Pre-bid Network Analysis for Transportation Procurement Auction under Stochastic Demand

by

Qian Wang

B.Eng. Industrial and Systems Engineering (2006)
National University of Singapore

Submitted to the School of Engineering
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Computation for Design and Optimization

at the
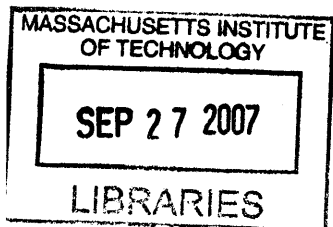
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2007

Author . . . .
School of Engineering
August 16, 2007

Certified by . . . .
David Simchi-Levi
Professor of Civil and Environmental Engineering
Thesis Supervisor

Accepted by . . . .
Jaime Peraire
Professor of Aeronautics and Astronautics
Codirector, Computation for Design and Optimization Program

Pre-bid Network Analysis for Transportation Procurement Auction under Stochastic Demand

by

Qian Wang

Abstract

Transportation procurement is one of the most critical sourcing decisions to be made in many companies. This thesis addresses a real-life industrial problem of creating package bids for a company's transportation procurement auction. The purpose of offering package bids is to increase the carriers' capacity and to improve the reliability of services. In this thesis, we investigate the possibility of forming packages using the company's own distribution network.

Effective distribution of packages requires balanced cycles. A balanced cycle is a cycle containing no more than 3 nodes with equal loads (or volume of package) on every link in the cycle. We develop mixed-integer programs to find the maximum amount of network volume that can be covered by well-balanced cycles. These models are deterministic models that provide a rough guide on the optimal way of package formation when loads are known in advance.

Since demand is random in real life, we perform a stochastic analysis of the problem using various techniques including simulation, probabilistic analysis and stochastic programming. Results from the stochastic analysis show that the effectiveness of package distribution depends on how we allocate the volumes on the lanes to create balanced cycles. If we always assign a fixed proportion of the lanes' volumes to the cycles, then it is only possible to have well-balanced cycles when the average volumes on the lanes are very large, validating the advantage of joint bids between several companies. However, if we assign a different proportion of the lanes' volumes to the cycles each time demand changes, then it is possible to create cycles that are balanced most of the time. An approximated solution method is provided to obtain a set of balanced cycles that can be bid out.

Thesis Supervisor: David Simchi-Levi
Title: Professor of Civil and Environmental Engineering

3

Acknowledgments

I am gratefully indebted to the following people during the course of this thesis research.

Professor David Simchi-Levi, my advisor, for giving me this opportunity to work with him. I am deeply thankful to him for all his guidance, patience and understanding. It has been an enriching experience working with him.

Professor Teo Chung Piaw, my co-advisor from the National University of Singapore, for his insightful suggestions and comments. He has shown great patience and spent a lot of time talking to me online. I learnt a lot from the discussions we had.

Jim, from the collaboration company, who shared a lot of his practical experience and insights with me.

Laura Koller and Jocelyn Sales, for their administrative support so that our student experience in MIT has been smooth and trouble-free.

The fellow students in my program, whose company during these 7 months have been really precious. Special thanks to my roommate, Rhea, and my office-mates, Lu Yun and Jiachuan.

My other friends in MIT and Singapore, who supported me all along, especially Linjuan, with whom I spent my most enjoyable time in Boston, and Ye Gang, whose encouragement helped me go through the upset days. I would also like to thank Song Miao and Fang Zhe for their help when I first started learning CPLEX.

My parents, for their unconditional love and faith in me.

And finally, the Singapore-MIT Alliance, for sponsoring my graduate study.

Contents

8

List of Figures

List of Tables

Chapter 1

Introduction

Transportation plays a key role in the economy, as producers rely on a variety of transportation modes to deliver products to customers efficiently. From the perspective of a global supply chain, transportation serves as a critical link between the different supply chain stages from raw material suppliers to the end customers; it also contributes to a significant portion of the total supply chain cost. Therefore, making the appropriate transportation procurement decision is very important to a company's success, in terms of keeping costs down and improving supply chain efficiency. The 2006 Standard and Poor's Industry Survey [1] on commercial transportation indicates that commercial freight transportation in the United States accounted for $720B of revenues in 2004. Out of the $720B, $671B of the market goes to trucking operations; in other words, the trucking industry occupies 87.1% of the entire domestic freight market in the U.S.

1.1 Overview of the U.S. Trucking Industry

In freight transportation, shippers, such as manufacturers and distributors, are the beneficial owners of freight. Carriers are transportation companies and service providers, such as trucklines, airlines and ocean transport providers [2]. A load is a collection of items with a common origin and destination which also move in the same vehicle for some portion of

their total trip [3]. A lane is a one-way origin destination pair on which loads need to be moved [4].

Trucking is an important mode of freight transportation. In the year 2003, the trucking industry hauled 68.9% of the total tonnage of freight in the U.S., equating to 9.1 billion tons. Since the deregulation of the motor carrier industry by the Motor Carrier Act of 1980, the number of interstate motor carriers has increased drastically. By July 2004, there were more than 524,000 U.S. carriers on file with the U.S. Department of Transportation. Among all the trucking companies, 96% of them operate 20 trucks or less and more than 87% operate six trucks or less [5]. Therefore, the trucking industry is mainly comprised of small businesses, which makes this industry highly fragmented and hence competitive.

1.1.1 Different Types of Fleet

Carriers can be classified into for-hire carriers, private fleets, dedicated fleets and fleets that haul U.S. Mail, according to how they are operated and who the owner is.

Private Fleets

Private fleets are operated by companies whose primary business is not hauling freight for-hire, but who own or lease a fleet of trucks in support of their primary business. Usually, companies use private fleets if they need to make frequent and timely deliveries to the customers. For example, the wholesale food service industry favors private fleets very much. As a large sum of fixed capital investment needs to be made to own a private fleet, it is only cost effective if the required level of customer service and demand are both high.

Dedicated Fleets

Dedicated fleets are not owned by the shipper; however, they provide exclusive service to the shipper just like the private fleets, on a contractual basis. A lot of the large carriers such as Schneider National have dedicated fleet business. A private fleet that is not large enough

may be replaced by a dedicated fleet, as the dedicated fleet does not require a large amount of initial capital outlay.

For-hire Carriers

For-hire carriers are motor carriers that offer freight transportation services to shippers at a pre-specified rate on a contractual basis. For-hire carriers can be further divided into different segments, namely, truckload, less-than-truckload, tank, refrigerated, etc. Truckload (TL) is defined as the quantity of freight required to fill a truck. When used in connection with freight rates, TL also refers to the quantity of freight necessary to qualify a shipment for a truckload rate. Usually, TL is greater than 10,000 pounds. Less-than-truckload(LTL), on the other hand, refers to a quantity of freight less than that required to apply a truckload rate, which is usually less than 10,000 pounds [5].

Consolidation is the most important characteristic of the less-than-truckload operation. A LTL carrier usually travels on a regular route. Along the route, it collects freight from different shippers and consolidate the freight onto a line-haul truck which moves the freight to a delivering terminal or to a hub terminal where the freight will be further sorted and consolidated for additional line-hauls. A line-haul is defined as a vehicle trip between two different locations with no intermediate stops [3].

As opposed to LTL carriers, truckload(TL) carriers perform direct line-hauls from origin to destination. The advantages of TL carriers over LTL carriers are shorter transit times and greater reliability due to less amount of handling. However, using TL carriers inevitably incurs a much higher cost than using LTL carriers. There were around 45,000 for-hire TL carriers in the U.S. in 2004. More than 60% of them had annual revenues of less than $1M [1]. Again, it can be observed that the market for for-hire TL carriers is highly segmented, mainly due to low entry barriers. In later parts of this thesis, the term 'carrier' is used to refer to for-hire TL carrier in general.

17

1.1.2 Carrier Economics

It is necessary to understand carrier's economics and how shippers can influence the carrier's cost structure. Economies of scale usually refer to the decrease in production cost as volume increases; while economies of scope are related to the cost reduction as different products (usually complementary) are produced. Among most commodities, economies of scale are exhibited - price goes down as the buyer purchases a larger volume of goods from the seller. However, transportation is special. Jara Diaz [6] [7] defines transportation as a multi-dimensional output process. It has 4 dimensions: origin, destination, freight and timing.

In transportation, economies of scale refer to the decrease in average cost per load given that the load volume on a given lane increases. By having economies of scale, we assume lane independence with respect to the carriers' cost. On the other hand, economies of scope looks at how the freight on one lane is related to the freight on the other lanes, i.e. lane interdependence. For example, assume that the carrier currently ships 10 loads per week from A to B, if the shipper increases the weekly volume to 20 from A to B in Scenario 1 (giving economies of scale), while the shipper gives another 10 weekly loads from B to A in Scenario 2 (giving economies of scope), usually the carrier will be more willing to give a better rate in Scenario 2 because Scenario 2 helps the carrier to balance its network.

It was also shown in literature that economies of scale are absent in the trucking industry [8]; whereas economies of scope have a significant influence [6] [7]. We define follow-on load as the load which originates from the destination location of the current lane on which the carrier is serving. Carriers exhibit strong economies of scope because with a high probability of finding a follow-on load, empty miles can be reduced, equipment and drivers can also be better managed - which all contribute to a reduced cost of service. This is also illustrated by the example in the last paragraph.

1.1.3 Truckload Procurement Process

Shippers buy transportation services from carriers using a request-for-proposal(RFP) process, which is also called an auction. Caplice and Sheffi [9] described the transportation procurement process in three steps: *Bid Preparation* where the shipper selects the carrier base and determines the lanes to be bid out. *Bid Execution* which involve an exchange of information between the shipper and the carriers. The shipper provides the invited carriers with details of the network being bid out; whereas the carriers respond with quotes. *Bid Analysis and Assignment* where the shipper analyzes the carriers' proposals and assigns them to the network respectively.

In the past, the RFP process is usually done in the following way. The shipper first estimate the freight that need to be hauled in the coming year from historical data. Then, it provides the carriers with lane information such as origin, destination, estimated volume per week, and days of shipments. After the carriers submit quotes on the prices at which they are willing to ship the loads, the shipper evaluate the bids lane by lane to determine the winner, usually based on a single criterion - price. They may also negotiate for bundles of lanes with one carrier at a time. This process can be considered as a repeated set of simple sealed-bid auctions, ignoring lane interdependencies.

However, as explained in section 1.1.2, the carriers exhibit strong economies of scope. Therefore, ignoring lane interdependencies in the transportation procurement auction is not cost effective. In recent years, several large shippers have implemented combinatorial auction mechanisms. In a combinatorial auction, all lanes are made available for bidding simultaneously. The carriers are allowed to quote prices on packages of lanes, in addition to individual lanes. The carriers can thus form their own packages based on their existing service network, their drivers' hometowns, and their maintenance networks. Both the carrier and the shipper can benefit if these packages contribute to cost cutting or better utilized capacity. Moreover, the carriers are also allowed to submit bids on partial packages, i.e. a specified percentage of the volume on a package.

