that at the beginning of the day the trucking company know only 40% of the loads that will be carried that day (see Powell [1988]).

In the stochastic model, we explicitly incorporate the uncertainty of demands and supplies by modelling them as a random elements. Their distributions can be obtained by forecasting models. For forcast models, it is beyond the scope of this thesis. Gendreau, M., Crainic, T.G. Dejax, P., and Steffan, H. [1991] present a preliminary study on forecast models for container demands. Here we assume that we already have the supplies and demands for empty containers at each depot on each time period. The objective of stochastic models is to find a best container allocation strategy at each time period while the uncertainties of future empty container demands and supplies are considered. The major difference between stochastic models and deterministic models is that in the stochastic one we have to find a best allocation strategy while we only have probabilistic knowledge of the demands and supplies for empty containers in future periods. In this section, two stochastic models are presented. One involves both empty and loaded flows, the other involves just the empty flow.

### 2.5.1 Stochastic Models with Both Empty and Loaded Flows

Before presenting the model formulation, let us define the followings : On each particular time period, assume that we have the following input data:

$\mathcal{N}$ = set of depots in the system,

$N$ = number of time periods in a planning horizon,

$c_{ij}^0$ = average cost of moving an owned empty container from depot $i$ to $j$

$h_i^0$ = average inventory cost of holding an owned empty container over the next time period of depot i.

$r_{ij}^0$ = average net revenue (price minus transportation cost) earned by a loaded containers moving from depot $i$ to j.

$\xi_{ij}(t)$ = random variable denoting the number of loads that will be called from $i$ to $j$ to be picked up at time t, $t = 2, ..., N+1$,

$\bar{\xi}_{ij}(1)$ = actual number of loads known at time $t = 1$ to be available moving from $i$ to $j$ at the first time period.

$b_i^0(t)$ = number of empty containers becoming available for the first time in depot $i$ at time t. Note $b_i^0(t) = 0$ for $t = 2, ..., N+1$.

For input data on the leasing network, we have the followings:

$e_{ij}^k$ = the rental cost over the travel time of moving a leased container leased from depot $i$ and moved to j, where $i \in \mathcal{N}$, $j \in \mathcal{N}$.

$c_{ij}^k = c_{ij}^0 + e_{ij}^k$ the total cost for an leased empty container moving from $i$ to j.

$r_{ij}^k = r_{ij}^0 - e_{ij}^k$ the net revenue earned by moving a loaded leased container from depot $i$ to j.

$b_i^k(t)$ = number of empty leased containers available for lease in depot $i$ at time t, $b_i^k(t) = 0$ for $t = 2, ..., N+1$. Those containers are leased in previous planning horizon, and becomes empty in the current horizon.

$c_{sj}^k$ = "on-hire" cost when an empty container is leased at $j^{th}$ time-space node in the leasing network.

$c_{iw}^k$ = "off-hire" cost when an empty container is drop off in $i^{th}$ time-space node in the leasing network.

$g_{sj}^k(t)$ = maximum number of containers that can be leased at $j^{th}$ time-space node from the leasing company,

$g_{iw}^k(t)$ = maximum number of containers that can be returned at $i^{th}$ time-space node to the leasing company,

$b_s^k$ = maximum number of containers that can be leased from the leasing company over the planning horizon.

Now let us define the decision variables for the stochastic model :

$x_{ij}^0(t)$ = number of loaded containers to be moved from depot $i$ to $j$ at time $t$ ($i \neq s$, $j \neq w$, $i \neq j$),

$y_{ij}^0(t)$ = number of empty containers to be moved from depot $i$ to depot $j$ at time (t). ($i \neq s$, $j \neq t$, $i \neq j$),

$y_{ii}^0(t)$ = number of empty containers to be hold at depot $i$ over the next time period at time t,

$y_{sj}^k(t)$ = number of empty containers leased from the leasing company in depot $j$ at time t

$x_{ij}^k(t)$ = number of loaded containers (using leased containers) to be moved from depot $i$ to depot $j$ at time $t$.

$y_{ij}^k(t)$ = number of empty leased containers to be moved from depot $i$ to j at time t.

$y_{ii}^k(t)$ = number of empty leased containers to be hold at depot $i$ over the next time period at time t.

$y_{sj}^k(t)$ = number of containers leased from leasing company in depot $j$ at time t.

$y_{iw}^k(t)$ = number of containers to be returned to kth leasing company in depot $i$ at time t.

Now we describe the formulation of the stochastic model. For simplicity of presentation of the model, it is assumed that the travel time between any two depots is one time period.

**Constraints**

Flow conservation constraints in service network:

$$\sum_{j\in\mathcal{N},i\neq j}^{N} x_{ij}^0(t) + \sum_{j\in\mathcal{N}} y_{ij}^0(t) - \sum_{j\in\mathcal{N},i\neq j} x_{ji}^0(t-1) - \sum_{j\in\mathcal{N}} y_{ji}^0(t-1) = b_i^0(t) \quad \forall i \in \mathcal{N} \quad (2.3)$$

for $t = 1, 2, \ldots, N$.

$$-\sum_{j\in\mathcal{N},i\neq j} x_{ji}^0(N) - \sum_{j\in\mathcal{N}} y_{ji}^0(N) + y_{iw}^0(N+1) = b_i^0(N+1) \quad (2.4)$$

for $i \in \mathcal{N}$.

Demand of the sink node in service network should equal to the total external supplies of all other nodes in the service network. Therefore, flow conservation constraint for sink node w is,

$$-\sum_{i\in\mathcal{N}} y^0_{iw}(N+1) = -\sum_{t=1}^{N+1}\sum_{i\in\mathcal{N}} b^0_i(t) \tag{2.5}$$

Flow conservation constraints in the leasing network.

$$\sum_{j\in\mathcal{N},i\neq j} x^k_{ij}(t)+\sum_{j\in\mathcal{N}} y^k_{ij}(t)+y^k_{iw}(t)-\sum_{j\in\mathcal{N},i\neq j} x^k_{ji}(t-1)-\sum_{j\in\mathcal{N}} y^k_{ji}(t-1)-y^k_{si}(t) = b^k_i(t) \tag{2.6}$$

for $i \in \mathcal{N}$ and $t = 1$.

$$-\sum_{j\in\mathcal{N},i\neq j} x^k_{ji}(N) - \sum_{j\in\mathcal{N}} y^k_{ji}(N) + y^k_{iw}(N+1) = b^k_i(N+1) \tag{2.7}$$

for $i \in \mathcal{N}$.

For the source node, the flow conservation constraint is :

$$\sum_{i\in\mathcal{N}}\sum_{t=1}^{P} y^k_{si}(t) = b^k_s \tag{2.8}$$

For the sink node:

$$-\sum_{j\in\mathcal{N}}\sum_{t=1}^{P+1} y^k_{jw}(t) - y^k_{sw} = -b^k_w \tag{2.9}$$

where $(b^k_w = b^k_s)$, and

$y^k_{sw}$ = number of containers that is not leased during the planning horizon, which is actually the flow on the arc $(s, w)$ (the cost of this arc is zero and capacity of this arc is infinite.)

The constraints for the maximum number of containers that can be leased or dropped off at each depot at each time period are as follows:

$$y^k_{sj}(t) \leq g^k_{sj}(t) \tag{2.10}$$

for $j \in \mathcal{N}$ and $t = 1, \ldots, N$.

$$y_{iw}^k(t) \leq g_{iw}^k(t) \tag{2.11}$$

for $i \in \mathcal{N}$ and $t = 1, \ldots, N$.

The bundle constraints among decision variables in service network and those of the leasing network are as follows:

$$\sum_{k=0}^{k} x_{ij}^k(1) \leq \bar{\xi}_{ij}(1) \tag{2.12}$$

for $i \in \mathcal{N}$, $j \in \mathcal{N}$.

$$\sum_{k=0}^{k} x_{ij}^k(t) \leq \xi_{ij}(t) \tag{2.13}$$

for $i \in \mathcal{N}$, $j \in \mathcal{N}$ and $t = 2, \ldots, N$.

The last set of constraints are nonnegative constraints:

$$x_{ij}^k(t) \geq 0, \quad y_{ij}^k(t) \geq 0 \tag{2.14}$$

for $i \in \mathcal{N}$, $j \in \mathcal{N}$, $t = 2, \ldots, N$and $k = 0, \ldots, K$.

**Objective Function:**

$$z = \min \left\{ -\sum_{k=0}^{K} \left\{ \sum_{j \in \mathcal{N}, i \neq j} r_{ij}^k x_{ij}^k(1) + \sum_{j \in \mathcal{N}} c_{ij}^k y_{ij}^k(1) \right\} + \sum_{j \in \mathcal{N}} [c_{sj}^k y_{sj}^k(1) + c_{jw}^k y_{jw}^k(1)] + \right.$$
$$E_{\xi(2)} \left\{ \min \left\{ -\sum_{k=0}^{K} \left\{ \sum_{j \in \mathcal{N}, i \neq j} r_{ij}^k x_{ij}^k(2) + \sum_{j \in \mathcal{N}} c_{ij}^k y_{ij}^k(2) \right\} + \sum_{j \in \mathcal{N}} [c_{sj}^k y_{sj}^k(2) + c_{jw}^k y_{jw}^k(2)] \right\} + \cdots + \right.$$
$$\left. E_{\xi(N)} \left\{ \min \left\{ -\sum_{k=0}^{K} \left\{ \sum_{j \in \mathcal{N}, i \neq j} r_{ij}^k x_{ij}^k(N) + \sum_{j \in \mathcal{N}} c_{ij}^k y_{ij}^k(N) \right\} + \sum_{j \in \mathcal{N}} [c_{sj}^k y_{sj}^k(N) + c_{jw}^k y_{jw}^k(N)] \right\} \cdots \right\} \right.$$

subject to the constraints presented above.

Here $z$ equal to the maximum total expected profits over the $N$ period planning horizon, given the initial container allocation $b_i^0(1)$, $b_i^k(1)$. Thus the random elements apear on the right-hand sides in the constraint set. From the formulation, it is clear that the flow on each arc in the network not only depends on the supply of its tail node but also depends on on the demand on the adjacent arcs through the network conservation constraints. Therefore the resulting formulation is a multistage stochastic program with the recourse. For each random outcome of the random elements in the model, the underlying optimization problem is a multicommodities flow problem. Even we approximate above multistage stochastic program with recourse to a two stage stochastic program by assuming that all random elements reveal their outcomes at the beginning of stage two, it is numerically intractable for large scale network to solve this model. Furthermore, the average travel time of loaded containers is random and its variance is relatively large compare to travel time of an empty move. By using the average travel time as an approximation in our model, it is likely to give poor results. Because of this drawback and complexity of multicommodity formulation of above stochastic model, we will mainly focus on the stochastic model with empty flow only, which is presented in the following section.

### 2.5.2 Stochastic Model with Only Empty Flow

This model is similar to its deterministic counterpart presented in section 2.4.1 except that the demands, supplies and route capacities in future time period are not known with certainties. Here we only describe the model inputs and decision variables which are differ from those of the deterministic model. Let

$\bar{\xi}_j(1)$ = demand for empty containers at depot $j$ in period 1.

$\bar{\eta}_j(1)$ = external supply for empty containers at depot $j$ in period 1.

$\bar{\eta_{v}}_j(1)$ = available limit on number of containers that can be requested for leasing at depot $j$ at time 1.

$u_{ij}(1)$ = remaining capacity for empty containers on traffic route form depot $i$ to depot $j$ at time 1.

$z_j(0)$ = number of empty containers leased in previous planning horizon at depot $j$.

$z_{j\beta}(0)$ = leasing quota carried over from previous planning horizon at depot $j$.

**Stochastic Elements :**

For each sample point $\omega \in \Omega$, we define the random outcome of the following coefficients :

$\xi_j(t,\omega)$ = market demand for empty containers at depot $j$ in period $t$ for $t = 2,3,...,N$.

$\eta_j(t,\omega)$ = external supply of empty containers at depot $j$ in period $t$ for $t = 2,3,...,N$.

$\bar{\eta_{vj}}(t,\omega)$ = available limit on number of containers that can be requested for leasing at depot $j$ at time $t$.

$\bar{u}_{ij}(t,\omega)$ = remaining capacity for empty containers on traffic route form depot $i$ to depot $j$ at time $t$.

**Decision Variables :**

$x_{ij}(1)$ = empty container allocated from depot $i$ to depot $j$ at time 1.

$y_j(1)$ = demand satisfied at depot $j$ at time 1.

$z_j(1)$ = the request of number of empty containers the shipping company send to the leasing company for leasing at depot $j$ at time 1, they will arrive at shipping company's depot for use at next time period.

For stage two variables, they are defined for each random realization $\omega$ and for each time period $t = 2,3,\cdots,N$. For each $\omega \in \Omega$, let

$x_{ij}(t,\omega)$ = empty container allocated from depot $i$ to depot $j$ at time $t$.

$y_j(t,\omega)$ = demand satisfied at depot $j$ at time $t$.

$z_j(t,\omega)$ = the request of number of empty containers the shipping company send to the leasing company for leasing at depot $j$ at time $t$, they will arrive at shipping company's depot for use at next time period.

These variables are called "corrective actions" taken at following stages when the random coefficients reveal their realizations. However, we only implemented the allocation decisions at stage 1, these "corrective actions" variables are not implemented..