In combinatorial auctions, the optimization problem to identify the best set of winning

bidders is named as the *Winner Determination Problem*. The winner determination problem in transportation procurement auction is modeled as a set-covering problem (see Caplice and Sheffi [9] for the formulations), which is NP-hard.

There are a few software vendors that provide optimization based bidding tools to help shippers make strategic sourcing decisions, such as CombineNet, Emptoris, OptiBid, etc. With these tools, the shipper can specify its business rules in a expressive and precise manner. For example, the shipper can state that "Limit the total number of winners to below 200", "Ensure that Carrier X wins at least 30 loads per week", "Any one carrier cannot win more than 15 percent of the total bid volume", etc. On the other hand, carriers can also express their offers and conditions in a convenient manner, such as capacity constraints, discount schedules and various conditional discounts. These business rules or side constraints are translated into optimization constraints by the software solutions. The shipper also has the flexibility to navigate multiple scenarios to find the optimal allocation decision under each scenario, where each scenario is solved as a large-scale mixed-integer program by the software.

1.1.4   Issues and Challenges

The American Transportation Research Institute conducts annual studies to investigate the critical issues in the U.S. trucking industry. In the 2006 report [10], it listed the top three critical issues as driver shortage, fuel costs, and driver retention.

Driver Shortage

The capacity of the trucking industry is facing a severe shortage problem as the economy grows. In another study conducted by Global Insight [11], it was estimated that there is currently a shortage of 20,000 long-haul drivers in the U.S.; it was also projected that the driver shortage will exceed 110,000 by 2014.

Except for low wages, there are a few other reasons that made the job of being a truck driver unattractive, the most important ones being "quality of life" issues [12]. These is-

sues include not being able to go home regularly and having to follow unpredictable route schedules, especially for long-haul truck drivers.

The issue of driver shortage is already forcing trucking companies to turn down orders and raise driver compensation to historic levels. It may also be worthwhile for the carriers to redesign their shipping networks or to adjust their operating practices so that better work schedules can be made available to the drivers.

Fuel Costs

In the past few years, fuel prices have shown a general rising trend. The average diesel prices in the U.S. have increased by 91% between January 2003 and April 2006 [12]. Given the volatility of fuel prices due to political instability and OPEC production cuts, this issue will continue to be one of the major concerns for carriers.

Driver Retention

The cost of hiring and training a driver is estimated to be $4000 to $8000 [10]. Due to some of the unattractive aspects of this job as explained above, most carriers are very concerned about high turn-over rate of drivers. Various measures such as improved trainings and driver compensation schemes are implemented to raise the level of driver retention.

1.2   Thesis Objective and Motivation

This thesis analyzes a real-life industrial problem faced by company ABC [1], who relies heavily on truckload transportation in its supply chain. The objective is to develop a new fleet modeling tool, which can be used in bid preparation. This new tool creates small packages on the shipper's behalf and input information of the pre-bundled lanes into the bidding system. By doing so, we give carriers the extra option to bid on these pre-bundled lanes, in addition to the current system of only offering individual lanes to them and relying

---

[1]Company ABC is a large US-based manufacturer. We changed the name of the company to protect its identity.

entirely on their own ability to package the lanes in the bidding process. We also investigate the effectiveness of these pre-bundled lanes under demand uncertainty.

1.2.1 Motivation

Less Dependence on Carrier's Ability to Package Good Lanes

As described in the beginning of section 1.1, the trucking industry is highly fragmented. It is sometimes difficult for many carriers to construct well-designed packages which are cost efficient, and yet do not compromise on the level of compliance to meet demands. Especially for carriers who has less experience with the shipper, they often tend to bid low on bundles of lanes, not fully aware of the shipper's demand variability and hence the balance of workload. After being awarded the lanes, they face with the problem of meeting the required level of compliance and often have to turn back loads, which is undesirable for both the shipper and the carriers.

However, it is believed that although some carriers, as mentioned above, have a high turndown ratio, they do not necessarily perform worse than the largest, most sophisticated carriers when given well-designed routes, especially for short-hauls. By offering these pre-bundled lanes, we encourage all carriers to compete on a more even platform, regardless of their own ability to create packages.

Imperfect Collaborative Bidding with Other Shippers

Sometimes, company ABC also collaborates with other shippers (e.g. its suppliers) to conduct a joint bid in which the carriers are invited to bid on both the company and the other shippers' lanes. It is expected that this joint bid will allow the carriers to enjoy the benefit of having back-hauls or multi-legged routes. However, this collaborative bidding system does not always work perfectly as the volumes promised by the other shippers may not be realized all the time.

Therefore, the new fleet modeling tool is trying to explore opportunities to package lanes based on the company's own demand and then offer the carriers with this extra flexibility of

bidding on the pre-packaged lanes.

Direct Comparison on the Pricing

Another motivation of pre-bundling the lanes is that if the carriers can now bid on the same packages, it gives the shipper a more direct comparison on the prices that are offered by different carriers.

1.3 Literature Review

There is an abundant amount of literature relating to the transportation procurement problem. Both quantitative and qualitative approaches have been adopted by researchers to understand and improve the efficiency of transportation procurement activities. This section reviews three main areas of the relevant literature.

1.3.1 Identifying the Best Carrier Selection Criteria

There is a great deal of qualitative studies focused on identifying the best carrier-shipper relationships and carrier selection criteria. This belongs to the Bid Preparation step [9] as explained in section 1.1.3.

Bardi, Bagchi and Raghunathan [13] developed a questionnaire which found the most important carrier selection determinants, the top five being transit-time reliability, transportation rates, total transit time, willingness of the carrier to negotiate rates, and financial stability of the carrier. They also used factor analysis to determine the impact of the 1980's deregulation on the emphasis placed on carrier selection determinants - rate-related factor and customer service factor experienced the greatest changes in emphasis.

Abshire and Premeaux [14] studied how different shippers and carriers perceived the importance of various carrier selection criteria. It was found that four out of thirty-five selection criteria that were important to shippers were underrated by carriers, while another four that were less important to shippers were overrated by carriers. Different opinions

between shippers and carriers resulted in reduced shipper satisfaction and subsequent carrier losses.

In the research program conducted by Gibson, Sink and Mundy [15], it was found that transportation procurement strategies and shipper-carrier relationships have changed from price-based and short-term to cooperation-based and long-term. Greater emphasis on strategic partnerships between shippers and carriers provides opportunities for future cost reduction and service improvement.

1.3.2 Applying Combinatorial Auctions in Transportation Procurement

The application of combinatorial optimization in transportation procurement problem has attracted much industrial and academic attention, due to the increasing popularity of using combinatorial auctions in the sourcing process. This is involved in the Bid Analysis and Assignment step [9] in the bidding process (section 1.1.3).

In the late 1980s, the Reynolds Metals Company developed a centralized dispatching system. In Moore, Warmke and Gorban's paper in 1991 [16] which explained this system, they described an optimization-based approach to assign carriers to different locations. It was one of the earliest applications of combinatorial auctions.

Caplice and Sheffi [9] presented a detailed analysis in using optimization-based techniques for transportation procurement, by including two carrier assignment optimization models. The authors emphasized the importance of carrier's economics in lowering the cost to carriers, and hence the advantage of an optimization-based approach which took carrier's economies of scope into consideration.

Song and Regan [4] studied the benefits of combinatorial auctions from the carrier's perspective. Simulation experiments in their study showed that carrier also had cost reduction under combinatorial auctions; and the carrier's cost reduction was closely related to the distribution density of new lanes, in addition to its current lanes.

In addition, Sheffi [2] summarized the development, applications and benefits of combi-

natorial optimization in transportation procurement.

1.3.3 Incorporating Demand Variability into Procurement Decisions

Most of the optimization based bidding software find optimal solutions based on a deterministic approach using average demand. However, data uncertainty is present in real life. Allocation of shipments based on the average results in carriers having either excess capacity or insufficient capacity at times. The following literature, although focused on different transportation procurement issues, accounted for variability in their methodology or analysis.

Harding [17] designed a robust transportation planning methodology to minimize the shipper's cost due to unplanned events. A variety of transactional data were analyzed and integrated with the optimization software. The author also developed a simulation model to test the robustness of an optimized transportation plan.

Mulqueen [12] also considered the issue of variability in addressing the problem of creating transportation policy. He adopted an iterative process that used both optimization and simulation to ensure that variability within the distribution network was taken into account. In his study, it was also found that variability adjusted volume should be used for each lane in the planning so as to achieve an optimal confidence level.

In Chapter 5 of Caplice [3], the probability of matching an inbound load to an outbound load was assessed using a simple expression, derived based on properties of the poisson process.

Finally, Powell *et al.* [18] also incorporated probabilistic analysis in their design of a daily load dispatching methodology.

1.4 Outline of Thesis

The remainder of this thesis is organized as follows. Chapter 2 gives a more detailed description of the problem and analyzes data from the given distribution network of company ABC.

In Chapter 3, we present four optimization models to solve for the lane bundling problem with a deterministic assumption; the computational results from the various models are also shown and compared. Chapter 4 employs several techniques to analyze how the optimal set of packages should be created under demand uncertainty; these techniques include simulation, probabilistic analysis, and stochastic programming. Finally, the findings of this thesis are summarized and future work directions are outlined in Chapter 5.

Chapter 2

Problem Description and Data Overview

2.1 Proposed Package Models

The objective of this thesis is to develop feasible and reliable pre-bundled packages for the carriers to bid on. Caplice [3] discussed four forms of lane bundling: *Reloads* which consist of two lanes, where the destination of one lane is the origin of the other; *Open Tours* which are collections of more than two lanes that begin and end at different locations; *Closed Tours* which are collections of lanes that begin and end at the same location; and *Local Tours* which are groups of short haul lanes with a common origin and/or destination. Based on past experience with carrier-constructed package bids, and also due to the fact that the carriers' shipping networks are unknown to the shipper, we propose the following two package models which are more likely to be successful.

2.1.1 Model 1: Two-way and Triangular Cycles

In the first model, we consider two-way and three-way cycles (i.e. closed tours) between nodes. Each node here represents a market that is defined as the aggregation of a few nearby locations.
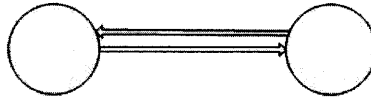
Figure 2-1: Two-way cycle
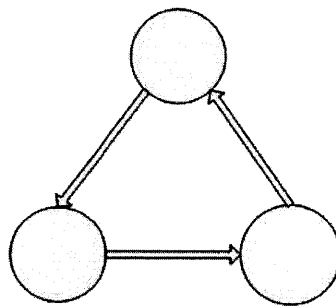
Figure 2-2: Three-way cycle
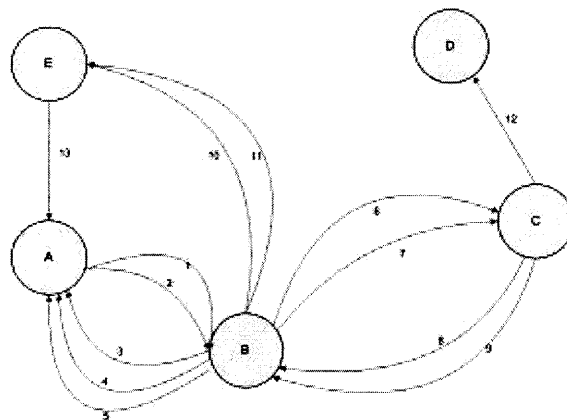
Figure 2-3: Illustration of a small network

Figure 2-1 and Figure 2-2 illustrate a two-way cycle and a three-way cycle respectively. In Figure 2-3, a small network with 5 markets is shown. There are one or more lanes connecting each pair of markets. Two-way cycles may be formed by lanes between markets A and B, or markets B and C; while three-way cycles may be formed by the lanes connecting markets A, B and E.

We define a balanced cycle as a cycle with no more than 3 nodes having equal loads on every lane in the cycle. We only consider cycles with two or three nodes, because for cycles containing more than three nodes, it is usually difficult to ensure balance of flows in the cycles and hence carriers' performance on those multi-legged cycles is not guaranteed to be more stable than on ungrouped individual lanes. In constructing the packages, we do not necessarily need to assign the full demand of a lane to the cycle - the demand on the cycle can come from some fraction of a lane's total demand.

The objective of this model is to maximize the volumes shipped on all balanced cycles in the network. This requirement of balanced cycles contributes to balancing the carriers' networks. This is important to the carriers, as they can always have follow-on loads (defined in section 1.1.2) in balanced networks, allowing them to make better utilization of their drivers and equipment.

2.1.2 Model 2: Regional Dedicated Fleet
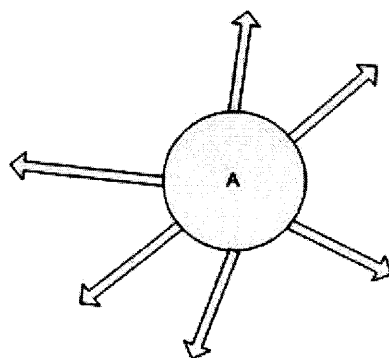
Figure 2-4: Regional dedicated fleet

In this model, we construct a fleet which serves an area around a market or a plant where

there are a lot of ship-outs. This model is similar to the idea of local tours, as mentioned in the first paragraph of this chapter. It is appropriate because company ABC's transportation network usually has more outflows than inflows. The radius of the area served by the regional dedicated fleet should not exceed a given number of miles, for example, each lane should have a transit time of less than 2 days, as we prefer short-hauls in this model.

As this model involves a dedicated fleet, we introduce another parameter which is the number of drivers in the fleet. For example, if we need to have a 10-drivers fleet serving the region with center A, we shall pick the subset of lanes out of A such that the number of loads which can be served by the 10 drivers is maximized. In other words, the objective is to maximize utilization of the 10 drivers, while maintaining the level of compliance above a certain standard, such as 80%.

This thesis, however, will only focus on developing the packages described in Model 1, which is to create balanced cycles.

2.2 Advantages of the Proposed Package Models

2.2.1 Increased Capacity

These small packages with regular routes increase carriers' capacity, as these packages offer relatively fixed schedules and allow drivers to go back home regularly. As we reported in section 1.1.4, the trucking industry is severely challenged by the problem of recruiting and retaining drivers. Allowing drivers to go home regularly is one of the most crucial factors that count towards driver satisfaction. Therefore, with these packages, it makes hiring extra drivers a lot easier for the carriers.

2.2.2 Improved Reliability

Pre-bundling also improves the reliability of the shipper's transportation network. As these packages are constructed by shipper, it is less likely to have scheduling conflicts or imbalanced workloads on the lanes in these pre-bundled packages.

Another factor which also contributes to reliability is that we usually have almost the same group of drivers traveling the same set of lanes for the small packages. Familiarity with the lanes not only adds to driver satisfaction, but also improves on the reliability of delivery.

2.3 Data Overview

In this section, we will give an overview of the data we received from company ABC to which this research applies. We will also be able to get an insight to the company's distribution network from the data.

The data comes from a bidding exercise conducted two years ago. The network consists of 5335 lanes. Each lane is defined by an origin point and a destination point, based on the 5-digit postal code. There are 119 distinct origin locations and 1626 distinct destination locations. As mentioned in section 2.1.1, we aggregate a few locations to form a market. In the given data, each location is already assigned to a market; and it is observed that the general rule of forming markets is to aggregate locations sharing the first 2 digits of zip code. After aggregating the locations into markets, we can also combine lanes starting from and ending at the same markets. As a result, we have 1718 aggregated lanes, originating from 57 distinct markets and terminating at 157 distinct markets. The projected average weekly volume over the entire network is 13,826 loads, transported over 8.9 million miles per week. So each load travels about 640 miles on average.

2.3.1 General Orientation of the Network

We have observed a much smaller number of origin locations than destination locations, i.e. starting from a few origin points, the loads are distributed into a large number of destination points. This implies that the distribution network of company ABC is primarily for outbound loads, which is also claimed by the company's transportation personnel.

Figure 2-5 describes each shipping location by the number of its outbound and inbound loads. The dotted line is a 45 degree line on which the number of outbound and inbound loads

31
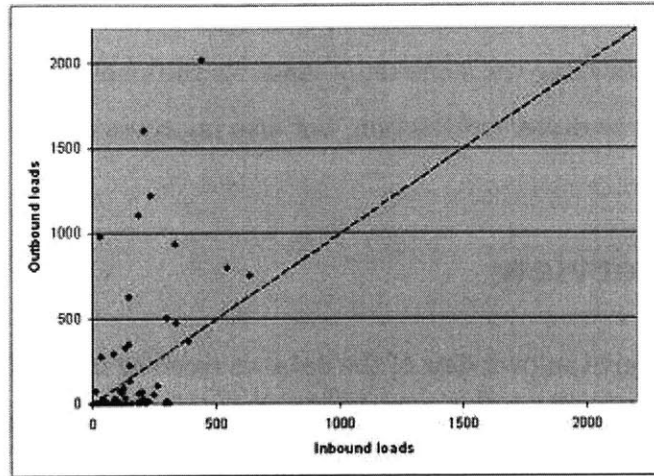
Figure 2-5: Classification of company ABC's shipping locations by the number of outbound and inbound loads

are equal. Shipping points located above the line handle mainly outbound loads, and those below the line have more inbound loads. From the figure, we can observe a large number of small-volume, inbound points and several large-volume, outbound points, again validating the above notion that company ABC has an outbound oriented distribution network.

2.3.2 Analyzing Load Volume by Lane
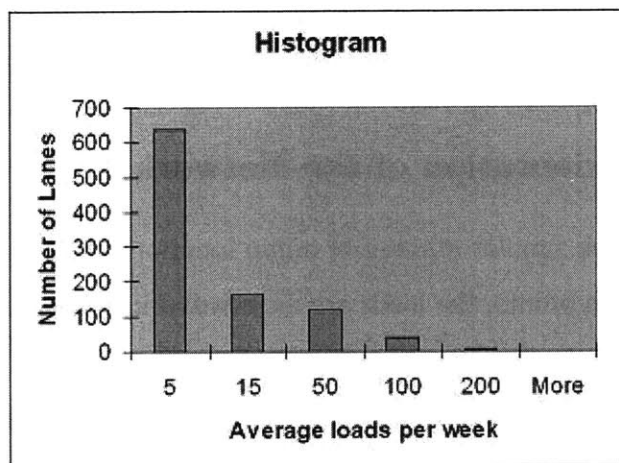
Figure 2-6: Frequency of lane volumes

The average load volume per lane is 2.59 per week. The histogram shown in Figure 2-6 reflects the variation of shipment volumes on different lanes. We can see that more than 600 out of the total 1718 aggregated lanes have less than 5 loads per week, and less than 10 lanes have more than 100 loads per week.
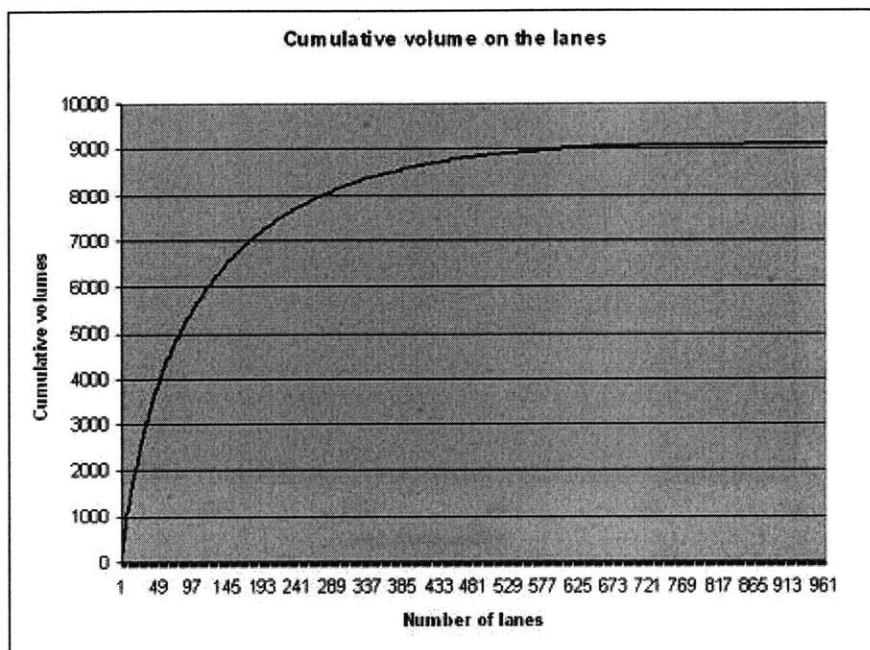
Figure 2-7: Cumulative volume on the lanes

It was also found that a large proportion of the loads transported comes from relatively few lanes. 50% of the total volume is transported by only 7% of the lanes, and 80% of the total volume is transported by 20% of the lanes (shown in Figure 2-7).

Therefore, there are a large number of lanes in the company's network that have very few loads and cumulatively account for only a small proportion of the total volume.

2.3.3 Investigating Demand Variability

Demand Variation with respect to Lane Volume

Besides average weekly volume, we are also given standard deviation of the weekly volume on each lane. We use the coefficient of variation $\frac{\sigma}{\mu}$ to assess demand variability on the lanes.