If we assume all random elements reveal their realizations at the beginning of stage two, the problem becomes a two stage stochastic program with network recourse. For simplicity of presentation, we assume the travel time between any two depots are one period. The formulation is given :

$$\min_{x(1),y(1),z(1)} c^T x(1) + r^T y(1) + a^T z(1) + E_\omega Q(S(1),\omega) \tag{2.15}$$

subject to

$$\begin{aligned}
\sum_{k\in\mathcal{N}} x_{jk}(1) + y_j(1) &= \bar{\eta_j}(1) + z_j(0) \quad \forall j \in \mathcal{N} \\
z_j(1) + z_{j\beta}(1) &= \bar{\eta_{vj}}(1) + z_{j\beta}(0) \quad \forall j \in \mathcal{N} \\
\sum_{i\in\mathcal{N}} x_{ij}(1) - S_j(1) &= 0 \qquad \forall j \in \mathcal{N} \\
0 \le x_{ij}(1) &\le \bar{u}_{ij}(1) \qquad \forall i \in \mathcal{N} \quad \forall j \in \mathcal{N} \\
0 \le y_j(1) &\le \bar{\xi}_j(1) \qquad \forall j \in \mathcal{N}
\end{aligned}$$

where

$$Q_2(S(1),\omega) = \min_{x(t,\omega),y(t,\omega),z(t,\omega)} \sum_{t=2}^{N} \left\{ c^T x(t,\omega) + r^T y(t,\omega) + a^T z(t,\omega) \right\} \tag{2.16}$$

s.t.

$$\begin{aligned}
\sum_{k\in\mathcal{N}} x_{jk}(t,\omega) + y_j(t,\omega) - \sum_{i\in\mathcal{N}} x_{ij}(t-1,\omega) &= \eta_j(t,\omega) + z_j(t-1,\omega) \quad \forall j \in \mathcal{N}, \quad \forall t = 2,\ldots,N \\
z_j(t,\omega) + z_{j\beta}(t,\omega) &= \eta_{vj}(t,\omega) + z_{j\beta}(t-1) \quad \forall j \in \mathcal{N} \\
0 \le x_{ij}(t,\omega) &\le u_{ij}(t,\omega) \qquad \forall i \in \mathcal{N} \qquad j \in \mathcal{N} \\
0 \le y_j(t,\omega) &\le \xi_j(t,\omega) \qquad \forall j \in \mathcal{N}, \ \forall t = 2,\ldots,N
\end{aligned}$$

A dynamic network with two depots and three time periods planning horizon is shown in Figure 2-5. The solid component of the network is the stage one network and the dotted component is the stage two network. Adding sink and source nodes, we have the following network shown in Figure 2-6.

In next chapter, we will discuss the solution methods for this dynamic stochastic model.

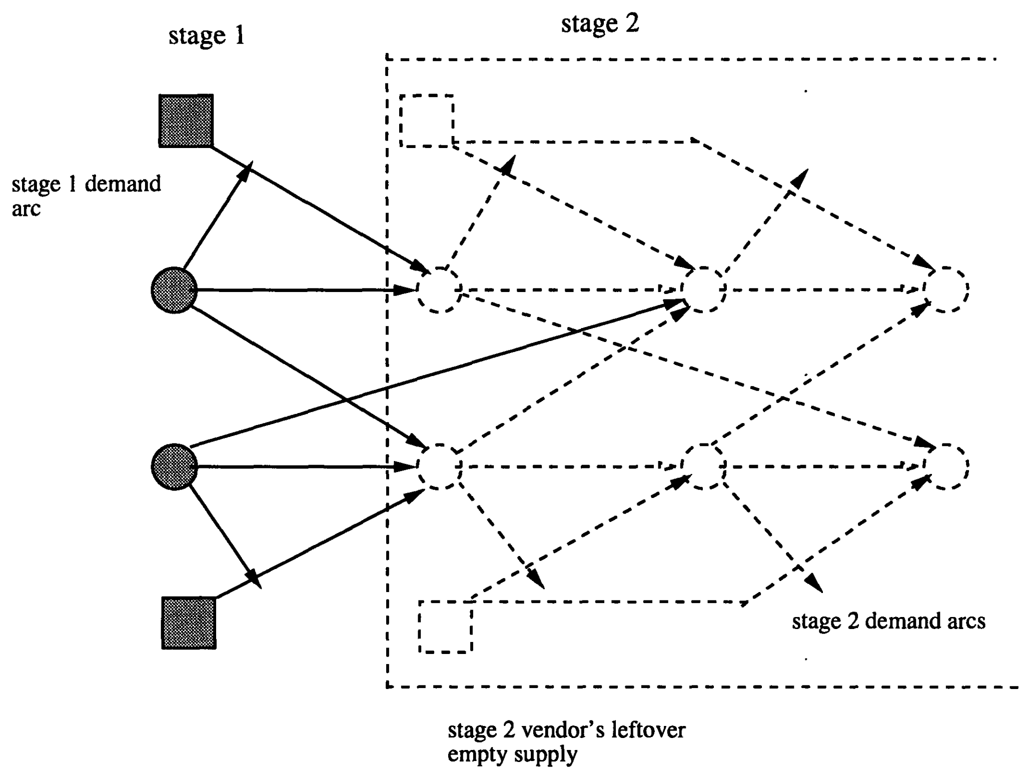


Figure 2-5: An example of network with two depots and three time periods planning horizon

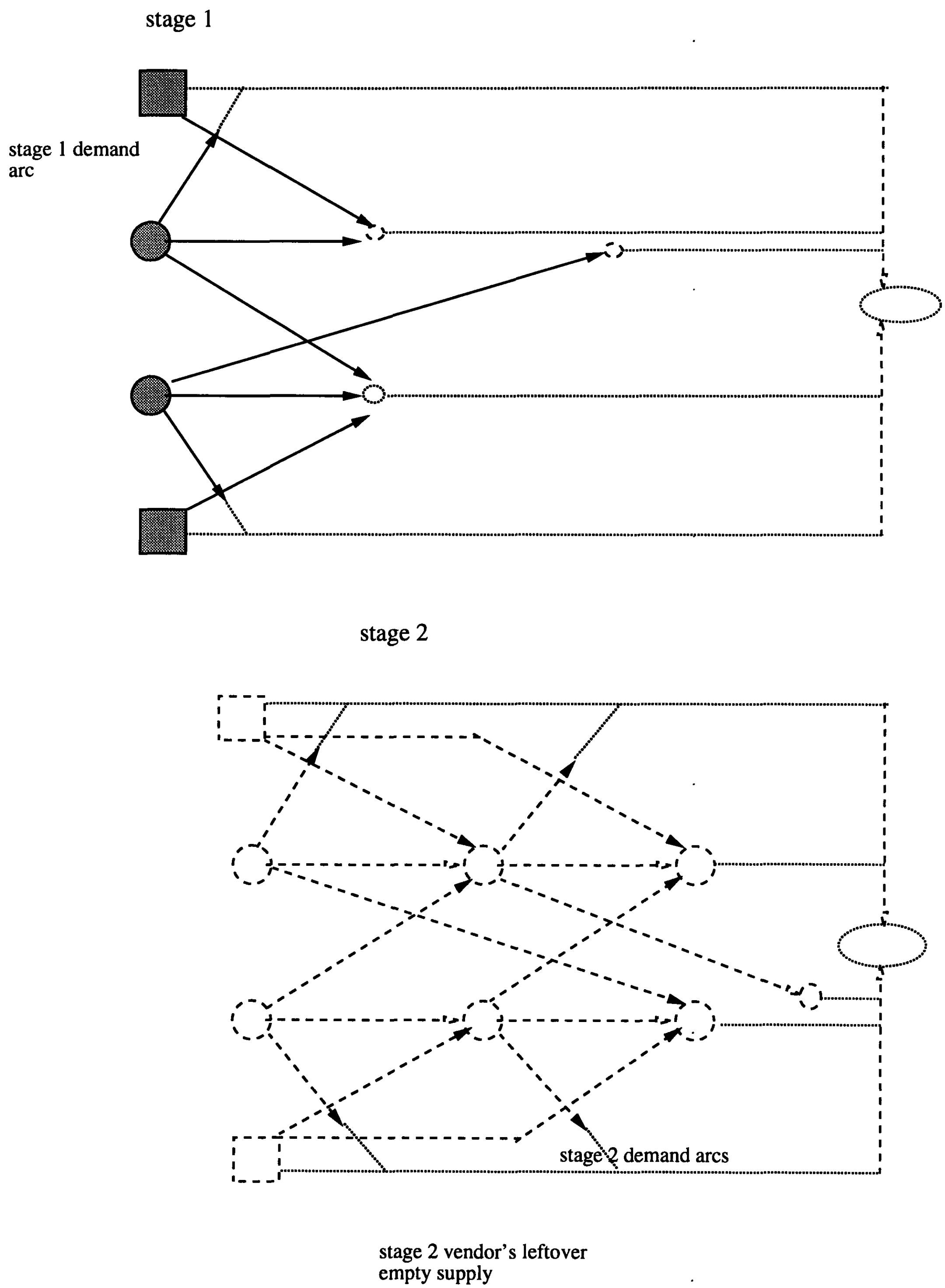


Figure 2-6: An example of network with two depots and three time periods planning horizon

# Chapter 3

# Solution Methods for the Stochastic Model

## 3.1 Introduction

In section 2.5.2, we present a mathematical formulation of a dynamic stochastic model for empty container allocation. The underlying optimization problem is a stochastic linear program with network recourse. In this thesis all random elements in the stochastic model, are assumed to be independently and discreetly distributed. Let us write the stochastic model presented in last chapter in matrix format:

$$\begin{aligned} z = \min_{x}\ & cx + E\{Q(x,\omega)\} \\ s.t.\quad & Ax = b \\ & 0 \le x \le u \end{aligned} \tag{3.1}$$

where, for each $\omega \in \Omega$, the recourse function $Q(x, \omega)$ is obtained by solving the following recourse problem.

$$\begin{aligned} Q(x,\omega) &= \min_{y(\omega)} qy(\omega) \\ s.t.\quad Wy(\omega) &= \eta(\omega) - Tx \\ 0 \le y(\omega) &\le \xi(\omega) \end{aligned} \tag{3.2}$$

In this chapter, we describe three methods for solving the stochastic program with recourse. They are the L-shaped method, the dual method for solving the equivalent deterministic problem and stochastic quasigradient methods. Section 3.2 , 3.3 and 3.4 describe these three methods respectively. Section 3.5 gives the application of stochastic quasigradient method to solve a two-stage stochastic program with recourse , which is the formulation of the stochastic model for container allocation. In section 3.5, we discuss the implementation issues of stochastic quasigradient method.

## 3.2 L-shaped Method:

L-shaped decomposition method proposed by Van Slyke and Wets [1969] is among the most popular methods for solving two-stage linear program with fixed recourse. It is a type of partial decomposition methods and uses cutting planes to estimate the expected recourse function from below, while maintaining primal feasibility. Consider the formulation of the stochastic model for empty container allocation presented in the beginning of this chapter, i.e. problems defined by ( 3.1) and ( 3.2), for every feasible solution of stage one problem ( 3.1), there is always a solution to stage two problem (the stage two problem is defined by ( 3.2)). So at every iteration, we do not need to concern about the feasibility of stage two problem for a given stage one solution $x$. Then, the basic steps of L-shaped method consists of

**Step 0**. Set $v = s = 0$

**Step 1**. Set $v = v + 1$ and solve the linear program

$$z = \min_{x} cx + \theta \tag{3.3}$$

$$s.t. \quad Ax = b \tag{3.4}$$

$$0 \leq x \leq u \tag{3.5}$$

$$E_t x + \theta \geq e_t \quad for \quad t = 1, ..., s \tag{3.6}$$

Let $(x^v, \theta^v)$ be an optimal solution of above problem. If there are no constraints of type ( 3.6), the variable $\theta$ is ignored in the computation of the optimal $x^v$, and set

$\theta^v = -\infty$.

**Step 2** For every random realization $\omega \in \Omega$ solve the linear program given by

$$
\begin{aligned}
(P4) \quad & \min_{y(\omega)} qy(\omega) \\
s.t. \quad & Wy(\omega) = \eta(\omega) - Tx^v \\
& -y(\omega) \geq -\xi(\omega) \\
& y(\omega) \geq 0
\end{aligned}
\tag{3.7}
$$

Let $(\pi(v,\omega), \alpha(v,\omega))$ be the simplex multipliers associated with the optimal solution of problem $(P4)$ . Set $t = t+1$ and define

$$E_t = E\{\pi(v,\omega)T\}, \tag{3.8}$$

$$e_t = E\{\pi(v,\omega)\eta(\omega) - \alpha(v,\omega)\xi(\omega)\} \tag{3.9}$$

and

$$w^v = e_t - E_t x^v \tag{3.10}$$

**Step 3** If $\theta^v \geq w^v$, we stop; $x^v$ is declared to be an optimal solution. Otherwise, we return to Step 1 with a new constraint of type ( 3.6).

In step 2, dual solutions to problem $(P4)$ need to be obtained, one approach is to use network simplex method to get $(\pi(\omega), \alpha(\omega))$. Here, we describe this procedure as follows. For simplicity of notation we drop $v$ superscript and rewrite $(P4)$ as:

$$
\begin{aligned}
(P4) \quad & \min_{y(\omega)} qy(\omega) \\
s.t. \quad & Wy(\omega) = \eta(\omega) - Tx \\
& -y(\omega) \geq -\xi(\omega) \\
& y(\omega) \geq 0
\end{aligned}
\tag{3.11}
$$

Its dual problem is:

$$\max_{y(\omega)} \pi(\omega)(\eta(\omega) - Tx) - \alpha(\omega)\xi(\omega) \tag{3.12}$$

$$s.t. \quad \pi_i(\omega) - \pi_j(\omega) - \alpha_{ij}(\omega) \leq q_{ij} \quad \forall (i,j) \in \mathcal{A}$$

$$\alpha_{ij}(\omega) \geq 0 \quad \forall (i,j) \in \mathcal{A}$$

$$\pi_j(\omega) \ \ unrestricted \ \ in \ \ signs, \ \ for \ \ all \ \ j \in \mathcal{N}$$

Since the objective function coefficient of $\alpha_{ij}(\omega)$ is $-\xi_{ij}(\omega)$ ($\xi_{ij}(\omega) \geq 0$), to maximize the objective function, $\alpha_{ij}(\omega)$ need to be as small as possible. Therefore we have

$$\alpha_{ij}(\omega) = \max\{0, -q_{ij} + \pi_i(\omega) - \pi_j(\omega)\} \tag{3.13}$$

Using network simplex we obtain $\pi_j(\omega)$, $j \in \mathcal{N}$, so $\alpha_{ij}(\omega)$ can be calculated easily by equation ( 3.13).

The idea of L-shaped method is to replace the second stage problem defined by ( 3.2) by a set of inequalities expressed only in terms of first stage variable $x$, i.e. the cutting planes and a scalar variable $\theta$. Since this cutting planes is used to check the optimality and hence they are called *optimality cuts.* The scalar variable $\theta$ represents the contribution of second stage to the objective function $z$. The convergence of the algorithm under the appropriate nondegeneracy assumption, to an optimal solution of the problem defined by ( 3.1) and ( 3.2), is based on the fact that there are only a finite number of optimality cuts. For each cut is corresponding to a basis of matrix $W$ and there are only finite number of basis for any given matrix.