It was found that different lanes have different amount of variability, with coefficients of variation ranging from 0.02 to 16.
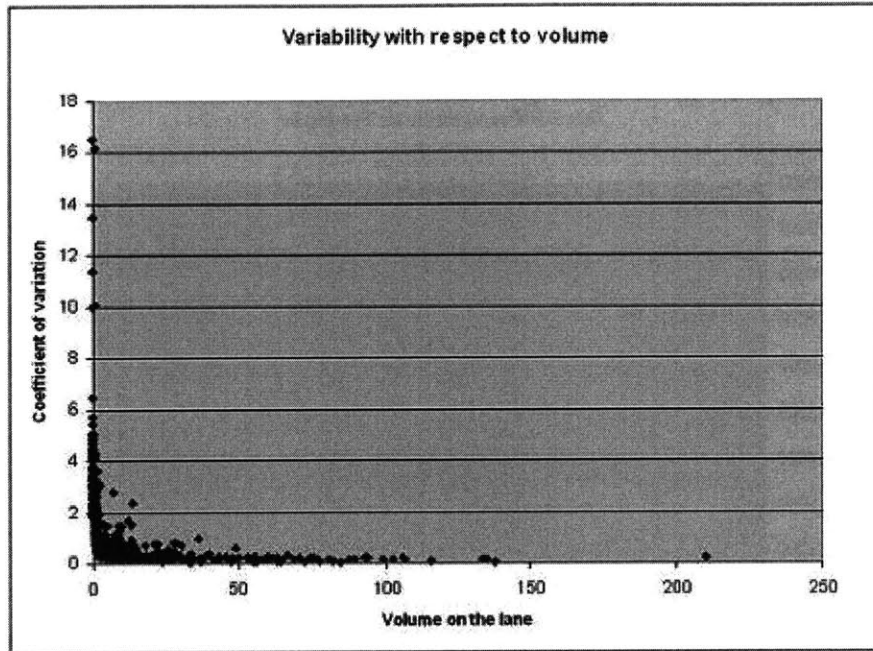
Figure 2-8: Variability with respect to volume

Figure 2-8 shows the relationship between the coefficient of variation and the average weekly volume on each lane. We can see that lanes with higher average volumes tend to have smaller variability. The highest coefficients of variation usually occur on lanes with a weekly demand of less than 10.

Demand Variation with respect to Transit Time

We are given the transit time on each lane in number of days. In Figure 2-9, we plot the coefficient of variation against the transit time on each lane. Although there are a few outliers for lanes with a transit time of less than 3 days, in general, we observed that the coefficient of variation may not be related to the transit time on the lane, which is also indicative of the relationship between load variability and lane distance.

34
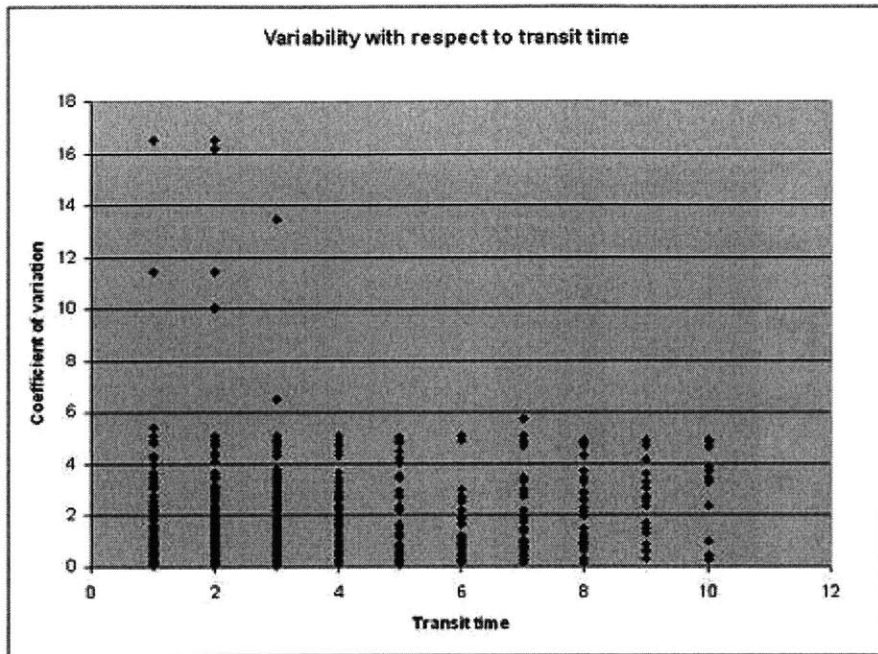
Figure 2-9: Variability with respect to transit time

Demand Variation with respect to Day of Week

Figure 2-10 plots the average number of loads shipped on each day of week, aggregated over the entire network. It is quite obvious that the volume on weekends (Sundays and Saturdays) is less than the volume on weekdays.

In addition, the lanes not only differ from one another in average volume and load variability, they also differ in the day of week shipment pattern. While Figure 2-10 shows the shipments on each day of week by aggregating the volumes on all lanes; Figure 2-11 and Figure 2-12 are examples of two individual lanes, where lane X does not have a distinct pattern for weekend loads but lane Y has very few weekend loads.
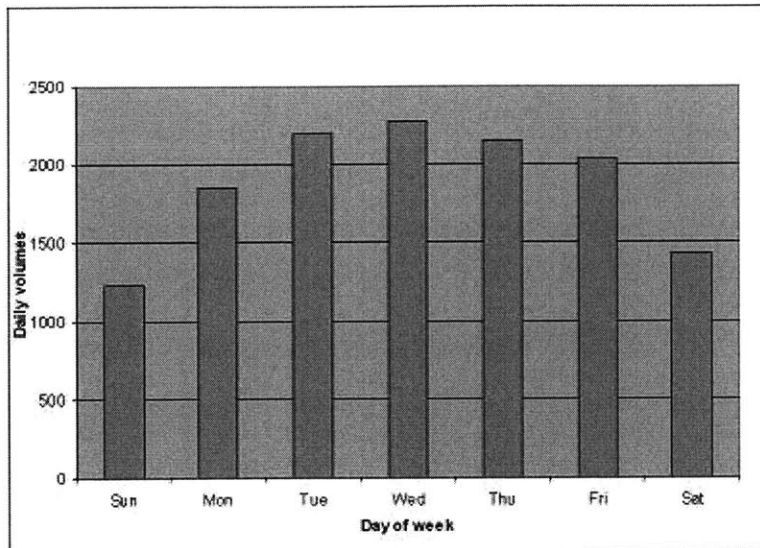
Figure 2-10: Aggregated volume on each day of week
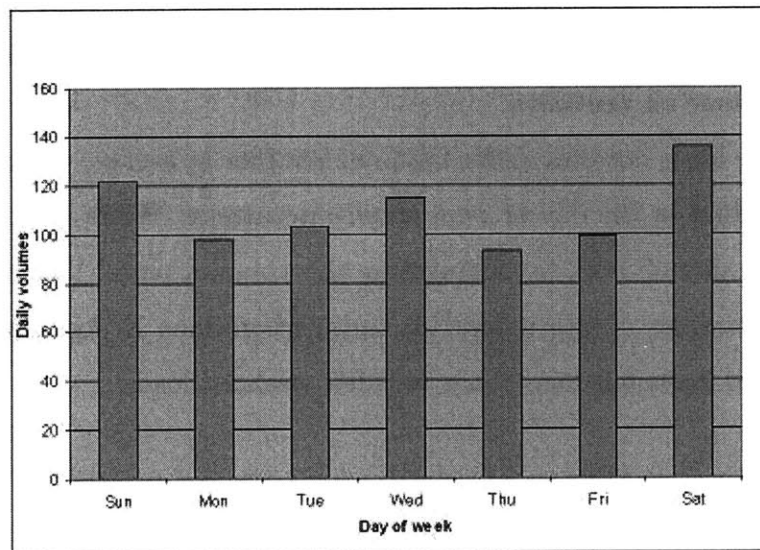
Figure 2-11: Volume on lane X on each day of week
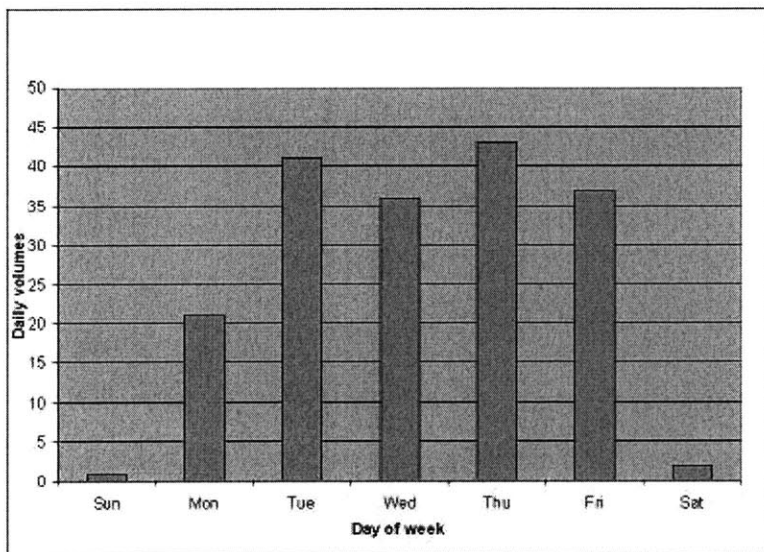
Figure 2-12: Volume on lane Y on each day of week

Chapter 3

Deterministic Models

In this chapter, we will propose optimization models to solve for the lane bundling problem deterministically. The objective of these models, as stated in section 2.1.1, is to create two-way or three-way balanced cycles out of the company's existing distribution network, and to maximize the volumes shipped on all such cycles. We will also present and analyze the deterministic solutions.

3.1   General Assumptions

3.1.1   Full Truckload Movements

The unit of shipment in the data given by company ABC is truckload. Therefore, we assume that all loads in our model have full truckload movements.

3.1.2   Empty Miles

In our model, we aggregate nearby shipping locations into markets. Our two-way and three-way cycles are formed based on markets. In actual operations, we will have empty miles as the follow-on load may be in another location within the same market. However, the effect of empty miles are ignored here, as the focus of our study is on balancing flows on the cycles, rather than minimizing the cost incurred due to empty miles.

3.1.3 Transit Times

We are given the transit time on each lane in number of days, without knowing the exact time for each pick-up and delivery. It is not necessary to know the pick-up and delivery times precisely, as we only want to model the flow balances in terms of loads, instead of the movement of individual drivers.

3.1.4 Vehicle Compatibility

It is assumed in our model that the vehicles used to carry each load are all compatible. Hence, we have less restriction in forming the cycles, as the vehicle used to ship one load will always be compatible with the equipment required to ship the follow-on load on the subsequent link of the same cycle.

3.2 Compliance Requirement

The most important constraint of this problem is to have balanced flows on the cycles. We use the *compliance requirement* to assess whether sufficient balance has been achieved on the cycles. Compliance, defined by company ABC, is the percentage of demand that the carrier is able to satisfy using the drivers who just came from the previous node in the cycle. The compliance requirement is set at 92% by the company. For example, given a two-way cycle with nodes A and B, to satisfy the compliance requirement, we want to have 92% of the loads from node B to node A shipped by the drivers who just delivered loads from A to B; and vice versa.

Mathematical Interpretation of the Compliance Requirement

In the two-way cycle in Figure 3-1, let x denote the number of loads from A to B, let y denote the number of loads from B to A.
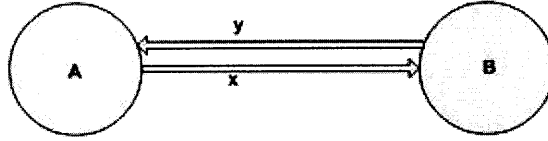
Figure 3-1: Two-way cycle

By having 92% compliance, we have:

$$0.92y \leq x$$

(The loads from A to B cover at least 92% of the loads from B to A, i.e. we have enough drivers for at least 92% of the loads from B to A.)

$$0.92x \leq y$$

(The same compliance condition applies in the reverse direction.)

Therefore, we have:

$$x \leq \frac{1}{0.92}y = 1.087y$$

$$y \leq \frac{1}{0.92}x = 1.087x$$

This is the same as having volumes on both lanes to differ by less than 8.7%; we call this "almost-balanced". In other words, the amount of loads from A to B should be within a certain range ($0.92y \leq x \leq \frac{1}{0.92}y$) of the amount of loads from B to A; and vice versa.

In the three-way cycle in Figure 3-2, denote the number of loads on the three lanes by x, y and z respectively.
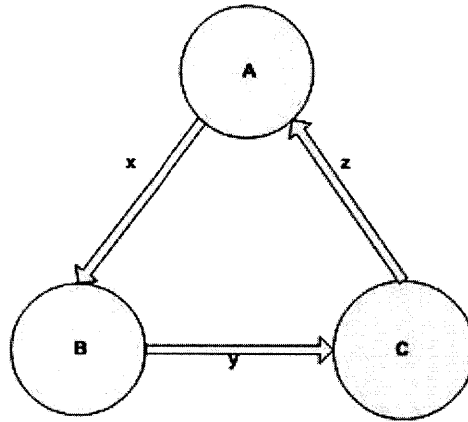
Figure 3-2: Three-way cycle

By having 92% compliance, we have:

$$0.92x \leq z$$

$$0.92y \leq x$$

$$0.92z \leq y$$

$$\Rightarrow \begin{cases} 0.92x \leq z \Rightarrow x \leq 1.087z \\ (0.92)^2x \leq 0.92z \leq y \Rightarrow x \leq 1.181y \end{cases}$$

Similarly for y and z.

Therefore, if the loads on each lane do not differ from one another by more than 8.7%, then the 92% compliance requirement is satisfied too. In other words, we also want the volumes on lanes of the same cycle to be "almost-balanced".

3.3 Finding Balanced Cycles

It is not meaningful to construct cycles with very little demand on each lane, because they give little or no marginal benefit. Thus, we require the average weekly demand on each link of a cycle to exceed 2 loads. This parameter can be changed if necessary.

First, we aggregate shipping locations in the given network into markets (the second paragraph in section 2.3 described the general rule of market formation in this problem). The average load volumes on lanes sharing the same origin and destination markets are summed up, giving us aggregated lane volumes; assuming demand independence on each individual lane, we also compute the standard deviation of the weekly demand on the aggregated lanes.

Next, we select lanes (lanes, from here onwards, refer to the aggregated lanes between markets) which have an average weekly demand of more than 2 loads, and enumerate all possible two-way and three-way cycles with the following search algorithm:

1. Number each lane in the network in sequence.

2. For each lane i, search through all lanes starting from lane $(i + 1)$. If no lane has the same destination as the origin of lane i, repeat Step 2 with the next lane $(i + 1)$; if the destination market of some lane j is the same as the origin market of lane i, do the following:

   - If the origin market of lane j is the same as the destination market of lane i, lanes i and j form a two-way cycle.

   - If the above is not true, search through all lanes starting from lane $(i + 2)$, if we can find some lane k whose origin overlaps with the destination of lane i and whose destination overlaps with the origin of lane j, then lanes i, j and k form a three-way cycle.

   - Continue with Step 2 if neither of the above two conditions can be satisfied.

After we have enumerated all possible two-way and three-way cycles, the following optimization models are solved, depending on our balance requirement.

43

3.3.1   Formulation 1: Strictly Balanced Weekly Demand

$$\max \quad \sum_{\forall j} a_j V_j$$

$$\text{s.t.} \quad V_j \geq 2Y_j \qquad \forall j \qquad (1)$$

$$V_j \leq M \cdot Y_j \qquad \forall j \qquad (2)$$

$$\sum_{\forall j} V_j \cdot I_{ej} \leq D_e \qquad \forall e \qquad (3)$$

$$V_j \geq 0$$

$$Y_j = \{0, 1\}$$

Decision variables:

V_j: Volume of shipment on each lane of cycle j;

Y_j: 0-1 integer variable; 1 if $V_j > 0$, 0 otherwise.

Parameters:

a_j is the number of lanes on cycle j. a_j is either 2 or 3 in this problem.

D_e is the weekly demand on lane e.

I_{ej} is a 0-1 parameter, indicating whether lane e belongs to S_j (i.e. cycle j).

M is an arbitrary large number which may take the value of the maximum demand on the lanes.

In this model, each lane in a cycle is required to have the same weekly demand. Therefore, we use V_j to denote the volume along the cycle. Constraints (1) & (2) restrict that any cycle that is chosen to be offered as a package bid will have at least 2 weekly demands along the cycle. Constraint (3) implies that the sum of volumes assigned to different cycles from a lane does not exceed the total weekly demand on that lane.

3.3.2 Formulation 2: Almost-Balanced Weekly Demand

$$\text{max} \quad \sum_{\forall j} \sum_{\forall e} V_{ej}$$

$$\text{s.t.} \quad V_{ej} \geq 2I_{ej}Y_j \qquad \forall e\&j \qquad (1)$$

$$V_{ej} \leq D_e I_{ej} Y_j \qquad \forall e\&j \qquad (2)$$

$$\sum_{\forall j} V_{ej} \leq D_e \qquad \forall e \qquad (3)$$

$$V_{e_1,j} \leq 1.087 \times V_{e_2,j} \qquad \forall j\&\{e_1, e_2\} \in S_j \quad (4)$$

$$V_{e_2,j} \leq 1.087 \times V_{e_1,j} \qquad \forall j\&\{e_1, e_2\} \in S_j \quad (5)$$

$$V_{ej} \geq 0$$

$$Y_j = \{0, 1\}$$

Decision variables:

V_{ej}: Volume of shipment on lane e of cycle j;

Y_j: 0-1 integer variable; 1 if cycle j is selected, 0 otherwise.

In this model, weekly demands on lanes of a cycle are allowed to differ by at most 8.7%, according to the 92% compliance requirement (see section 3.2). As a result, the weekly demand on a cycle is "almost-balanced". Constraints (1) & (2) imply that V_{ej} can only be positive given that lane e belongs to cycle j and cycle j is selected as a potential package in the bidding; and when V_{ej} is positive, it should be larger than 2. Constraint (3) implies that the sum of all the split volumes on one lane does not exceed the total weekly demand on that lane. Constraints (4) & (5) impose the less-than-8.7% almost-balanced conditions. The 92% compliance requirement is satisfied from the weekly stand point.

3.3.3   Formulation 3: Strictly Balanced Daily Demand

$$\max \quad \sum_{\forall j} \sum_{\forall i} \sum_{\forall e} V_{eij}$$

$$\text{s.t.} \quad \sum_{\forall i} V_{eij} \geq 2 I_{ej} Y_j \qquad \forall e \& j \qquad (1)$$

$$V_{eij} \leq d_{ei} I_{ej} Y_j \qquad \forall e, i \& j \qquad (2)$$

$$\sum_{\forall j} V_{eij} \leq d_{ei} \qquad \forall e \& i \qquad (3)$$

$$V_{e_1,(i+t_2),j} = V_{e_2,i,j} \qquad \forall j, i \& \{e_1, e_2\} \in S_j \quad (4)$$

$$V_{e_2,(i+t_1),j} = V_{e_1,i,j} \qquad \forall j, i \& \{e_1, e_2\} \in S_j \quad (5)$$

$$V_{eij} \geq 0$$

$$Y_j = \{0, 1\}$$

Decision variables:

V_{eij}: the volume of shipment on lane e of cycle j on the ith day of week.

Y_j: 0-1 integer variable; 1 if cycle j is selected, 0 otherwise.

Parameter $d_{e_1,i}$ represents the demand on the ith day on lane e_1. Also, parameter t_1 or t_2 is the transit time on lane e_1 or e_2.

This model takes the time factor into consideration, as we want daily demand to be balanced. We only have information on the average demand on each day of week, so we will assume that each day of week have a distinct demand pattern. By daily balance, it means that given a two-way cycle, if x loads are shipped from node A on Monday and delivered to node B on Wednesday, we need to have x loads at node B on Wednesday so that the same amount is carried back to A on Friday; at A on Friday, load matching is checked again, and this process continues between the two nodes on different days, depending on the transit

times. The decision variable V_{eij} is a three-dimensional variable, each dimension representing information on lane, cycle, and day of week.

Similar to Formulation 2 above, constraints (1) & (2) imply that V_{eij} can only be positive given that lane e belongs to cycle j and cycle j is selected as a potential package in the bidding; and when V_{eij} is positive, the weekly sum should be larger than 2. Constraint (3) implies that the daily sum of all the split volumes on one lane does not exceed the total daily demand on that lane. Constraints (4) & (5) require the daily volumes on lanes of the same cycle to be balanced, taking transit times into account.

3.3.4   Formulation 4: Almost-Balanced Daily Demand

$$\text{max} \quad \sum_{\forall j} \sum_{\forall i} \sum_{\forall e} V_{eij}$$

$$\text{s.t.} \quad \sum_{\forall i} V_{eij} \geq 2I_{ej}Y_j \qquad \forall e\&j \qquad (1)$$

$$V_{eij} \leq d_{ei}I_{ej}Y_j \qquad \forall e, i\&j \qquad (2)$$

$$\sum_{\forall j} V_{eij} \leq d_{ei} \qquad \forall e\&i \qquad (3)$$

$$V_{e_1,(i+t_2),j} \leq 1.087 \times V_{e_2,i,j} \qquad \forall j, i\&\{e_1, e_2\} \in S_j \quad (4)$$

$$V_{e_2,(i+t_1),j} \leq 1.087 \times V_{e_1,i,j} \qquad \forall j, i\&\{e_1, e_2\} \in S_j \quad (5)$$

$$V_{eij} \geq 0$$

$$Y_j = \{0, 1\}$$

This model requires the daily demand on each cycle to be almost balanced - the differences among lanes of the same cycle do not exceed 8.7%. The formulation is the same as Formulation 3, except for constraints (4) & (5), which impose the almost-balanced conditions and make sure that the 92% compliance rate is fulfilled, i.e. 92% of the loads from B to A can be shipped by drivers who just came from A to B if A and B form a two-way cycle; and

vice versa.

3.4 Computational Results

The optimization models presented in the previous section are solved using *CPLEX Callable Library, version 9.1*.