Birge [1985] described his implementation of this method and extend it to solve multistage stochastic program with recourse. Wallace[1986] applies L-shaped algorithm to solve stochastic programs with network recourse. The formulation of the container model is a stochastic program with network recourse. The recourse problem presented by Wallace [1986] is the one with only equality constraints, and the random elements is only on the right hand side of these constraints. In the container allocation model, random elements appears on the right hand sides of all flow conservation constraints and the upper bounds of the capacity constraints. The L-shaped algorithm presented in this section applies to case where the right sides of both flow conservation constraints and arc capacity constraints are random elements.

However, as the problem size increases, solving the two-stage problem by the L-shape algorithm is usually numerically intractable due to the proliferation of possible realizations of random variables. For the container allocation model, the underlying problem size is very large, it is numericaly infeasible to use the L-shape method to solve the model.

## 3.3 Dual Methods for the Equivalent Deterministic Problem

Another approach is to write the equivalent deterministic problem of the two-stage stochastic program with recourse, and exploit the underlying special structure of the dual problem of this equivalent deterministic problem. Assume the random elements in the two stage program described at the beginning of this chapter is independently and discretely distributed, and there are total of $K$ possible outcomes, with corresponding probabilities of $(p_1, \ldots, p_K)$. Then, the equivalent deterministic problem of the stochastic program defined by ( 3.1) and ( 3.2) is given by:

$$\min_x cx + \sum_{k=1}^{K} p_k q y(\omega^k) \tag{3.14}$$

subject to

$$\begin{aligned} Ax &= b_1 \\ 0 \le x &\le u \\ Tx + Wy(\omega^k) &= \eta(\omega^k) \quad k = 1, \ldots, K \\ y(\omega^k) &\le \xi(\omega^k) \quad k = 1, \ldots, K \qquad (3.15) \\ y(\omega^k) &\ge 0 \quad k = 1, \ldots, K \qquad (3.16) \end{aligned}$$

Adding slack variable to capacity constraints, i.e.

$$x + Is = u$$

$$y(\omega^k) + Is(\omega^k) = \xi(\omega^k)$$

and combining with the equality constraints by changing matrices $A$, $T$ and $W$ to $\bar{A}$, $\bar{T}$ and $\bar{W}$ and letting $b = (b_1, u)^T$, $h(\omega^k) = (\eta(\omega^k), \xi(\omega^k))^T$, the above deterministic problem becomes,

$$\begin{aligned} &\min_x cx + \sum_{k=1}^{K} p_k q y(\omega^k) \\ s.t. \quad &\bar{A}x = b \\ &\bar{T}x + \bar{W}y(\omega^k) = h(\omega^k) \quad k = 1, \ldots, K \\ &x \geq 0, \quad y(\omega^k) \geq 0, \quad k = 1, \ldots, K \end{aligned} \tag{3.17}$$

Its dual problem is given

$$\begin{aligned} &\max_{(\sigma, \pi(\omega^k))} \sigma b + \sum_{k=1}^{K} p_k \pi(\omega^k) h(\omega^k) \\ s.t. \quad &\sigma \bar{A} + \bar{T} \sum_{k=1}^{K} p_k \pi(\omega^k) \leq c \\ &\pi(\omega^k) \bar{W} \leq p_k q. \\ &\sigma, \pi(\omega^k) \ \ unrestricted \ \ for \ \ k = 1, \ldots, K \end{aligned} \tag{3.18}$$

Let

$$\pi(\omega^k) = \hat{\pi}(\omega^k)/p_k$$

we have a block angular structure, i.e.

$$\begin{aligned} &\max_{(\sigma, \hat{\pi}(\omega^k))} \sigma b + \sum_{k=1}^{L} \hat{\pi}(\omega^k) h(\omega^k) \\ s.t. \quad &\sigma A + \sum_{k=1}^{L} \hat{\pi}(\omega^k) \bar{T} \leq c \\ &\hat{\pi}_k \bar{W} \leq q \quad k = 1, \ldots, K \end{aligned}$$

$$\sigma, \hat{\pi}(\omega^k) \ \ unrestricted, \quad k = 1, \ldots, K$$

Dantzig and Madansky [1961] apply Dantzig-Wolfe decomposition to above dual problem. Let $\beta^i(\omega^k), i \in \bar{J}_k$ be extreme points of polyhedron

$$\{\hat{\pi}(\omega^k) | \hat{\pi}(\omega^k)\bar{W} \leq q, \hat{\pi}(\omega^k) \ \ unrestricted\} \tag{3.19}$$

Then form the following restricted master problem:

$$\begin{aligned} \max \quad & \sigma b + \sum_{k=1}^{K} p_k h(\omega^k)(\sum_{i \in J_k} \lambda_k^i \beta^i(\omega^k)) \\ s.t. \quad & \sigma \bar{A} + \bar{T} \sum_{k=1}^{K} p_k(\sum_{i \in J_k} \lambda_k^i \beta^i(\omega^k)) \leq c \\ & \sum_{i \in J_k} \lambda_k^i = 1 \quad for \quad k = 1, \ldots, K \\ & \lambda_k^i \geq 0, \quad for \quad i \in \bar{J}_k, \quad k = 1, \ldots, K \end{aligned} \tag{3.20}$$

Let $x, w^1, w^2, \ldots, w^k$ be corresponding simplex multipliers to above constraints. Then solve the following $K$ subproblems:

$$\begin{aligned} (P_k) \qquad v^k = \max \ \ & (p_k \eta(\omega^k) - x)\pi(\omega^k) \\ s.t. \quad & \pi(\omega^k)\bar{W} \leq q \\ & \pi(\omega^k) \ \ unrestricted \end{aligned} \tag{3.21}$$

If $v^k \leq w^k$ for $k = 1, \ldots, K$, then the solution to the restricted master problem is optimal. If, however, $v > w^1$, we obtain a new extreme point $\beta^l(\omega^1)$ adding to the restricted master problem by introducing a new variable $\lambda_1^l$. This procedure goes on until $v^k \leq w^k$ for all $k = 1, \ldots, K$. Then the solution to restricted master problem is optimal.

However, this method is applicable to a stochastic program with relatively large $K$ since we need to solve $K$ number subproblems at each iteration, and $K$ is usually very large for practical problems.

## 3.4 Stochastic Quasigradient Methods

Stochastic quasigradient (SQG) methods are stochastic algorithmic procedures for solving general constrained optimization problems. SQG methods allow us to solve optimization problems with objective functions and constraints of such a complex nature that it is impossible to calculate the precise values of these function values, let alone of their derivatives. The basic idea of this approach is to use statistical estimates of objective function values and derivatives. Consider the following stochastic optimization problem:

$$\min_{x \in \mathcal{X}} F(x) = E_\omega f(x, \omega) \tag{3.22}$$

where $\mathcal{X} \subseteq \mathcal{R}^n$ is a feasible set and $\omega$ is an outcome of the probability space $(\Omega, \mathcal{F}, \mathcal{P})$. SQG methods typically involve first choosing an initial value $x^0 \in \mathcal{X}$ and then generating a sequence $x^k$ by :

$$x^{k+1} = x^k + a_k d^k(x^k, \omega^k) \tag{3.23}$$

where $\omega^k \in \Omega$ is particular outcome. The vector $d^k(x^k, \omega^k)$ is called the step direction which depends on statistical estimates of the gradient or subgradient of $F(x)$, denoted by $g^k(x^k, \omega^k)$. A statistical estimate, $g^k(x^k, \omega^k)$, of the gradient or subgradient of $F(x)$ is defined as follows:

If $F(x)$ is differentiable, then $g^k(x^k, \omega^k)$ should satisfy :

$$E_\omega(g^k(x^k, \omega^k)|x^0, \ldots, x^k) = \nabla F(x^k) + b_k \tag{3.24}$$

If $F(x)$ is nondifferentiable, then $g^k(x^k, \omega^k)$ should satisfy :

$$E_\omega(g^k(x^k, \omega^k) \mid x^0, \ldots, x^k) = \hat{F}(x^k) + b_k \tag{3.25}$$

where $\hat{F}(x^k)$ is a subgradient of $F(x)$ at $x^k$. The vector $b_k$ is the error terms and

$\| b_k \| \rightarrow 0$ as $k \rightarrow 0$.

If $b_k \equiv 0$, $g^k(x^k, \omega^k)$ is called *stochastic generalized gradient* (or *stochastic gradient* when $F(x)$ is differentiable), otherwise it is called *stochastic quasigradient.*

Stochastic quasigradient methods extend the stochastic approximation methods for solving unconstrained stochastic problems to solving constrained ones. For stochastic approximation method, Robbin and Monro [1951] first proposed an iterative method to find the root of monotonic regression function. Kiefer and Wolfowitz [1952] extend the Robbin-Monro algorithm to finding the maximum of a regression function. They differ in how they estimate the gradient of the objective function. The Robbin-Monro algorithm estimates the gradient directly, whereas the Kiefer-Wolfowitz algorithm uses finite differences to estimate the gradient. Robbin and Monro [1952] showed in their paper the sequence $\{x^k\}$ produced by the algorithms converges in probability to the optimal solution $x^*$. In 1954, Blum [1954] proved the multidimensional analog of the Kiefer-Wolfowitz methods with weekend assumptions, and more strong convergence, i.e. convergence with probability one. The monograph by Wasan [1969] reviews the progress of stochastic approximation methods during the 1950's and the early 1960's. Stochastic approximation methods presently experienced a resurgence of interest due to the recent development of efficient gradient estimation techniques(see Gassmann [1991]).

SQG methods generalize stochastic approximation methods for unconstrained problems to problems involving constraints and nondifferentiable objective functions. These methods are driven by sample gradients (or sample subgradient). By virtue of using only statistical estimates of gradients or subgradients, SQG methods can be applied to complex problems where analytical approximation methods and scenario methods failed to work. Ermoliev [1968] studies a stochastic analog of deterministic projection methods in the context of constrained stochastic optimization and proves the convergence. Gupal and Bazhenov [1972] propose a stochastic linearization method and try to solve a convex stochastic program by solving a sequence

of linear programs. Ruszczynski [1986] extends the original stochastic linearization method by imposing a quadratic term in the linearization subproblem, attempting to produce a more well-behaved sequence of solutions during the iterations. Higle and Sen [1991] combined the cutting plane algorithm with SQG methods and develop the stochastic decomposition method. Recently, Culioli and Cohen [1990] proposed an auxiliary function method, which extends the auxiliary principle of Cohen [1980] for deterministic optimization to stochastic optimization. The penalty method of Pflug [1982] puts the constraints set to the objective function via some convex differentiable functions, turning a constrained problem into an unconstrained problem.
Two major methods of SQG methods are the stochastic projection method, the stochastic linearization method. We present those two methods in the following sections.

### 3.4.1 The Projection Method

Consider the following minimization problem:

$$\min_{x \in \mathcal{X}} F(x) = E_\omega f(x, \omega) \tag{3.26}$$

Let $\Pi_{\mathcal{X}} : \mathcal{R}^n \mapsto \mathcal{X}$ be a projection operator which maps a point in $\mathcal{R}^n$ to the feasible set $\mathcal{X}$. Then, given an initial point $x^0$, a stochastic projection method generates the sequence $x^1, x^2, \ldots,$ according to:

$$x^{k+1} = \Pi_{\mathcal{X}}(x^k - a_k d^k(x^k, \omega^k)) \tag{3.27}$$

If we choose the step direction $d^k(x^k, \omega^k)$ to be stochastic generalized gradient or quasigradient of $F(x)$ at $x^k$, then we have

$$E_\omega(g^k(x^k, \omega^k) \mid x^0, \ldots, x^k) = \hat{F}(x^k) + b_k \tag{3.28}$$

then the sequence $x^k$ generated by algorithm ( 3.27) converges to $x^*$ with probability one under the following conditions:

(1) $F(x)$ is a convex and continuous,

(2) $\mathcal{X}$ is a convex and compact set ,

(3) The stepsize $a_k$ satisfy :

$$\sum_{k=1}^{\infty} a_k = \infty, \sum_{k=1}^{\infty} a_k^2 < \infty, \text{ and } \sum_{k=1}^{\infty} E\{a_k \parallel b_k \parallel + a_k^2\} < \infty$$

(4)$E\{\parallel g^k(x^k, \omega^k) \mid x^0, \ldots, x^k \parallel^2\} \leq C$ , where $C$ is a constant.

When choose $d^k(x^k, \omega^k) = g^k(x^k, \omega^k)$, it may quite different from iteration to iteration since different samples are usually chosen at each iteration, producing an oscillated sequence $\{x^k\}$. An alternative way is to choose step direction as follows:

$$d^k(x^k, \omega^k) = \gamma_k g^k(x^k, \omega^k) - (1 - \gamma_k) d^{k-1}(x^{k-1}, \omega^{k-1}) \tag{3.29}$$

where $0 \leq \gamma_k \leq 1$, and

$$\frac{\gamma_k}{a_k} \to 0 \text{ and } \sum_{k=1}^{\infty} \gamma_k^2 < \infty$$

Using step direction given by equation ( 3.29) usually accelerates the convergence of the algorithm since it captures the momentum of successive stochastic gradients.

### 3.4.2 The Linearization Method

One of the first order method of SQG is the stochastic linearization method. Let $F(x)$ be continuously differentiable, consider the optimization problem defined by ( 3.26), then the standard linearization method is given as follows

**Step 1** choose an initial $x^0 \in \mathcal{X}$

**Step 2** at iteration $k$, find a subgradient $g^k(x^k, \omega^k)$

**Step 3** let

$$v^k = (1 - \gamma_k) v^{k-1} + \gamma_k g^k(x^k, \omega^k) \tag{3.30}$$

**Step 4** find

$$\bar{x}^k = arg \min_{x \in \mathcal{X}} v^k x \tag{3.31}$$

Let the feasible direction be

$$d^k(x^k, \omega^k) = x^k - \bar{x}^k \tag{3.32}$$

*Step 5* update

$$x^{k+1} = x^k + a_k d^k(x^k, \omega^k) \tag{3.33}$$

The sequence $x^k$ generated by this method converges to $x^*$ under the following conditions:

(1) $F(x)$ be continuously differentiable,

(2) $\mathcal{X}$ is a convex and compact set ,

(3) The stepsize $a_k$ satisfy :

$$\sum_{k=1}^{\infty} a_k = \infty,\ \sum_{k=1}^{\infty} a_k^2 < \infty, \text{ and } \sum_{k=1}^{\infty} E\{a_k \parallel b_k \parallel + a_k^2\} < \infty$$

$$\gamma_k > 0,\ \frac{\gamma_k}{a_k} \to 0 \text{ and } \sum_{k=1}^{\infty} \gamma_k^2 < \infty$$

(4)$E\{\parallel g^k(x^k, \omega^k) \mid x^0, \ldots, x^k \parallel^2\} \leq C$ , where $C$ is a constant.

This method has been extended to solve nondifferentiable problems with $F(x)$ satisfying local Lipschitz conditions (see Ermoliev and Gupal [1978] and Ruszczynski [1986]). Ruszczynski [1986] proposed a method for solving problems with nondifferentiable nonconvex objective functions. The method is given as follows:

**Step 1** choose an initial $x^0 \in \mathcal{X}$

**Step 2** at iteration $k$, find a stochastic generalized gradient $g^k(x^k, \omega^k)$, i.e.

$$E_\omega(g^k(x^k, \omega^k) \mid x^0, \ldots, x^k) = \hat{F}(x^k) \tag{3.34}$$

**Step 3** let

$$v^k = (1 - \gamma_k)v^{k-1} + \gamma_k g^k(x^k, \omega^k) \tag{3.35}$$

**Step 4** find

$$\bar{x}^k = arg \min_{x \in \mathcal{X}} v^k(x - x^k) + \frac{1}{2} \parallel x - x^k \parallel^2 \tag{3.36}$$

Let the feasible direction be

$$d^k(x^k, \omega^k) = x^k - \bar{x}^k \tag{3.37}$$

*Step 5* update

$$x^{k+1} = x^k + a_k d^k(x^k, \omega^k) \tag{3.38}$$

The sequence $x^k$ generated by this method converges to $x^*$ with the following conditions:

(1) $F(x)$ may not be differentiable and convex. $F(x)$ is Lipschitz continuous on $\mathcal{S} \supset \mathcal{X}$ and its subgradient has the following property: there exist a constant $\sigma$ such that for every $x \in \mathcal{X}$ and every subgradient of $F(x)$, one has

$$F(y) - F(x) \geq F\hat{(}x)(y - x) - \sigma \parallel y - x \parallel^2 \tag{3.39}$$

(2) $\mathcal{X}$ is a convex and compact set ,

(3) The stepsize $a_k$ satisfy :

$\sum_{k=1}^{\infty} a_k = \infty$, $\sum_{k=1}^{\infty} a_k^2 < \infty$, and $0 < a_k < min(1, \frac{1}{a})$ where $a$ is positive constant . $E\{\parallel g^k(x^k, \omega^k) \parallel^2\} \leq C$

$$\gamma_k = a_k a \tag{3.40}$$

The major difference is in the direction finding subproblem, we need to solve a quadratic programming problem instead a linear programming when $F(x)$ is not differentiable.

## 3.5 Application of SQG to the Stochastic Container Allocation Model

The calculation of an exact gradient involves first finding the distribution of $f(x, \omega)$ as a function of $x$ and then performing multi-dimensional integrations. However, usually we only know the values of $f(x, \omega)$ for each particular $x$, not the function form. Even if we know the function form, the calculation of the exact gradient is computational tractable for a few special cases. In general, statistical estimates of gradients or subgradients are much easier to find.

Consider the formulation of the stochastic allocation model (given by ( 3.1) and ( 3.2), which is a two stage stochastic program with recourse :

$$\min_{x \in \mathcal{X}} F(x) = E_\omega \{cx + Q(x, \omega)\} \tag{3.41}$$

where

$$\mathcal{X} = \{x | Ax = b, 0 \leq x \leq u\} \tag{3.42}$$

$$Q(x, \omega) = \min_y q(\omega)^T y \tag{3.43}$$

subject to

$$\begin{aligned} W(\omega)y(\omega) &= h(\omega) - T(\omega)x \\ -y(\omega) &\geq -\xi(\omega) \\ y(\omega) &\geq 0 \end{aligned} \tag{3.44}$$

For a given outcome $\omega^k$ and stage one decision $x^k$, the dual of the recourse problem is given by

$$\max_{\pi(\omega), \alpha(\omega)} \pi(\omega) \{h(\omega) - T(\omega)x\} - \alpha(\omega)\xi(\omega) \tag{3.45}$$

subject to

$$\pi(\omega)^T W(\omega) - \alpha(\omega)^T \leq q(\omega)$$
$$\alpha(\omega) \geq 0 \quad \pi(\omega) \quad \textit{unrestricted}$$

By duality, we have for given $(x^k, \omega^k)$, and for any other $x$ :

$$Q(x, \omega^k) - Q(x^k, \omega^k) \geq (x - x^k)^T \{-T(\omega)^T \pi(x^k, \omega^k)^T\}$$

Then a stochastic subgradient $g^k(x^k, \omega^k)$ is given by :

$$g^k(x^k, \omega^k) = c - T(\omega)^T \pi(x^k, \omega^k)^T \tag{3.46}$$

It can be shown that under any reasonable choice of $\pi(x^k, \omega^k)$ ( see Ermoliev [1983]), we have

$$E\{g^k(x^k, \omega^k) \mid x^0, \ldots x^k\} = \hat{F}(x^k) \tag{3.47}$$

The objective function $F(x)$ is convex and nondifferentiable in general, since the minimization operator is present under the integral sign. A statistical estimate of $F(x^k)$ can be obtained by solving the recourse problem for given $x^k$ i.e.

$$f(x^k, \omega^k) = cx^k + Q(x^k, \omega^k) \tag{3.48}$$

To calculate the exact value of $F(x^k)$, we need to find the distribution of $Q(x^k, \omega^k)$ and get the expected recourse function $E\{Q(x^k, \omega^k)\}$, which is only numerically feasible in few special cases. This is also true for the calculation of the subgradients or gradients of $F(x^k)$.

The peculiarity of SQG methods is their highly oscillatory behavior. It is difficult to judge whether the algorithm has already approached a neighborhood of the optimal point since the exact values of the objective function is difficult to obtain. Another property of the algorithm is its slow convergence. The asymptotic rate of convergence is given by (Gaivoronski [1988]):

$$e(x^k) = \| x^* - x^k \| \leq \frac{C}{\sqrt{M}} \tag{3.49}$$

where $M$ is the number of samples $\omega$ chosen so far. It usually depends on the number of iterations $k$. However, no other algorithm can do better than this for stochastic program defined by ( 3.1) and ( 3.2) (see Gaivoronski [1988]).

For implementation of SQG methods, there two major tasks, i.e. choosing a step direction $d^k$ and a step size $a_k$. We discuss these two tasks respectively as follows :

**Step Direction Chioces**

First is to get statistical estimates of gradients or subgradient. There are several ways of obtaining these estimates, here we present two of them.

*1. Sample gradients*

For a two-stage stochastic programming problem with recourse, it is easy to get a sample of subgradient which is given equation ( 3.46). We can directly use this estimate in the projection or lineraization methods described above. To reduce the variance of estimation, we can use the sample mean given by :

$$\bar{g}^k(x^k) = \frac{1}{N} \sum_{i=1}^{N} g^k(x^k, \omega^k) \tag{3.50}$$

as our estimation. It may stabilize step directions among successive iterations, but it is more expensive computational. In this case, it is not necessary to chose $\gamma_k \to 0$ as $k \to \infty$, because we have convergence for any $0 \leq \gamma_k \leq 1$.

In cases where the calculation of the sample gradients is computational expensive or impossible, we can rely on approximations of the sample quasigradients or subgradients.

*2. Finite-difference Approximation*

When $F(x)$ is differentiable, one possibility is to use forward finite differences

$$g^s = \sum_{i=1}^{n} \frac{f(x_s + \delta_s e_i, \omega_{i,1}^s) - f(x^s, \omega_{i,2}^s)}{\delta_s} e_i \tag{3.51}$$

In order to ensure convergence with probability one, it is sufficient to take any sequence $\delta_s$ such that $\sum_{i=1}^{\infty} a_i^2/\delta_i^2 < \infty$. If it is possible to take $\omega_{i,1}^s = \omega_{i,2}^s$, then any sequence such that $\delta_s \to 0$ will ensure convergence with probability one.

When $F(x)$ is nondifferentiable, Gupal [1979] proposed the following subgradient estimation

$$g^s = \sum_{i=1}^{n} \frac{f(x_s + \delta_s e_i, \omega_{i,1}^s) - f(x^s, \omega_{i,2}^s)}{\delta_s} e_i \tag{3.52}$$

$\gamma_s$ and $\delta_s$ should be chosen such that $\frac{\gamma_s}{\delta_s} \to \infty$ and $\frac{(\gamma_s - \gamma_{s+1})}{a_s} \to 0$ to ensure convergence with probability one.

Once the estimate of qusigradient or subgradient is obtained, the step direction can be determined by either using the $g^k(x^k, \omega^k)$ directly or using the successive averaging procedure described by (( 3.29)). The successive averaging procedure allows us to use information obtained at previous iterations instead of only information at iteration $k$. This is one way of smoothing out the oscillations inherent in the the sample estimate of gradients or subgradients. Methods of this type may be viewed as stochastic analogues of conjugate methods and they were first proposed by Gupal and Basenov [1972]. For $F(x)$ is differentiable, it is shown that $\| v^k - \nabla F(x^k) \| \to 0$ under very general conditions. For $F(x)$ is nondifferentiable, we need to introduce the smoothing procedure for the objective function [see Gaivoronski [1988] or use the method proposed by Ruszczynski [1986].

### Stepsize Choices

For the sequence $x^k$ generated by SQG to converge, the stepsize $a_k$ has to satisfy certain conditions (see condition (3) for the projection method and the linearization method). Clearly $a_k = \frac{C}{k}$ satisfies the conditions, but this choice of step size does not utilize the information accumulates during previous iterations. Experiments show that using this choice, the SQG gives poor results most of time. When initial choose $x^0$ is far away from the optimal solution $x^*$, the step size may becomes too small before $x^k$ reaches the vincity of $x^*$, causing the iterations stalled. If the initial choose $x^0$ is close the optimal solution $x^*$, the step size may becomes too large initially and making the solutions highly oscillated or even move farther away from $x^*$ than $x^0$. A better strategy is to choose the step size adaptively based on the performance measures gathered during the iterations. For example, estimate of objective function value can be valid performance measure. When estimate of objective value oscillates in large extent, then the stepsize may be too large, while the estimates change slowly, the stepsize is usually too small. If the estimates of $F(x^k)$ decreases steadily (meaning the algorithm behaves regularly), the step size is just right. In the following, we describe several performance measures that can be used as criteriain for adjusting the step sizes.

*1. Estimates of Objective Function Values*

Estimates of $F(x^k) = E\{f(x^k, \omega)\}$ can be obtained by following equation:

$$\tilde{F}(x^k) = \frac{1}{k}\sum_{j=0}^{k} f(x^k, \omega^j) \tag{3.53}$$

By choosing one sample per iteration and averaging over all iterations , $\tilde{F}(x^k)$ is a better stable estimates than by just using $f(x^k, \omega^j)$.

*2. Ratio of Changes in $F(x^k)$ and Changes in $x^k$*

Let $p(x^k, m)$ be a measure of relative changes in $F(x)$ made during the previous $m$ iterations. The stepaize rule can be chosen adaptively as follows for some predetermined parameters $\delta$ and $\alpha$.

$$a_{k+1} = \begin{cases} \alpha a_k & \text{if } p(x^k, m) < \delta \\ a_k & \text{otherwise} \end{cases} \tag{3.54}$$

where $0 < \alpha \leq 1$, $m > 1$ and the ratio $p(x^k, m)$ is given by :

$$p(x^k, m) = \frac{\tilde{F}(x^{k-m}) - \tilde{F}(x^k)}{\sum_{j=k-m}^{k-1} \| x^{j+1} - x^j \|} \tag{3.55}$$

This stepsize rule says that if the progress made during previous $m$ iterations is less than $\delta$, then we reduce the stepsize by a factor of $\alpha$, otherwise we keep the same stepsize.

*3. Inner Product of Two Consecutive Generalized Gradient Estimates*

The decision as to whether to change the step size may be based on the inner product of $g^{k-1}$ and $g^k$. If they form an acute angle, then it may indicate regular behavior prevail over oscillated behavior. One the other hand, these generalized gradient may fluctuate significantly in the neighborhood of the optimal solution. Kesten [1958] suggests the following step size rule.

$$a_k = \frac{1}{t_k} \tag{3.56}$$

where

$$t_k = \begin{cases} k & \text{if } k = 1, 2 \\ 2 + \sum_{n=3}^{k} I_{\{(g^{k-1})^T g^k \leq 0\}} & \text{if } k \geq 3 \end{cases} \tag{3.57}$$

where $I_{\{\cdot\}}$ is an indicator variable. This rule says that step size is reduced if the inner product of two consecutive gradients shows a negative sign. Andraottir [1990] shows this rule can speed up convergence drastically together with normalized step direction choice for stochastic approximation method.

We choose to apply the stochastic linearization method to the stochastic container allocation problem in this thesis. The successive averaging procedure is adopted for the step direction, and for the step size the inner product rule is used. In the next chapter, numerical experiements are presented and comparison of deterministic model with stochastic model is made based on the results of nemerical experiements.

# Chapter 4

# Experiment Design and Numerical Results

In order to evaluate the effectiveness of different container allocation models, an evaluation procedure need to be established. This procedure consists of setting up testing problems, conducting rolling horizon simulation and comparing the performance measures obtained from running each model. In this chapter, we first describe the testing problems, then present the rolling horizon simulation procedure and the performance measures.

## 4.1 Testing Problems

Testing problems are generated by a random problem generator. Two models for container allocation presented in chapter two, i.e. the deterministic model with empty flow and the stochastic model with empty flow presented in section 2.5.3, are evaluated using those testing problem so that we can make comparison under various conditions. The parameters for the random problem generator are the number of depots, length of planning horizon, the transportation , inventory and leasing costs, the coefficients for available traffic routes, and the demands and supplies of empty containers. The random problem generator creates testing problems according to the the following procedures:

**The Locations of Depots:** It first generates $R$ depots which are uniformly located in a 2500 by 2500 miles square. We simply take the Euclidean distance between each pair of depots as the corresponding travel distance. The average speed of a ship is assumed to be 500 miles per period and average speed for a train or a truck is assumed to be 800 mile per period. The travel time between two depots approximately equal to the integer value of the ratio of distance to speed.