3.4.1 Illustration of Results using a 3-Market Example

We use the following simple 3-market example to illustrate the results of cycle selection using the four optimization models respectively.
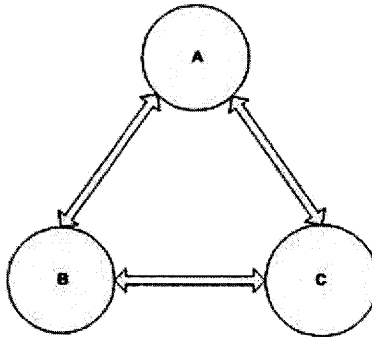
Figure 3-3: Simple 3-market example

In this simple example, there are altogether 6 lanes, connecting each pair of markets in both directions. We can form 3 two-way cycles between each pair of markets and 2 three-way cycles in both A-B-C-A and A-C-B-A directions respectively.

Table 3.2 shows the results from solving both weekly demand balance models. It can be observed that in both solutions, 3 cycles are chosen to be bid out from a total of 5 possible cycles; and the 3 chosen cycles from both models are the same for this small problem - the cycles may not necessarily be the same for other larger networks. However, the objective value (i.e. the total demand covered by the chosen cycles) from Formulation 2 is larger than that from Formulation 1, because there is more flexibility with the almost-balanced conditions in Formulation 2.

48

Table 3.1: Information on the 3-market example

	Lane ID					
	AB	BA	AC	CA	BC	CB
Weekly demand	47.933	2.0008	24.651	9.8769	3.9383	14.585
Std dev	11.081	2.2198	10.847	4.161	3.1263	3.3983
Transit time	2	2	2	2	2	1
Sun	3.7543	0.0785	0.2765	1.1077	0.9931	1.2
Mon	8.8183	0.3531	4.4245	1.8	0.4452	1.8923
Tue	7.7269	0.51	4.8986	2.3077	0.1712	3.1385
Wed	9.3858	0.3531	5.8072	1.3385	0.2397	2.7692
Thu	7.4213	0.3531	5.4911	1.2	0.5479	2.1231
Fri	6.5482	0.1569	3.0023	1.0154	0.6507	2.2154
Sat	4.2782	0.1962	0.7506	1.1077	0.8904	1.2462

Table 3.2: Comparison of results from Formulations 1 and 2 (3-market example)

Formulation 1 Results		Formulation 2 Results			
Cycle selected	Weekly demand	Cycle selected	Weekly demand		
A-C-A	9.8769	A-C-A	10.736(AC)	9.8769(CA)	
B-C-B	3.9383	B-C-B	3.9383(BC)	4.2809(CB)	
A-C-B-A	2.0008	A-C-B-A	2.1748(AC)	2.1748(CB)	2.0008(BA)
Sum	33.633	Sum	35.183		
% covered	32.66%	% covered	34.16%		

There is no solution for Formulation 3, as this example has too little demand on the network and daily balance is very difficult to be fulfilled, unless we relax the requirement to be almost-balanced.

Table 3.3: Results from Formulation 4 (3-market example)

Cycle selected	Lane ID	Daily demand						
		Sun	Mon	Tue	Wed	Thu	Fri	Sat
A-C-A	AC	0.2765	0.3861	0.539	0.7525	0.3267	0.4562	0.6369
	CA	0.4959	0.6923	0.3006	0.4197	0.5859	0.7548	0.3552
A-B-C-A	AB	0.2832	0.6474	0.5505	0.4287	0.3337	0.2599	0.2023
	BC	0.2825	0.2199	0.1712	0.2397	0.5479	0.4659	0.3628
	CA	0.5065	0.3943	0.307	0.2391	0.1861	0.2606	0.5956
Sum		14.463						
% covered		14.04%						

Table 3.3 presents the results from Formulation 4. Formulation 4 requires almost-balanced daily demand, in other words, 92% of the demand on one lane on some day is covered by the demand on the other lane on some other day, taking the transit times into account. It gives slightly more flexibility than Formulation 3; however, the total demand covered by the chosen cycles in this model is less than half of the amount in Formulations 1 and 2, taking up only 14.04% of the demand on the network. In addition, we can see that the average daily demand assigned to each lane of the cycle is very small, implying little practicality in actual operations, as such little average daily demand may be unlikely to balance out when uncertainly is present.

3.4.2 Results from Solving the Large-Scale Industrial Problem

Balanced cycles are selected from company ABC's distribution network using the four optimization models. We first aggregate the shipping locations into markets and find 1718 aggregated lanes, out of which 485 lanes have an average weekly demand exceeding 2 loads. These 485 lanes can form a total of 471 cycles, with 74 two-way cycles and 397 three-way cycles.

Table 3.4: Scales of the four optimization models (large network problem from the company)

Formulation	1	2	3	4
Total No. of decision variables	942	228,906	1,599,516	1,599,516
No. of binary variables	471	471	471	471
No. of constraints	1,427	459,885	1,840,248	1,840,248

Table 3.4 shows the scales of the four optimization models with company ABC's network information. All four models are mixed-integer programs, each having 471 binary decision variables. It can be seen that Formulations 3 and 4 have too many decision variables and constraints for CPLEX to solve. But we know there are actually a lot of redundant variables among all possible V_{eij}'s, as every lane e does not belong to every cycle j, and V_{eij} is redundant if e and j are unrelated. For the purpose of presenting the models clearly, we used the parameter I_{ej} in the constraints of Formulations 3 and 4 to denote whether lane e belongs to cycle j. In the actual coding in CPLEX, we can eliminate all the variables V_{eij} for which lane e does not belong to cycle j, and all the redundant constraints with $I_{ej} = 0$. After removing the redundancies, we only have 9844 decision variables and 23,480 constraints, which makes the problem readily solvable.

Table 3.5: Computational results (large network problem from the company)

Formulation	1	2	3	4
Computation time	15 sec	8 sec	5 sec	10 sec
Number of balanced cycles chosen	144	138	78	90
Identical number of chosen cycles	120		70	
% of entire network covered	18.60%	19.26%	7.95%	10.09%

The computational results for solving the company's large-scale problem are summarized in Table 3.5. Comparing Formulations 1 and 2, Formulations 3 and 4, most of the cycles selected by the two weekly balanced models are identical, and most of the cycles from the two daily balanced models are also the same. The model with almost-balanced conditions always gives slightly more coverage for the network's total demand than the corresponding model with strictly balanced conditions; however, the improvement is not significant (only

51

about 1% more coverage), due to the obligation to satisfy the 92% compliance requirement. It is also found that only a small proportion of the network's demand can be covered by the cycles; especially for the case of daily demand balance (Formulations 3 and 4), only 10% or even less can be covered. As our models are already solved under the simplified assumption that the demand can be deterministic and fractional, the results have indicated that there may not be a lot of opportunities for good, well-balanced package bids.

Chapter 4

Stochastic Analysis

In Chapter 3, we presented deterministic models to solve for the optimal set of cycles that can be bid out. However, in real life, demand variability exists, which may affect the optimality, and even the feasibility of the deterministic solutions, obtained using mean demand. We consider the following two approaches to assign volumes on the lanes to the cycles. The first approach assigns a fixed proportion of the random demand on each lane to the cycles. For example, if we assign 50% of lane 1's demand to cycle A, when the demand on lane 1 is 2 loads on Monday, 1 load is given to the carrier serving cycle A; when the demand becomes 4 loads on Wednesday, we assign 2 loads to the carrier. In the second approach, we may assign variable proportions of the lanes' volumes to the cycles, depending on how the volumes change each time. For example, given cycle A with two lanes 1 and 2, if lane 1 has 5 loads and lane 2 has 2 loads, then we assign 2 loads to cycle A to form a balanced cycle; if lane 1 has 5 loads but lane 2 has 4 loads, then we assign 4 loads to cycle A. We provide carriers with information on the expected amount of volume on each lane, without guaranteeing a fixed proportion of every demand arrival. In this approach, the cycles being offered in the bidding must be fixed; in other words, if cycle X is not chosen to be bid out, even if we get some balanced loads on cycle X some time in the future, we have to assign these loads to carriers serving the individual lanes.

4.1   Fixed Proportion Strategy

4.1.1   Using the Deterministic Models

We use solutions from the deterministic models in Chapter 3 to find the cycles as well as the proportion of demand that can be assigned from the lanes to the cycles. Then, we use simulation to test whether the 92% compliance requirement can be satisfied with these deterministic results.

Simulation Results for Daily Balance

We use the deterministic solutions from Formulation 4 in section 3.3.4 to obtain the cycles and the daily demands on the cycles. Assuming that each day's demand arrival follows a poisson distribution, we simulated 100 days' of demand and assigned every day's demand to the cycles proportionally. The compliance on each cycle was analyzed and a correlation was found between the average daily demand on the cycle and the compliance level.

Figure 4-1 shows a scatter plot of 21 data points from a small network example. It can be observed that the average compliances on the cycles that we obtained in our example are all much lower than 92%. In addition, there is a trend that as the average daily demand on the cycle increases, the corresponding compliance increases as well.

We then perform a regression analysis on the data points, as shown in Figure 4-2. Both linear and logarithmic regressions give R-squared values above 0.9.

The linear regression line is $y = 0.2489x + 0.2029$. By extrapolation, it is found that for the compliance of a cycle to exceed 92%, the average daily demand must be larger than 2.88. Out of a total of 78 balanced cycles in the solution for company ABC's distribution network, only 3 cycles have an average daily demand above 2.88. However, simulation shows that the average compliance on the 3 cycles is at most 75%. Therefore, linear extrapolation is inaccurate.

The logarithmic regression line is $y = 0.19 \ln x + 0.48$. Again by extrapolation, we find that the average demand has to be least 10.2 loads per day for the compliance level to
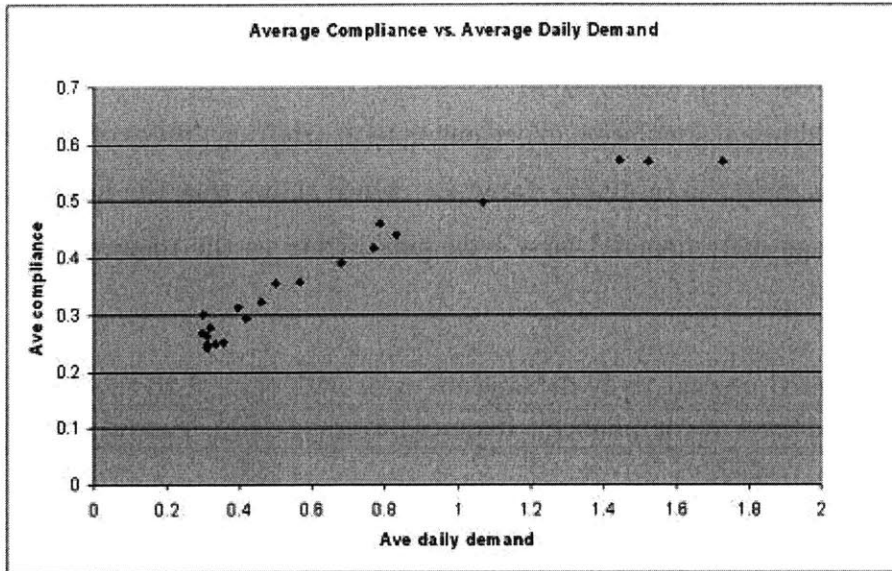
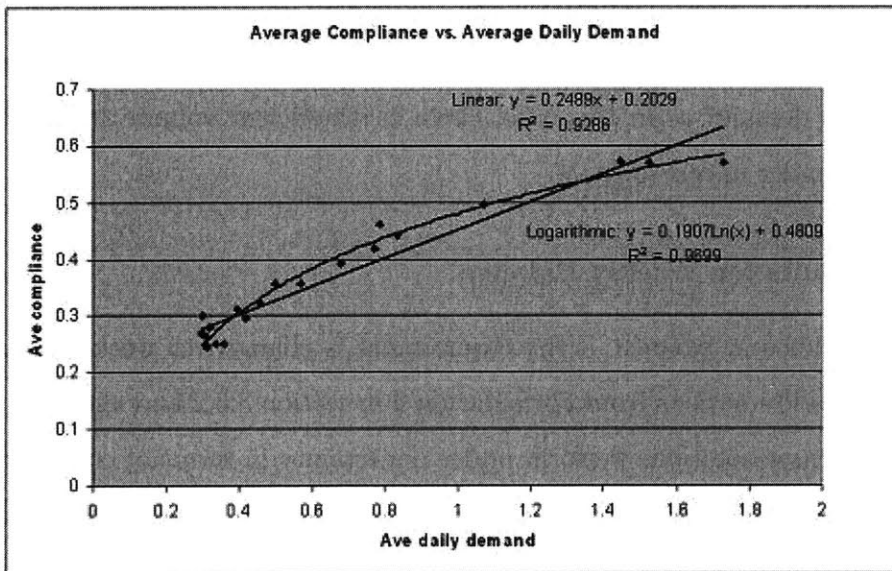Figure 4-1: Scatter plot: average compliance vs. average daily demand