**Customer Demand and External Supply of Empty Containers:** The generator assigns empty container demands and supplies at each depot on each day over the planning horizon. For the demand for empty containers at depot $j$ on period $t$, it assumed to follow Poison distribution (we can assume other probability distributions for the demand) with a mean value $\xi_j(t)$ which is given by :

$$\xi_j(t) = \beta_j \cdot \gamma_t \cdot \nu \qquad (4.1)$$

where

$\beta_j$ = outbound potential for depot $j$. The outbound potential for each depot capture the region's ability to generate the outbound loaded flows, i.e. the demand for empty containers at this depot. The value of $\beta_j$, is drawn uniformly between 0.2 and 1.8.

$\nu$ = an exponential random variable representing the average demand for empty containers over all depots in the system.

For the external supply at a depot $j$ on period $t$, we assume that it follows Poison distribution with a mean value $\eta_j(t)$ which is given by

$$\eta_j(t) = \alpha_j \cdot \gamma_t \cdot p_j \cdot \nu \qquad (4.2)$$

where

$\alpha_i$ = inbound potential for depot $i$. The inbound potential represents the depot's capability of attracting inbound loaded flow to the region around this depot, equivalently reflect the potential of for external empty supplies.

$p_j$ = a random variable uniformly distributed over 0 to 0.5 (this number is chosen arbitrarily), which represents the percentage of leased containers of all inbound loaded containers into depot $j$ that is returning to their leasing companies.

$\gamma_t$ = a daily variation factor which represents weekly pattern for demands and supplies : high demands and supplies in weekdays and low demands and supplies in weekends, for $t = 0, 2, ..., 6$.

$\nu$ = an exponential random variable representing the average external supplies for empty containers over all depots in the system.

In real world applications, regions with large inbound flows often have small outbound flows. Here, we set

$$\alpha_i = 2 - \beta_i \tag{4.3}$$

Therefore, these two potentials are negatively correlated. An important reason for this kind setting is that a myopic solution may produce a poor solution since a container may sent to a depot with high inbound potential but with very low outbound potential.

**Update Procedures for Demands and External Supplies and Route Capacities:** At any period $t$, the generator generates deterministic demands and external supplies for all depots at period $t$, and stochastic demands and supplies for all depots from period $t+1$ to period $t+N-1$. Then running the alternative models on this random generated network, and implementing only the distribution decisions at period $t$, the model updates the internal supplies of empty containers, and move the time to $t+1$ period. At period $t+1$, it generates deterministic demands and external supplies for all depots at period $t+1$ by drawing a outcome from the stochastic demands and supplies generated in period $t$, and stochastic demands and supplies for all depots from period $t+2$ to period $t+N$ *according to weighted average* of the demands and supplies for the same depots and corresponding to the same time period *but* generated in *previous* time periods. All models or methods on the same model are compared under a homogeneous environment.

**Available Traffic Routes :** Shipping companies and railroads usually offer a fixed schedule over each week. However for trucking companies, it is not the case.

In our case, since ship and rail routes consist of majority of the all available routes. Here we assume that available traffic routes are same during each week. To generate routes over a typical week, first we need to introduce a route density coefficient $d \in [0, D]$, representing the average number of routes emanenting out of a typical depot. It is used to control the arc density of the resulting dynamic network. For a low value of $d$, the future impacts by current decisions are accentuated, meaning that "bad" current decisions may induce more serious consequences, because with a sparse network, corrective actions are hard to make. The number of routes is determined by multiplying the density coefficient $d$ and number of depots in the system. To create each route, the generator chooses its origin uniformly among all depots, and the destination of the route over the rest of depots. Then the corresponding time period of this route leaving its origin is chosen uniformly over the time periods during week. The travel time is determined by the distance between its origin depot and destination depot and the average speeds. Each route is represented by an arc in the underlying dynamic network. Each route is then assigned as ship route with 0.5 probability, and as a rail route with 0.5 probability unless the travel time of this route is greater than 4 time periods (this value is chosen arbitrarily).

There is capacity limit for each ship route, but there is no such limit on a rail route. The reason is that railroads usually can accept any demand for each pair of origin and destination because they can run several trains between higher demand traffic lanes, whereas shipping companies usually only assign one ship on each route. For ship route, loaded containers have higher priority and the remaining capacity of the ship is used for moving empty containers. We choose the capacity of a ship, denoted by $v$, uniformly over the intervals lower and upper bounds of a ship capacity provided by the input data. For a particular route originates from depot $i$ at time period $t$ and terminates at depot $j$, its mean capacity of this ship route $u_{ij}(t)$ is given by

$$u_{ij}(t) = (1 - \beta_i \alpha_j \gamma_t) \cdot \nu \tag{4.4}$$

where

$\nu$ = an exponential random variable representing the average ship's capacity.

$\beta_i$ ,$\alpha_j$, $\gamma_t$ are defined as before.

Once the routes over the first week have been determined, we just repeat those traffic routes on following weeks until the end of planning horizon, because it is a usual practice for shipping lines and railroads to offer a relatively fixed schedule on each week.

## 4.2 Performance Measures

The performance of a model is the total costs of empty distribution and short-term leasing over a sufficient long simulation period of time when different models are employed as a decision support tool. However, we also solve a deterministic dynamic network with the $T$ simulation periods after all random demands are revealed. The total cost obtained from this problem is called the *posterior bound* (PLB) which is generally unreachable since no uncertainty is involved, and decisions in period $t$ are allowed to anticipate future events exactly. The percentage gap between the total cost of each model with respect to the PLB is an important performance measure. Additional measures can be the total empty container-miles over the whole simulation period and the percentage of time a container traveled empty over the average turnaround time of a container.

Finally, the total computation time for each model is another performance measure which indicates whether the model can be used as a real-time dispatching tool. These times include only the processing time needed to solve the model and does not include the time for updating input data.

## 4.3 Rolling Horizon Simulation

To evaluate the performance of a model or a solution method, we take a rolling horizon simulation approach. Let $N$ be the length of the planning horizon and $T$ be the length of the rolling horizon simulation. At any time period $t$, we draw a realization of demands and supplies of empty containers ( for the first stage it is deterministic, and for the remaining $N-1$ stages they are random variables ). Then solve the $N$ periods problem starting at time $t$ and its recommended actions are identified from the model outputs. However, we only implement the recommended actions for time $t$, and then we advance the clock to time period $t+1$ and solve the corresponding $N$ period problem starting at time period $t+1$. We repeat the above procedures for $t = 1, ..., T$. It is important to choose $T$ sufficiently large to capture the dynamic effects and to mitigate statistical sampling errors. As the rolling horizon simulation proceeds,the costs and other performance measures of the implemented recommended actions of the model are accumulated. When the rolling horizon simulation terminates, the total empty movement cost and leasing cost obtained by the model serve as evaluating criterion for the performances of those models for container allocation.

## 4.4 Evaluation of Altenative Models

In this section, we evaluate the performance of two alternative models for empty container allocation, i.e. the deterministic model and stochastic model. Both models are running on the same set of randomly generated testing problems described in section 4.1 over $T$ simulation periods. The outputs of each model on each time period, i.e. the container allocation decisions are implemented for the current time period only. The consequences of the allocation decisions, i.e. the costs incurred, number of empty containers leased and the gaps of those costs to the posterior lower bound. In the following sections we describe the implementation of these two models and present results of our experiements.

### 4.4.1 The Dynamic Deterministic Model

As we described in section 4.1, at time period $t$, only the demands and supplies and route capacities at this period are known with certainty. For all those inputs in remaining time periods of the current planning horizon, we only have their probability distributions. One possible approximation of those random elements is to use the their expected values. After substituting those random elements by their expected values, our container allocation becomes a model with all inputs known with certainty. Here we call this approximate model as "deterministic model". The underlying optimization problem is a minimum cost flow problem as shown in chapter two and we can solve it by classical network optimization techniques. In particular, the network simplex method is used here.

In our problem generator, we assume that all random elements are independent and discretely distributed, and their probability mass functions (PMF) assumed to be Poison distributions. Their expected values may not be integers, so we may not have all integer inputs when we solve the deterministic model. The solutions to the model may not be integer values. In practice, we cannot implement a noninteger solution as container allocation decisions, for example, we cannot send 1.75 containers from depot $A$ to depot $B$. The following roundup techniques are employed to get the approximate integer solutions. Let

$x_{ij}$ be optimal flow from depot $i$ to depot $j$ obtained from the solving the model.

$\tilde{x}_{ij}$ be the integer approximation of $x_{ij}$.

$$\tilde{x}_{ij} = \lfloor x_{ij} + \gamma_{ij} \rfloor \tag{4.5}$$

such that

$$\sum_{j \in A_i} \tilde{x}_{ij} = \sum_{j \in A_i} x_{ij} \tag{4.6}$$

where $0 \leq \gamma_{ij} \leq 1$.

### 4.4.2 The Dynamic Stochastic Model

The stochastic model formulation is a two-stage stochastic program with network recourse. The stochastic linearization method is employed to solve the model. Here we assume that all random elements from time periods $t = 2, 3, \ldots, N$ reveal their outcomes at the $t = 2$. The solutions obtained may also be noninteger solutions when using the stochastic linearization method, we use the same round up technique described in the last sections to get the approximate integer solution.

Step size is chosen adaptively and changes of step size is according to inner product rule given by ( 3.56) and ( 3.57). In each iteration, we use 6 samples to get an estimate of generalized gradient (see equation ( 3.50)). For step direction, the successive average procedure is adopted. The $\gamma_k$ is chosen to be equal to 0.02 as recommended by Gaivoronski [1988]. Therefore, in each iteration, six plus one minimum cost flow problems are solved. If on average, 40 iterations is needed to get a solution to the stochastic model, then the computation time is just approximately equal to the time for solving 280 minimum cost flow problems. This means that it is computational feasible for the stochastic model to be used as real-time decision supporting tool.

To see how the stochastic linearization behaves, we choose 3 cases and plot the estimates of objective function values over 40 iterations (see Figures $(4-1)$,$(4-2)$ and $(4-3)$). Two estimates of $F(x)$ are collected, they are defined as follows: Estimate one is the average over all previous iterations by taking one sample value $f(x,\omega)$ at each iteration, while estimate two is obtained by averaging over past 8 iterations. The solid curves present the estimate one, and the dotted curves represent estimate two. In the first 8 iterations, they are exactly equal, whereas in following iterations, estimate one is more stable estimate than estimate one. However, estimate two is in general a better estimate since what we need is the objective function value at current iteration $x^k$. From the figures, we find that for all three cases, the objective function values improve rapidly during the first few iterations, wheras later on they oscillates around a certain value. For the algorithm to converge with probability one, it is required to use diminishing step size given in chapter three. However, as we discussed in chapter three, if step size decreases too quickly, then algorithm may stall when $x^k$ is still far away from optimal. In our experiments, we use the inner product step size rule given by ( 3.56) and ( 3.57).

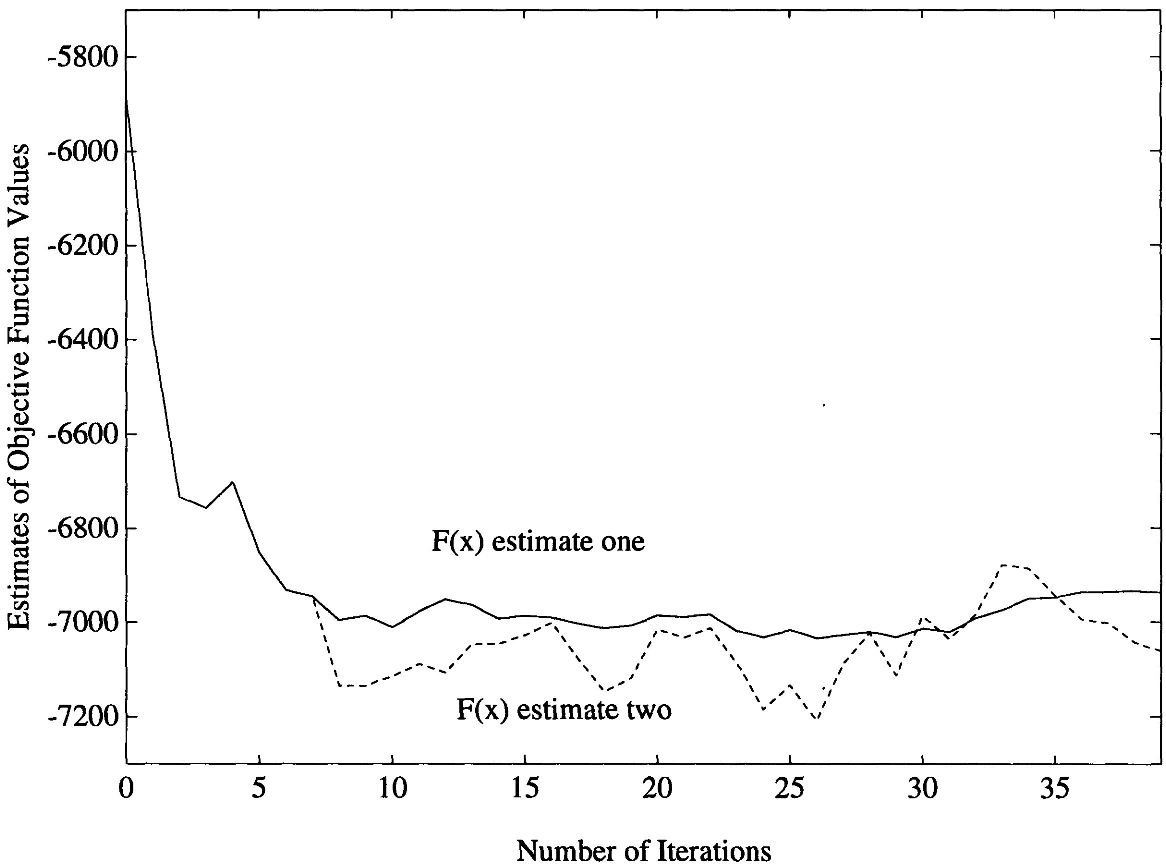


Figure 4-1: Changes of the estimate of $F(x)$ at different iterations: case one

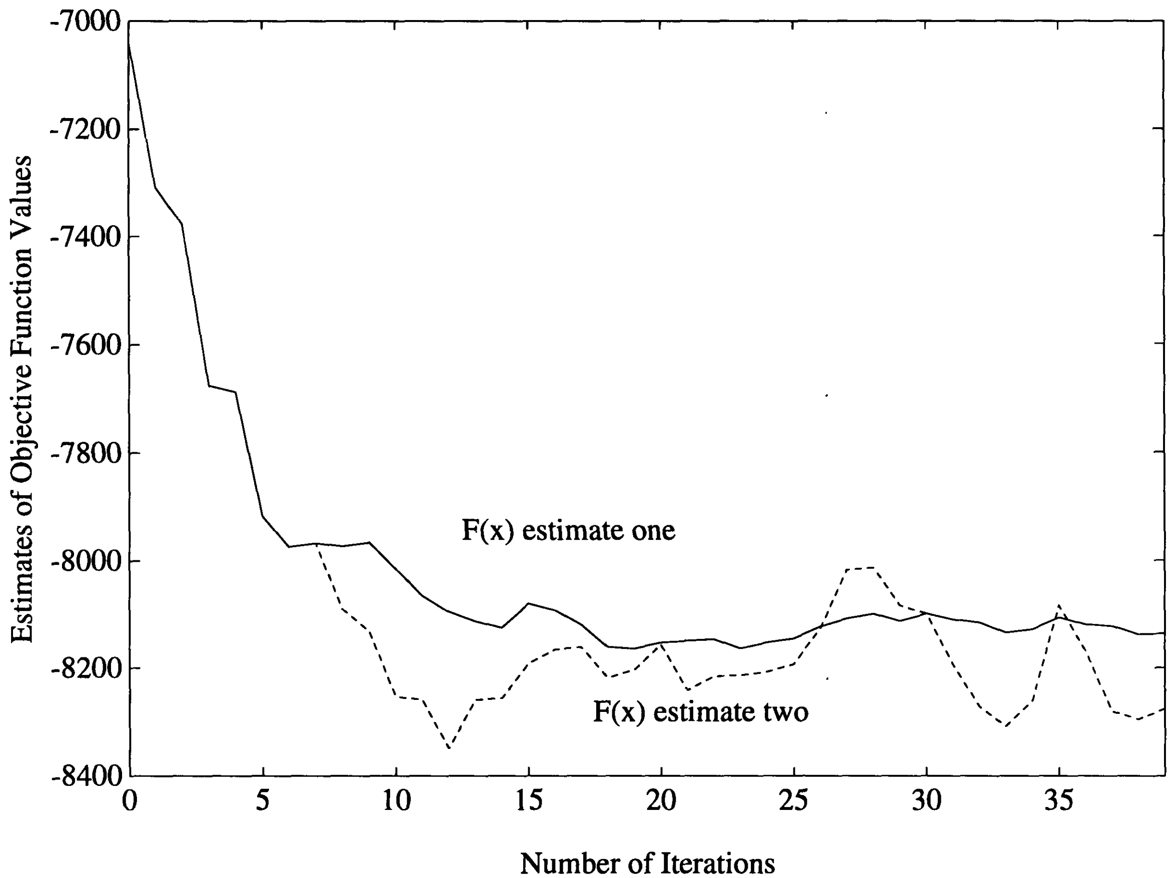


Figure 4-2: Changes of the estimate of $F(x)$ at different iterations: case two

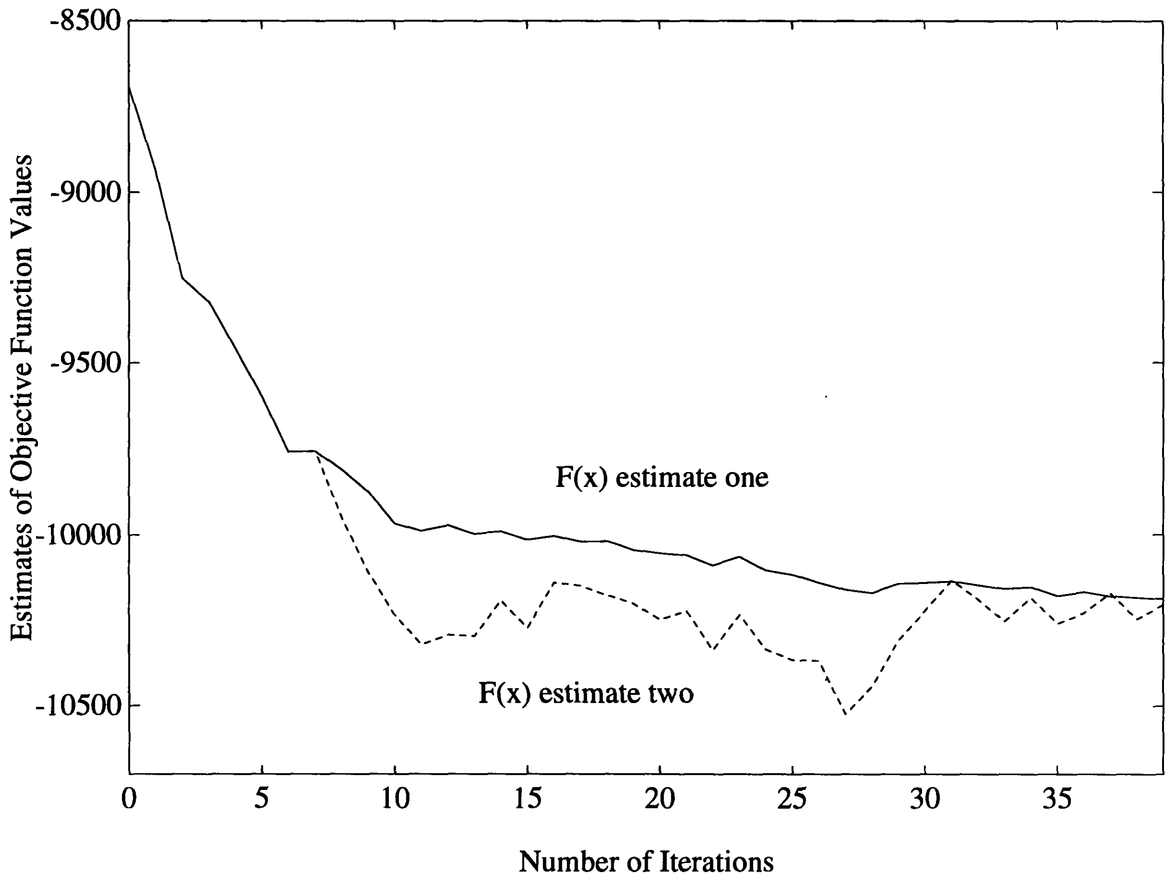


Figure 4-3: Changes of the estimate of $F(x)$ at different iterations: case three

Table 4.1: Parameters of Testing Problems

| Problems | P1 | P2 | P3 | P4 | P5 | P6 |
|---|---|---|---|---|---|---|
| Number of Depots | 15 | 16 | 17 | 18 | 19 | 20 |
| Planning Horizon | 7 days | 7 days | 7 days | 7 days | 7 days | 7 days |
| Simulation Time | 14 days | 14 days | 14 days | 14 days | 14 days | 14 days |
| Mean Supply | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 | 3.0 |
| Mean Demand | 5.0 | 5.0 | 5.0 | 5.0 | 5.0 | 5.0 |

### 4.4.3 Experiemental Results

**The Testing Problems:**

Six different sizes of testing problems are generated with 15 depots to 20 depots respectively. All testing problems has same length of planning horizon and length of simulation time. For simplicity, a time period is chosen to be one day. The problem parameters are described in Table 4.1.

**The Experimental Results :**

The objective of our experiments is to evaluate the performances of above two different models, i.e. deterministic model and stochastic model. Our major performance measure is the gap between total cost of each model and the *posterior lower bound* of total cost over the simulation period. This gap is expressed in percentage term, which is given by:

$$g_i = \frac{TC_i - PLB}{PLB} \times 100\% \qquad (4.7)$$

where

$TC_i$ = total cost incurred when using model $i$ over $T$ period simulation,

$PLB$ = posterior lower bound of total cost over the simulation period.

For each testing problem, i.e. $P1$, ..., $P6$ described in table 4.1, 10 different samples are generated and the sample mean of total costs incurred for each problem size are presented in the following table. The data on above table are plotted in Figiure $4-4$. From the figure, we find only one of six different problem sizes, the average costs of 10 samples for stochastic model is higher than that for deterministic

Table 4.2: Percentage Gaps of Deterministic and Stochastic Models

| Problems | Deterministic Model | Stochastic Model |
|---|---|---|
| P1 | 9.2570 | 8.3452 |
| P2 | 8.1909 | 8.4936 |
| P3 | 8.8882 | 7.7929 |
| P4 | 8.8320 | 8.3185 |
| P5 | 8.4073 | 8.2624 |
| P6 | 8.3373 | 7.3691 |

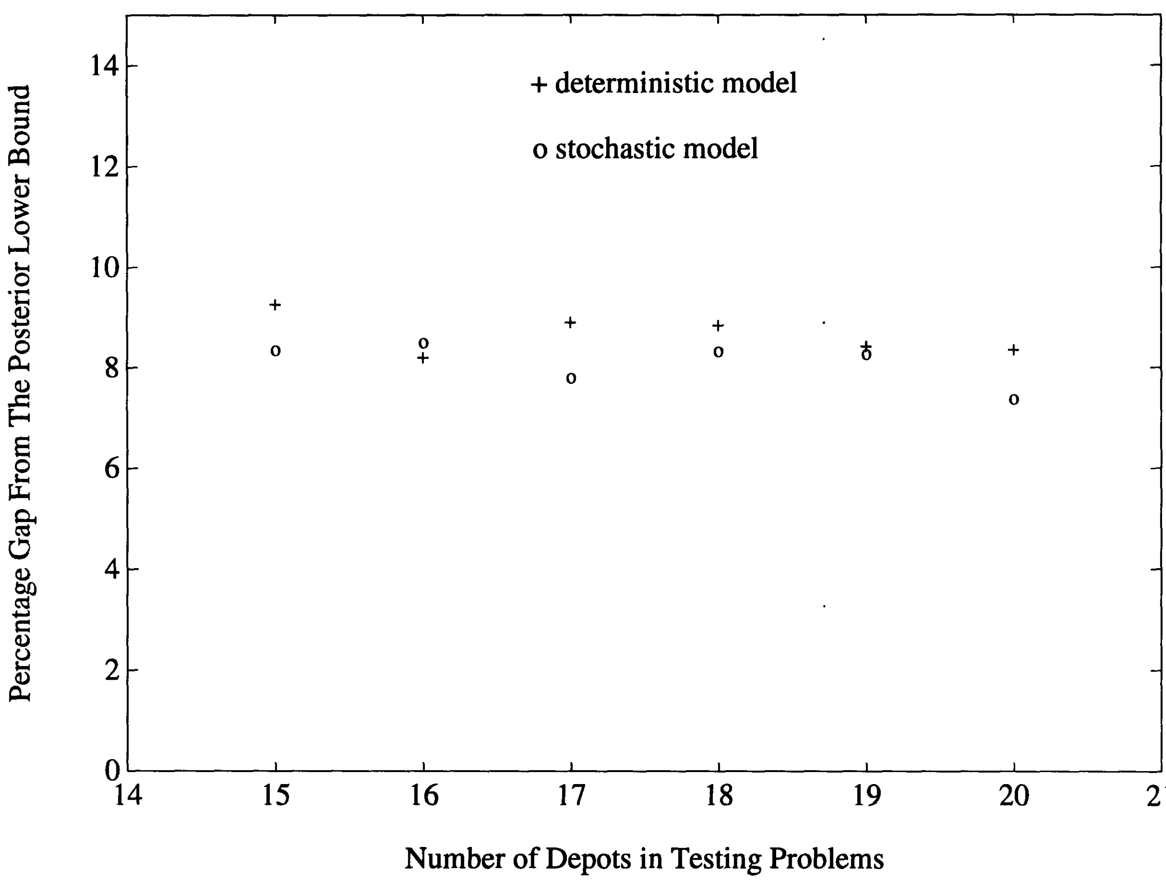


Figure 4-4: Comparison of the performance of the deterministic model and the stochastic model

model. Out of 60 samples, only in 10 samples the total cost of stochastic model is higher than that of deterministic model. In Figure $4-4$, all points are about 8% to 10% above the posterior lower bound. This is similar to the results obtained in Chueng and Powell [1992] for DVA problem for allocating truck motors. In general, stochastic model takes more information of future events into account, i.e. the distributions of the future random events, while the deterministic model only use the expected values. The stochastic model is expected to produce better solutions in terms of average total costs incurred than the deterministic on average.

Figure $4-4$ shows closeness of total costs of the two alternative models, it may due to following reasons :

1. The deterministic model is an approximation of the stochastic model. In fact, if we can obtained the expected recourse function $\bar{Q}(x) = E\{Q(x,\omega)\}$ of the stochastic model in terms of first stage decision variables $x$, then the solution of the stochastic model can be obtained by solving the following deterministic optimization problem :

$$
\begin{aligned}
z = \min_{x} \; & cx + \bar{Q}(x) \\
s.t. \quad & Ax = b \\
& 0 \leq x \leq u
\end{aligned}
\tag{4.8}
$$

   In the deterministic model, we simple use the minimum cost flow problem over the stage two network to approximate the expected recourse function. We denote this approximation by $\tilde{Q}(x)$. The expected values of future random events are used in this minimum cost flow problem. Both the exact $\bar{Q}(x)$ and the approximate one $\tilde{Q}(x)$ in deterministic model are piecewise linear functions in $x$. Since our problem generator generates coefficients of networks uniformly over all nodes and arcs, so that $\tilde{Q}(x)$ may be very close to $\bar{Q}(x)$ in term of the shape of the two functions. This fact is also reflected in Figure 4.3 in the costs curves of two models over simulation periods.