Figure 4-2: Linear regression and logarithmic regression: average compliance vs. average daily demand

reach 0.92. Unfortunately, there are very few lanes in the company's network which have an average of more than 10 loads per day (or 70 loads per week); and no balanced cycle with more than 10 average daily demand can be formed.

We perform additional simulation experiments with arbitrary values of the daily demand. These experiments yield the results in Table 4.1, which shows that the marginal increase in the required average daily demand on a cycle gets higher as the required compliance rate increases.

Table 4.1: Required average daily demand on cycle with respect to the compliance level

Required level of compliance	Required average daily demand on cycle
75%	4.6
85%	16
92%	70

This partly explains why small packages are more likely to fall apart. Services to bundled lanes are more reliable when carriers build densities along the corridors by supplying multiple shippers. As a result, they would have enough aggregated daily demand in their own network to fulfill the required level of compliance. However, for small packages which might rely on a single shipper's demand as in this case, there is insufficient volume to guarantee a high compliance rate under uncertainty.

Simulation Results for Weekly Balance

Instead of daily demand balance, if the requirement is changed to weekly balance, we can use the deterministic solutions from Formulation 2 in section 3.3.2 and again do a simulation to test how well these solutions perform under uncertainty in terms of compliance rate.

Assuming that the weekly demand on each lane follows a normal distribution with the given mean and standard deviation, we simulated 100 weeks' demand with a small network example. Then we calculated the proportions from the deterministic solutions and assigned demand on each lane to the corresponding cycles proportionally.

The simulation results have shown that the demand on one lane can usually cover 70%

to 80% of the demand on the other lane in the same cycle, but not as high as 92%. It is also found that the compliance on one lane is negatively correlated to the coefficients of variation of demand on this lane and the previous lane in the cycle.

A multiple regression is done with 24 data points. The regression line obtained is $y = 0.96 - 0.28x_1 - 0.06x_2$ where y is the compliance, x_1 and x_2 are the coefficients of variation of demand on the previous and current lanes respectively.

From the regression line, it is computed that the desired coefficients of variation need to be smaller than 0.12, in order for the compliance on all lanes to get above 92%. Unfortunately again, among company ABC's distribution network, no cycle can be formed with lanes having coefficients of variation smaller than 0.12.

In summary, simulation experiments have shown that the results from the deterministic formulations are not able to ensure a 92% compliance when demand is uncertain, both from the daily and the weekly perspectives. For the deterministic solutions to work well under uncertainty, we either need much larger volumes on the lanes or smaller coefficients of variation. As analyzed in section 2.3.3, lanes with greater volumes tend to have smaller coefficients of variation; therefore, large volume is the most important condition for the deterministic solutions to remain reliable in real life applications - the volumes supplied by one single company usually do not suffice.

4.1.2 Using Probabilistic Analysis

In this section, we conduct a probabilistic analysis to investigate if reliable packages can be built when demand is stochastic. We modify our deterministic formulations in Chapter 3 into the following:

$$\max \quad \sum_{\forall j} \sum_{\forall e} f_{ej} E(D_e)$$

$$\text{s.t.} \quad f_{ej} E(D_e) \geq 2 I_{ej} Y_j \qquad \forall e \& j \qquad (1)$$

$$f_{ej} \leq I_{ej} Y_j \qquad \forall e \& j \qquad (2)$$

$$\sum_{\forall j} f_{ej} \leq 1 \qquad \forall e \qquad (3)$$

$$P(0.92 f_{e_2 j} D_{e_2} \leq f_{e_1 j} D_{e_1} \leq 1.087 f_{e_2 j} D_{e_2}) \geq 0.9 \qquad \forall j, i \& \{e_1, e_2\} \in S_j \qquad (4)$$

$$0 \leq f_{ej} \leq 1$$

$$Y_j = \{0, 1\}$$

We use the decision variable f_{ej} to represent the fraction of volume on lane e being assigned to cycle j, as we are allocating a fixed proportion of the lanes' volumes to the cycles in this approach.

D_e is a random variable of the weekly demand on lane e; and $E(D_e)$ is the expectation of the weekly demand.

Constraints (1), (2), and (3) are direct modifications from the deterministic formulations. In constraint (4), we require the assigned volume from lane e_1 to cycle j to fall within some range of the assigned volume from lane e_2, i.e. the 92% compliance rate is fulfilled, with a high probability (set as 0.9 in this formulation). We want to find the ratio of $\frac{f_{e_1 j}}{f_{e_2 j}}$ so that constraint (4) can be satisfied, after which the optimization model can be solved.

The above formulation models a weekly balance case. If we want to have daily demand balances, we can just replace the random variable D_e by d_e, random variable of the daily demand, assuming for simplicity that each day's demand follows the same distribution.

To find the ratio of $\frac{f_{e_1 j}}{f_{e_2 j}}$, we have to approximate D_{e_1} and D_{e_2} to χ^2 distributions. Then

the probability expression in constraint (4) can be rewritten as

$$P(0.92 f_{e_2 j} D_{e_2} \leq f_{e_1 j} D_{e_1} \leq 1.087 f_{e_2 j} D_{e_2}) = P(\alpha r \leq \frac{D_{e_1}/\upsilon_{D_{e_1}}}{D_{e_2}/\upsilon_{D_{e_2}}} \leq \beta r)$$

where $\upsilon_{D_{e_1}}$ and $\upsilon_{D_{e_2}}$ are the degrees of freedom for the χ^2 distributions of D_{e_1} and D_{e_2} respectively; $\alpha = 0.92 \frac{\upsilon_{D_{e_2}}}{\upsilon_{D_{e_1}}}$, $\beta = 1.087 \frac{\upsilon_{D_{e_2}}}{\upsilon_{D_{e_1}}}$, and $r = \frac{f_{e_2 j}}{f_{e_1 j}}$

As D_{e_1} and D_{e_2} follow χ^2 distributions, $\frac{D_{e_1}/\upsilon_{D_{e_1}}}{D_{e_2}/\upsilon_{D_{e_2}}}$ follows an F-distribution with degrees of freedom $\upsilon_{D_{e_1}}$, $\upsilon_{D_{e_2}}$. We want to find the ratio r such that $F(\beta r) - F(\alpha r) \geq 0.9$ under the F-distribution.

However, there is no solution of r which satisfies the above probability inequality. In fact, even when the probability criteria is lowered to 0.5, i.e. $P(0.92 f_{e_2 j} D_{e_2} \leq f_{e_1 j} D_{e_1} \leq 1.087 f_{e_2 j} D_{e_2}) \geq 0.5$, there is no suitable ratio r such that $F(\beta r) - F(\alpha r)$ is above the criteria.

It is also found that we can only obtain solutions for r satisfying $F(\beta r) - F(\alpha r) \geq 0.9$ when the degrees of freedom in the F-distribution are very big, approximately 300 on average.

Having very big degrees of freedom for the F-distribution implies that the χ^2 random variables D_{e_1} and D_{e_2} must have very big degrees of freedom as well. The degree of freedom in the χ^2 distribution is equal to the mean. Therefore, it is also implied that we need to have a very high expected demand in order to obtain feasible solutions for our probability inequality as well as the optimization formulation on page 58.

In short, both methods in this section have shown that if we want to assign a fixed proportion of the lane's demand to the cycles, we will not be successful in forming reliable cycles that stay balanced most of the time unless we are able to accumulate a large amount of volumes on the lanes, which may be achieved by combining the volumes from several shippers. That is why carriers serving multiple shippers usually have a higher compliance rate, as they are able to build high densities along the corridors.

4.2 Variable Proportion Strategy

The previous section gave a stochastic analysis on the creation of balanced cycles in company ABC's network and showed that there is insufficient volume in the company's network if we want to assign a fixed proportion of the random demand on the lanes to the cycles. In this section, we try to explore the possibility of forming cycles under uncertainty by assigning variable proportions of the lanes' volumes to the cycles, as explained in the beginning of this chapter.

We only consider weekly balance in this section, due to data limitations. Moreover, the daily balance conditions are too restrictive with the given network volumes. More than 70% of the lanes out of the company's entire network has an average daily demand of less than 1 load, which implies high variability. The definition of compliance here requires lanes in a cycle to have almost the same daily demands, and drivers do not wait longer than one day. This makes the 92% compliance rate very hard to achieve under high daily demand variability.

4.2.1 Finding a Lower Bound Coverage

Assuming that the weekly demand follows a normal distribution with the given mean (μ) and standard deviation (σ), we compute the lower bound values for the demand on each lane as $Max(0, \mu - 2\sigma)$. Using these lower bound values, we solve Formulation 2 in section 3.3.2 to find the lower bound coverage of the network.

As a result, we find 7 two-way cycles and 20 three-way cycles in company ABC's network with the lower bound volumes; they cover 5.3% of the entire network's demand. Therefore, the most conservative way is to bid out these 27 cycles and fix the volumes on the cycles as the lower-bound solutions. We are almost guaranteed (more than 98% of the time) that the lane volumes that are offered in these package bids will be realized; and the 92% compliance requirement is satisified within each cycle.