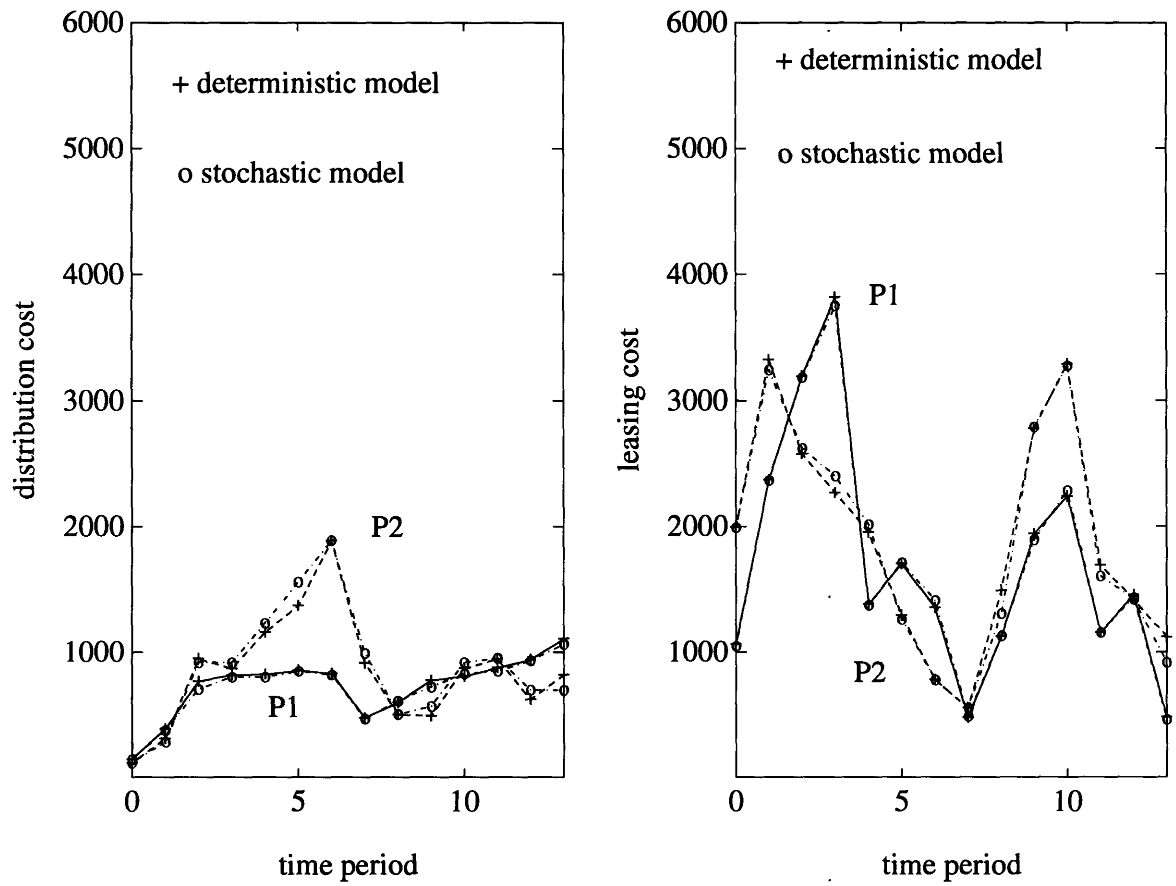


Figure 4-5: The distribution costs and leasing costs for testing problem P1 and P2

2. In the stochastic model, for each problem, only 40 iterations is conducted since it takes longer time to conduct more iterations.

In our experiements, we also collect the following performance measures for each model during the rolling horizon simulation.

1. total costs of distributing empty containers over the 14 days simulation time.

2. total costs of leasing empty containers over the 14 days simulation time.

3. number of containers leased over the 14 days simulation time.

Those costs over the simulation period are plotted in Figures 4−5 for testing problems $P1$ and $P2$. For testing problems $P3$ and $P4$, they are presented in Figures $4-6$.

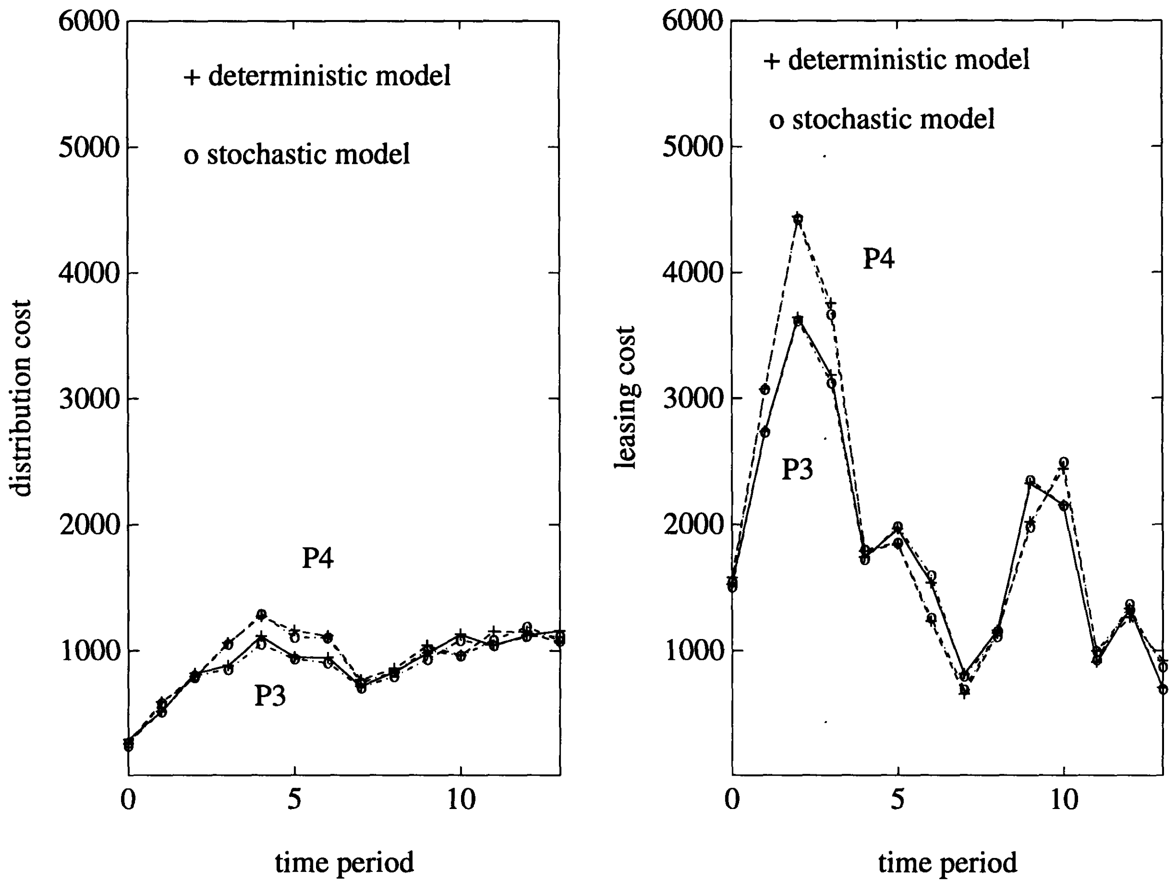


Figure 4-6: The distribution costs and leasing costs for testing problem P3 and P4

Table 4.3: The Daily Weights of Demands and Supplies

| Days | Sun | Mon | Tue | Wed | Thur | Fri | Sat |
|---|---|---|---|---|---|---|---|
| Daily Factors | 0.5 | 0.9 | 1.4 | 1.6 | 0.8 | 1.0 | 0.7 |

The figures on the left shows the changes of distribution costs over time, and those on the right give the shapes of leasing costs over time. From those graphs, we find the following results:

1. The weekly pattern is very clear from the changes of costs over two weeks simulation periods. In the experiments, day 0 is first Sunday, day 1 is the following Monday, etc. When generating testing problems, to reflect the weekly pattern of demands in practice, the following daily factors are used to adjust the demands generated.

2. The costs of deterministic model and costs of stochastic model on each day are very close. This means the deterministic model gives good approximation of the expected recourse function, for the testing problems generated by the problem generator described in section 4.1. But in real world, demands and supplies may not be so uniformly distributed over space and time in company's service networks. This indicates that simply using a generator that uniformly generates problem data may not reflect the real world problems. How to design a better problem generator could be a topic for future research.

3. In Figure $4-5$, $4-6$, the leasing costs shows more drastic weekly pattern than the distribution costs. This may due to costs of leasing a container is set to be much higher than costs of distributing a owned container in our problem generator. When leasing an empty container, the company has to pay a "pick up cost", rental fees and a "drop off cost" when the container is returned to the leasing company. In our model, this high leasing costs also ensure to use own containers to satisfy customer demands unless leasing costs less. This is also a strategy often adopted by shipping companies.

Table 4.4: Performance Measures of P1 Over a 14 Day Simulation Period

| Time | *Deterministic Model* | | | *Stochastic Model* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| 0 | 149.50 | 1064.90 | 22.60 | 152.50 | 1049.30 | 22.20 |
| 1 | 392.20 | 2370.70 | 48.00 | 382.20 | 2368.20 | 47.70 |
| 2 | 768.30 | 3191.90 | 65.20 | 706.60 | 3185.20 | 65.00 |
| 3 | 817.30 | 3819.40 | 78.30 | 802.80 | 3753.40 | 76.90 |
| 4 | 822.70 | 1384.40 | 31.60 | 802.20 | 1379.20 | 31.40 |
| 5 | 857.00 | 1703.70 | 36.10 | 848.20 | 1714.00 | 36.00 |
| 6 | 830.00 | 1355.50 | 26.20 | 820.40 | 1414.70 | 27.50 |
| 7 | 479.40 | 478.20 | 14.20 | 473.10 | 492.00 | 14.50 |
| 8 | 598.60 | 1124.20 | 31.70 | 614.10 | 1130.00 | 32.30 |
| 9 | 775.10 | 1939.70 | 55.60 | 724.20 | 1894.20 | 54.40 |
| 10 | 805.60 | 2240.00 | 62.50 | 830.10 | 2285.20 | 63.30 |
| 11 | 877.00 | 1155.90 | 33.20 | 849.90 | 1159.50 | 33.10 |
| 12 | 938.00 | 1458.20 | 40.50 | 932.80 | 1436.70 | 39.30 |
| 13 | 1106.60 | 484.60 | 15.80 | 1064.00 | 467.30 | 15.40 |

The results of the performance measures for the six testing problems are given in the following tables. In those tables, the columns under (1), (2) and (3) list the following data.

(1) total costs incurred for distributing empty containers during each time period,

(2) total leasing costs during each time period,

(3) total number of containers leased during each time period.

If historical real world data is available, we can using those data as inputs to alternative container allocation models. By comparing the performance measures obtained from each model with those in the historical data, i.e. the performance measures collected when the shipping company uses its current container allocation strategies, we can conclude whether the models perform better or not than the current strategies.

Table 4.5: Performance Measures of P2 Over a 14 Day Simulation Period

| Time | *Deterministic Model* | | | *Stochastic Model* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| 0 | 120.40 | 1995.00 | 42.70 | 124.40 | 1990.80 | 42.60 |
| 1 | 314.10 | 3321.80 | 73.90 | 289.30 | 3244.40 | 72.10 |
| 2 | 947.80 | 2576.70 | 56.70 | 922.00 | 2623.50 | 57.40 |
| 3 | 868.80 | 2268.10 | 45.60 | 918.00 | 2401.90 | 50.40 |
| 4 | 1161.70 | 1952.10 | 39.50 | 1236.20 | 2015.90 | 40.40 |
| 5 | 1372.00 | 1293.00 | 25.70 | 1561.80 | 1262.90 | 25.00 |
| 6 | 1886.80 | 778.80 | 16.40 | 1888.20 | 776.60 | 16.40 |
| 7 | 916.90 | 555.90 | 12.30 | 995.40 | 556.00 | 13.40 |
| 8 | 505.30 | 1491.70 | 45.80 | 508.60 | 1309.00 | 40.50 |
| 9 | 494.20 | 2777.80 | 80.30 | 571.00 | 2792.50 | 81.00 |
| 10 | 871.70 | 3289.70 | 87.80 | 918.70 | 3276.50 | 86.80 |
| 11 | 943.20 | 1693.10 | 45.70 | 953.50 | 1609.40 | 44.00 |
| 12 | 622.40 | 1419.80 | 43.80 | 698.00 | 1421.80 | 43.90 |
| 13 | 818.70 | 1121.30 | 27.00 | 698.70 | 918.80 | 22.20 |

Table 4.6: Performance Measures of P3 Over a 14 Day Simulation Period

| Time | *Deterministic Model* | | | *Stochastic Model* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| 0 | 292.20 | 1524.00 | 32.00 | 270.00 | 1505.90 | 31.50 |
| 1 | 514.80 | 2734.60 | 62.00 | 514.90 | 2730.10 | 61.60 |
| 2 | 818.00 | 3642.40 | 74.50 | 803.10 | 3619.00 | 73.80 |
| 3 | 881.20 | 3183.40 | 67.50 | 851.10 | 3123.90 | 66.20 |
| 4 | 1114.50 | 1738.20 | 35.80 | 1052.50 | 1722.50 | 35.20 |
| 5 | 949.70 | 1962.30 | 42.00 | 935.10 | 1983.60 | 42.20 |
| 6 | 944.80 | 1534.80 | 30.50 | 904.50 | 1594.80 | 31.80 |
| 7 | 715.20 | 810.70 | 22.00 | 706.30 | 795.10 | 21.40 |
| 8 | 830.00 | 1159.30 | 30.40 | 795.00 | 1113.30 | 29.40 |
| 9 | 980.80 | 2321.20 | 62.50 | 931.40 | 2349.90 | 62.70 |
| 10 | 1126.90 | 2150.50 | 63.90 | 1082.90 | 2146.60 | 63.80 |
| 11 | 1042.50 | 904.40 | 26.50 | 1040.40 | 925.20 | 27.00 |
| 12 | 1121.90 | 1330.30 | 39.30 | 1113.80 | 1367.50 | 40.00 |
| 13 | 1156.60 | 702.70 | 22.80 | 1117.20 | 694.50 | 22.70 |

Table 4.7: Performance Measures of P4 Over a 14 Day Simulation Period

| Time | *Deterministic Model* | | | *Stochastic Model* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| 0 | 257.80 | 1577.90 | 32.00 | 242.70 | 1543.20 | 31.40 |
| 1 | 592.20 | 3072.50 | 62.10 | 582.50 | 3069.60 | 62.00 |
| 2 | 793.90 | 4440.40 | 89.10 | 787.00 | 4426.30 | 88.70 |
| 3 | 1060.20 | 3753.70 | 80.00 | 1053.20 | 3667.30 | 78.00 |
| 4 | 1272.10 | 1786.30 | 37.90 | 1290.60 | 1799.50 | 37.90 |
| 5 | 1159.10 | 1843.20 | 38.60 | 1108.10 | 1853.50 | 39.00 |
| 6 | 1118.10 | 1230.10 | 24.20 | 1104.40 | 1258.90 | 24.50 |
| 7 | 767.00 | 653.40 | 20.30 | 746.80 | 689.40 | 20.90 |
| 8 | 856.70 | 1137.20 | 33.20 | 829.80 | 1154.20 | 34.40 |
| 9 | 1040.60 | 2017.60 | 58.20 | 1005.30 | 1978.50 | 57.90 |
| 10 | 961.00 | 2436.50 | 72.50 | 958.80 | 2492.60 | 73.70 |
| 11 | 1150.80 | 994.00 | 30.40 | 1084.30 | 994.60 | 30.10 |
| 12 | 1148.20 | 1259.80 | 35.20 | 1188.90 | 1321.60 | 36.80 |
| 13 | 1064.30 | 919.50 | 24.80 | 1079.10 | 871.00 | 23.50 |

Table 4.8: Performance Measures of P5 Over a 14 Day Simulation Period

| Time | *Deterministic Model* | | | *Stochastic Model* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| 0 | 257.20 | 2037.80 | 40.50 | 246.30 | 2004.00 | 39.80 |
| 1 | 519.60 | 2943.60 | 58.40 | 512.00 | 2927.30 | 58.00 |
| 2 | 815.60 | 4111.90 | 80.80 | 809.80 | 4121.30 | 80.70 |
| 3 | 1118.10 | 3802.80 | 80.30 | 1067.10 | 3806.10 | 80.20 |
| 4 | 1282.00 | 2081.20 | 40.90 | 1325.70 | 2131.90 | 41.80 |
| 5 | 1338.70 | 3558.30 | 70.50 | 1263.30 | 3430.30 | 67.80 |
| 6 | 1036.40 | 1608.00 | 32.50 | 1093.70 | 1623.30 | 32.80 |
| 7 | 1041.80 | 662.40 | 17.30 | 1019.00 | 656.50 | 17.50 |
| 8 | 686.70 | 1085.90 | 30.20 | 688.30 | 1013.20 | 28.40 |
| 9 | 921.70 | 1676.50 | 48.00 | 900.30 | 1816.60 | 51.50 |
| 10 | 921.60 | 2201.50 | 61.80 | 946.00 | 2232.00 | 62.10 |
| 11 | 1197.70 | 1092.30 | 30.40 | 1162.90 | 1104.10 | 30.20 |
| 12 | 994.70 | 1446.30 | 38.90 | 957.80 | 1459.40 | 39.00 |
| 13 | 860.40 | 642.70 | 18.50 | 912.20 | 664.90 | 18.40 |

Table 4.9: Performance Measures of P6 Over a 14 Day Simulation Period

| Time | *Deterministic Model* | | | *Stochastic Model* | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (1) | (2) | (3) |
| 0 | 251.90 | 1765.60 | 36.40 | 239.50 | 1760.50 | 36.40 |
| 1 | 524.80 | 3100.80 | 64.10 | 508.10 | 3073.20 | 63.30 |
| 2 | 621.60 | 4924.40 | 101.20 | 594.80 | 4911.80 | 101.20 |
| 3 | 1151.90 | 4398.50 | 93.50 | 1123.90 | 4323.10 | 91.90 |
| 4 | 1081.50 | 2316.20 | 50.70 | 1076.90 | 2333.50 | 51.10 |
| 5 | 1223.50 | 2441.30 | 51.20 | 1255.60 | 2444.40 | 50.70 |
| 6 | 1205.50 | 1441.80 | 30.70 | 1211.30 | 1443.80 | 30.80 |
| 7 | 982.70 | 775.80 | 21.80 | 938.70 | 743.40 | 20.70 |
| 8 | 1212.50 | 1587.10 | 47.20 | 1158.10 | 1488.70 | 44.70 |
| 9 | 959.00 | 2701.90 | 77.90 | 947.70 | 2762.20 | 79.50 |
| 10 | 1566.70 | 2095.30 | 59.20 | 1464.50 | 2153.80 | 60.40 |
| 11 | 1162.40 | 1008.40 | 27.80 | 1226.30 | 1012.70 | 27.50 |
| 12 | 1135.40 | 1559.20 | 45.80 | 1081.50 | 1517.10 | 44.10 |
| 13 | 1256.80 | 869.90 | 26.10 | 1279.60 | 882.70 | 26.30 |

# Chapter 5

# Conclusions and Recommendations

In this thesis, we study the container allocation problem which arises when a shipping company operates a fleet of containers to carry goods from shippers to receivers. The container allocation problem involves dispatching available empty containers to meet requests by shippers and redistribution of other empty containers to other depots or ports in anticipation of future demands. Our main objective is to develop optimization models for this problem which are numerically tractable and can be used as a real-time dispatching tool for the container fleet of a shipping company. We have developed both deterministic and stochastic models for container allocation problem. The performances of those models are evaluated through rolling horizon simulation on randomly generated testing problems and the results of the model evaluation is very encouraging.

We first present a new description of the container allocation problem in the context of a transportation network consisting of coastal ports and inland depots of the shipping company. We study the relationship of this problem with the dynamic vehicle allocation problem (DVA) which arises in other modes of transportation. We propose a modified definition for a more general DVA problem in which a leasing option is available to the shipping company when its own supply of empty containers fall short. We present the basic properties of the DVA problem, i.e. it has time and space dependency and its allocation decisions have to be made when we only have probabilistic knowledge of future demand and supply of vehicles. We present a literature review on the DVA problem and find that very little has been done for the container allocation problem arising in shipping industry. Specifically, only container allocation models for landside operation of a shipping company appear in the literature, and, to our knowledge, no numerical results of those models have been presented in the literature.

We use dynamic networks to model the container allocation process and develop dynamic deterministic and stochastic models on depot basis rather than on depot and customer basis which was used by Crainic and Dejax [1993]. Although the depot basis approach used in this thesis to model the allocation process is chosen according to the author's actual experience with an international shipping company, it offers a general modelling framework for this class of problems. The deterministic model with only empty container flow and one type of containers described in chapter two can be solved by using classical network optimization techniques. For the stochastic model of the same type, we formulate it as a stochastic program with full network recourse.

In order to solve the stochastic model for container allocation, we present possible solution methods to the stochastic program with network recourse. We find that it is numerically intractable to solve exactly this type of stochastic optimization problem for any real world size problem. We use the stochastic quasigradient methods, in particular the stochastic linearization method, and address the implementation issues of the stochastic linearization method for the dynamic stochastic model for container allocation. The major tasks of implementation are the proper choices of the step direction and step size in order to accelerate the convergence of this method. Successive averaging procedure is adopted in our experiments to choose the step direction. The inner product rule is chosen to change the stepsize during the iterative procedure of this method. Rapid improvement of the objective function values is observed in early iterations, which shows that the stepsize rule works well when the solutions is relatively away from optimal. In later iterations, the improvement becomes smaller as shown in Figure $4-1$, $4-2$ and $4-3$.

Then, we propose a rolling horizon simulation procedure which is specially designed for the evaluation of container allocation models. A random problem generator is developed to generate different testing problems for different problem parameters. Under the above framework for model evaluations, we obtain a posterior lower bound of total costs incurred for allocating containers over the simulation time which is generally unreachable since there is no uncertainty involved in the container allocation process. The total costs of using each model for container allocation are collected and compared with the posterior lower bound during the simulation and the results show that the stochastic model gives better solutions than the deterministic model in most of the cases. By comparing the results with a posterior lower bound of the total costs over the simulation period of our experiments, we also find that the stochastic linearization method produces a reasonable approximate solution to the stochastic model. The total costs for each testing problems are about 8% to 10% above the posterior lower bound. This is similar to the results obtained by Cheung and Powell [1992] for their DVA models for allocating truck motors to carry loads.

Our work in this thesis provides important insights on the container allocation problem. This work represents a building block for an extended research on this problem. Several immediate lines of investigation stimulated by this thesis are described below:

1. The models proposed in this thesis are very close to the real world operations of container allocation based our actual experience with an international shipping company and communications with the management of this company. However, further evaluations on these models need to be performed by using real world data before the models are applied to assist the decision making process for container allocation.

2. From our experimental results, we find that the total costs of the deterministic

model and stochastic model are very close over the simulation period for all the testing problems. The deterministic model is actually an approximation of the stochastic model in the sense that it uses only the expected values of the random elements instead of their probabilistic distributions. This approximation may be a good one for the testing problems generated in this thesis. Therefore, it is of interest to study the deterministic approximation of the stochastic model in other types of testing problems.

3. The planning horizon of the container allocation models is designed to capture the downstream effects of the allocation decisions made at current time period. Long planning horizons will obviously increase the size of the underlying dynamic networks of the models, whereas short planning horizons may not fully reflect the downstream effects. So, it is of interest to study the optimal planning horizon for the container allocation problem. Such research is particular important when we apply our model to real-world problems since long planning horizon can create a tremendous computational burden.

4. In our experiments, we only make 40 iterations when using the stochastic linearization method to solve the stochastic model. Further experiments are expected to study the relation of performance of the stochastic linearizaton method when we apply it to the stochastic container allocation models and the number of iterations under different stepsize and step direction rules.

5. The problem generator generates coefficients of networks uniformly over space and time dimensions of the underlying dynamic networks of the testing problems. The solutions of the deterministic model may be a good approximation of the stochastic model for this type of testing problems. Further study on the problem generator can be carried out in order to generate other cases of testing problems for which the deterministic model may give poor approximations.

6. In this thesis, we assume that the probability distributions of future demand and supply of empty containers are given. Developing forecasting models for

demands and supplies is a very important research topic since the quality of those forecasts affects the solution of the container allocation models presented in this thesis.

# Bibliography

[1] Andraottir, S. *A new algorithm for stochastic optimization* . In Proceedings of the 1990 winter simulation conference, p 364-366, 1990.

[2] Birge, J. R. *Decomposition and partitioning techniques for multistage stochastic linear programs.* Operations Research, 33(5):989-1007, 1985.

[3] Blum, J.R. *Approximation methods which converge with provability one.* Annals of Math. Stat.., 25:382-386, 1954.

[4] Cohen, G. *Auxiliary problem principle and decomposition of optimization problems.* J. Optim. Theory Appl., 32:277-305, 1980.

[5] Crainic, T., Gendreau, M. and Dejax, P. *Dynamic stochastic models for the allocation of empty containers.* Operations Research, 41(1):102-126, 1993.

[6] Culioli, J.C. and Cohen, G. *Decomposition/coordination algorithms in stochastic optimization.* SIAM Journal of Control and Optimization, 28:1372-1403, November 1990

[7] Dantzig , G.B. and Madansky, A. *On the solution of two-stage linear programs under uncertainty.* In Proc. 4th Berkerly Symp. on Math. Stat. and Prob. , p165-176, Berkerly, 1961. University of California Press.

[8] Dejax, P.J., and Crainic, T.G. *A review of empty flows and fleet management models in freight transportation.* Transportation Science, 21:p227-247, 1987.

[9] Ermoliev, Y. *On the stochastic quasi-gradient methods and stochastic quasi-Feyer sequence.* Kibernetika, 1, 1968.

[10] Ermoliev, Y. and Gupal, A.M. *An analog of the method of linearization in problems of minimizing nondifferentiable functions.* Kibernetika, No 1. p65-68, 1978.

[11] Ermoliev, Y. *Stochastic quasigradient methods an their application to system optimization* Stochastics, vol 9, p1-36, 1983.

[12] Gaivoronski, A. *Stochastic quasi-gradient methods and their implementation.* In Numerical Techniques for Stochastic Optimization, 313-351. Springer-Verlag, 1988.

[13] Gassmann, E.*Gradient Estimation via Perturbation Analysis.* Kluwer Academic Publishers, Norwell, Massachusetts, 1991.

[14] Gendreau, M., Crainic, T.G. Dejax, P., and Steffan, H. *Forecasting short-term demand for empty containers: a case study* Trans. Res. Record 1328, 98-103, 1991.

[15] Gupal, A.M. and Bazhenov, L.G. *A stochastic methods of linearization*, Cybernetics, p482-484, 1972.

[16] Higle, J.L. and Sen, S. *Stochastic Decomposition: an algorithm for two stage linear programs with recourse.* Math. of Oper. Res., 16(3):650-669, 1991.

[17] Kesten, H. *Accelerated stochastic approximation.* Annals of Math. Stat., 29:41-59, 1958.

[18] Kiefer, J. and Wolfowitz, J. *Stochastic estimation of the maximum of a regression function.* Annals of Math. Stat., 23:462-466, 1952.

[19] Misra, S.C. *Linear Programming of Empty Wagon Disposition.* Rail International, 3:151-158, 1972.

[20] Pflug, G.C. *On the penalization method in convex stochastic programming.* In Operations Research in Progress, p49-55, 1982.

[21] Powell, W.B. *A Comparative Review of Alternative Algorithms for the Dynamic Vehicle Allocation Problem.* In Vehicle Routing: Methods and Studies, Golden, B.L. and Assad, A.A. (eds). North Holland, Amsterdam, 249-291, 1988.

[22] Powell, W.B., Sheffi, Y. and Thiriez, S. *The Dynamic Vehicle Allocation Problem with Uncertain Demands.* in 9th International Symposium on Transportation and Traffic Theory, p357-374, J. vormuller and R. Hamerslag(eds), VNU Science Press, The Netherlands, 1984.

[23] Powell, W.B. *An Operational Planning Model for the Dynamic Vehicle Allocation Problem with Uncertain Demands.* Transportation Research, 21B:217-232, 1987.

[24] Powell, W.B. and Cheung, K.M. *Network recourse decomposition method for dynamic networks with random arc capacities.* Technical Report SOR-92-03, Dept. of Civil Engineering and Operations Research, Princeton University, 1992.

[25] Powell, W.B. and Cheung, K.M. *An algorithm for multistage dynamic networks with random arc capacities, with application to dynamic fleet management* Technical Report SOR-92-11, Dept. of Civil Engineering and Operations Research, Princeton University, 1992.

[26] Robbins, H. and Monro, S. *A stochastic approximation methods.* Annals of Math. Stat., 22:400-407, 1951.

[27] Ruszczynski, A. *A linearization method for nonsmooth stochastic programming problems.* In Stochastic Programming. Springer-Verlag, 1986.

[28] Van Slyke, R. M. and Wets, R. *L-shaped linear programs with applications to optimal control and stochastic programming.* SIAM J. Appl. Math., 17(4):638-663, 1969.

[29] Wallace, S.W. *Solving stochastic programs with network recourse.* Networks, 16:295-317, 1986.

[30] Wasan, M. *Stochastic Approximations.* In Cambridge Transactions in Math. and Math. Phys. 58. Cambridge University Press, Cambridge, 1969.

[31] White, W.W and Bomerault, A.M. *A network Algorithm for Empty Freight Car Allocation.* IBM System Journal, 8:147-169, 1969.

[32] White, W.W *Dynamic Transhipment Network Algorithm and Its Application toe the Distribution of Empty Containers.* Networks, 2:211-236, 1972.