60

4.2.2 Using Stochastic Optimization

The method presented in section 4.2.1 only provides a lower bound solution; to find the optimal set of cycles that can be bid out under demand uncertainty, we propose a stochastic optimization formulation. Decision variables Y_j's are used as first stage variables where we determine whether each possible cycle will be selected to be bid out; while $V_{ej\omega}$'s are second stage variables that determine the volume allocated from every lane to the corresponding cycle under each scenario ω.

$$\max \quad \sum_{\forall \omega} \alpha_\omega \sum_{\forall j} \sum_{\forall e} V_{ej\omega}$$

$$\text{s.t.} \quad \sum_{\forall \omega} \alpha_\omega V_{ej\omega} \geq 2 I_{ej} Y_j \quad \forall e, j \quad (1)$$

$$V_{ej\omega} \leq M I_{ej} Y_j \quad \forall e, j, \omega \quad (2)$$

$$\sum_{\forall j} V_{ej\omega} \leq D_{e\omega} \quad \forall e, \omega \quad (3)$$

$$V_{e_1, j, \omega} \leq 1.087 \times V_{e_2, j, \omega} \quad \forall j, \omega, \& \{e_1, e_2\} \in S_j \quad (4)$$

$$V_{e_2, j, \omega} \leq 1.087 \times V_{e_1, j, \omega} \quad \forall j, \omega, \& \{e_1, e_2\} \in S_j \quad (5)$$

$$V_{ej\omega} \geq 0$$

$$Y_j = \{0, 1\}$$

The above formulation also has the following parameters:

α_ω: The probability that scenario ω occurs;

$D_{e\omega}$: The weekly demand on lane e under scenario ω;

I_{ej}: 0-1 parameter, indicating whether lane e belongs to cycle j;

M: An arbitrary large number.

This formulation is also similar to the deterministic formulations in Chapter 3. Constraint

(1) implies that if cycle j is selected as a cycle to be bid out, the expected amount of volume assigned to each lane of this cycle is larger than 2; whereas constraint (2) implies that if cycle j is not selected as a cycle to be bid out, the volume assigned onto this cycle should be 0 in any scenario. The sum of all the split volumes on one lane should not exceed the total weekly demand for that lane in all scenarios, as indicated by constraint (3). Finally, constraints (4) & (5) require the weekly volumes for lanes of the same cycle to differ by less than 8.7% in every scenario. In other words, the weekly volumes are almost balanced for lanes belonging to the same cycle, satisfying the 92% compliance requirement from the weekly stand point.

The major challenge in solving the above stochastic program exists with the total number of scenarios. If we assume that the weekly demand on each lane takes only 2 values - one high and one low, the total number of possible scenarios is $2^{number\_of\_lanes}$, which grows exponentially with the number of lanes in the network. If we use more than 2 possible values for the demand on each lane, then the total number of scenarios is even greater. As the number of variables and constraints in our stochastic formulation is positively related to the number of scenarios, the problem becomes too big to be solvable for large networks.

The 3-market simple example given in section 3.4.1 is solved using this stochastic mixed-integer program. The demand on each lane takes one high value and one low value with a probability of 0.5 each; so the total number of scenarios is $2^6 = 64$, resulting in 773 variables (out of which 5 are binary variables) and 1182 constraints.

Table 4.2: Results from the stochastic optimization model (3-market example)

Cycle selected	Expected weekly demand on lanes		
A-C-A	10.709(AC)	9.8769(CA)	
B-C-B	3.9383(BC)	4.2809(CB)	
A-C-B-A	2.171(AC)	2.1748(CB)	2.0008(BA)
Sum	35.152		
% covered	34.13%		

Table 4.2 summarizes the results. As compared to the deterministic solution in Table 3.2,

the stochastic model recommends the same cycles to be bid out; but the expected amount of volume that can be covered with these cycles is slightly less than the deterministic solution.

Company ABC's entire distribution network cannot be solved by this stochastic optimization model, due to an overflowing number of scenarios as explained earlier on. However, when we enumerate the total number of scenarios, we are actually assuming that volumes on the lanes are independent. This may not be true, because volumes on the lanes can be dependent on factors such as seasonality as well as regions where the lanes belong to. For example, if the production at a certain major plant decreases due to fewer orders in winter, all lanes originating from that plant have reduced demand; the scenario that some lanes have high demand while some others have low demand is not likely to occur. If provided with more historical data, we can do some analysis on the correlation between the volumes on different lanes and reduce the total number of scenarios effectively.

Another approximated solution is to fix the values of the first stage variables Y_j using the cycles chosen from the deterministic model. Then every week when there is a new scenario (i.e. new demand on the lanes), we solve for the following linear program to allocate volumes onto each lane of the cycles.

$$\max \quad \sum_{\forall j} \sum_{\forall e} V_{ej}$$

$$\text{s.t.} \quad V_{ej} \leq D_e I_{ej} Y_j \qquad \forall e \& j \qquad (1)$$

$$\sum_{\forall j} V_{ej} \leq D_e \qquad \forall e \qquad (2)$$

$$V_{e_1,j} \leq 1.087 \times V_{e_2,j} \qquad \forall j \& \{e_1, e_2\} \in S_j \quad (3)$$

$$V_{e_2,j} \leq 1.087 \times V_{e_1,j} \qquad \forall j \& \{e_1, e_2\} \in S_j \quad (4)$$

$$V_{ej} \geq 0$$

This model above modifies Formulation 2 in section 3.3.2 by eliminating the decision

variables Y_j and replacing with pre-determined values of 0 or 1. Moreover, the first constraint in the original formulation is removed because we only want the expected volume on each lane to be above 2 loads, not necessarily in every scenario.

This linear program can be solved very fast. Experiments with 20 sets of different demands within the allowable range have shown that the total volume covered by the cycles under each scenario do not differ by more than 15% from the deterministic solution. And the optimization model itself makes sure that the 92% compliance is satisfied in each scenario.

In this section, we presented a stochastic programming formulation to find the best strategy of creating package bids and allocating volumes to the cycles as demand changes. However, this optimization model is too complex to be solved for the company's large network. Therefore, we used a lower-bound method to find the minimum amount of volumes that can be shipped with balanced cycles. We also proposed to use the deterministic model with mean values to obtain an approximated solution for the cycles which can be bid out. In this approach, we assign variable proportions of the lanes' volumes to the cycles according to how demand changes every time, ensuring that the 92% compliance requirement is fulfilled in each allocation.

Chapter 5

Conclusion

5.1 Summary

In this thesis, we have incorporated optimization models to help companies create package bids in their transportation procurement. The focus of this thesis to form two-way or three-way cycles out of the company's own distribution network, which can be offered to the carriers in the bidding. Mixed-integer programs were developed to solve the problem deterministically. However, as demand is stochastic in real life, we also presented several methods to analyze how the cycles should be created and demand allocated under uncertainty, including simulation, probabilistic analysis and stochastic optimization.

As we need balanced cycles to reap the benefits of having package bids, we developed models for both the weekly balanced and the daily balanced case. Results from the deterministic models as well as simulation have shown that there are more opportunities of forming cycles if we only enforce the weekly balanced conditions. The compliance requirement defined for the daily balance case is found to be too strict for reliable cycles to be constructed, especially since the variability of daily demand is quite high.

When creating package bids, it is easier in practice if we always allocate a fixed proportion of the lanes' volumes to the cycles where they belong to. However, both simulation experiments and probabilistic analysis have shown that this approach of assigning loads is

only feasible if we have a large amount of volumes on the lanes, which is not possible with one single company's shipment. Therefore, joint bids between non-competitive companies should still be encouraged for the benefit of having more well-balanced cycles.

We then proposed another approach to allocate different proportions of the lanes' volumes to the cycles according to the actual demand each time. We computed a lower bound solution to obtain the minimum amount of loads that can be shipped along balanced cycles. However, as the lower bound solution is too conservative, we suggested another stochastic programming model to attain the optimal set of cycles that can be offered under this approach. As solving this stochastic program for large networks is a big challenge, we presented an approximated solution method instead. Using this approach, we can offer at least 5% of the total demand on the network as package bids, and the approximated optimal solution can cover about 15% to 20% of the total demand.

5.2   Future Work Extensions

It was already mentioned that solving the stochastic programming formulation in section 4.2.2 for large networks is a challenging task. Solution techniques to solve such large-scaled problems can be very helpful. Alternatively, future attention may be devoted to modify the stochastic formulation so that it is tractable for large-scaled problems.

More understanding of the correlation of demands between different lanes is required, so that the investigation on forming well-balanced cycles under demand uncertainty can be done more accurately. With the correlation information, we may be able to reduce the size of the stochastic programming formulation by having fewer scenarios. Therefore, a robust demand forecasting methodology is going to be very useful in this problem.

In addition, the idea of regional directed fleet, briefly introduced in section 2.1.2, is also a major area for future work. We foresee opportunities in creating package bids with short tours originating from the same regional centers.

Bibliography

[1] Standard & Poor's Corporation. Industry surveys, Transportation: Commercial. Feb 9, 2006.

[2] Yossi Sheffi. Combinatorial auctions in the procurement of transportation services. *Interfaces*, 34(4):245–252, 2004.

[3] Christopher G. Caplice. *An Optimization Based Bidding Process: A New Framework for Shipper-Carrier Relationships*. PhD thesis, Massachusetts Institute of Technology, 1996.

[4] Jiongjiong Song and Amelia Regan. Combinatorial auctions for transportation service procurement: The carrier perspective. *Transportation Research Record*, 1833:40–46, 2003.

[5] American Trucking Associations. *American Trucking Trends: 2004*.

[6] Sergio R. Jara Diaz. *Transportation Cost Functions: A Multiproduct Approach*. PhD thesis, Massachusetts Institute of Technology, 1981.

[7] Sergio R. Jara Diaz. Freight transportation multioutput analysis. *Transportation Research*, 17A(6):429–438, 1983.

[8] Curtis M. Grimm, Thomas M. Corsi, and Judith L. Jarrell. U.S. motor carrier cost structure under deregulation. *The Logistics and Transportation Review*, 25:231–249, 1989.

[9] Chris Caplice and Yossi Sheffi. Optimization-based procurement for transportation services. *Journal of Business Logistics*, 24(2):109–127, 2003.

[10] The American Transportation Research Institute. Top industry issues 2006. Oct 2006. $http : //atri - online.org/research/results/top\_industry\_issues\_2006.pdf$.

[11] Global Insight. The U.S. truck driver shortage: Analysis and forecasts. May 2005.

[12] Michael J. Mulqueen. *Creating Transportation Policy in a Network that utilizes both Contract Carriers and an Internally Managed Fleet*. Master thesis, Massachusetts Institute of Technology, 2006.

[13] Edward J. Bardi, Prabir K. Bagchi, and T. S. Raghunathan. Motor carrier selection in a deregulated environment. *Transportation Journal*, 29(1):4–11, 1989.

[14] Roger Dale Abshire and Shane R. Premeaux. Motor carrier selection criteria: Perceptual differences between shippers and carriers. *Transportation Journal*, 31(3):31–35, 1991.

[15] Brian J. Gibson, Harry L. Sink, and Ray A. Mundy. Shipper-carrier relationships and carrier selection criteria. *Logistics and Transportation Review*, 29(4):371–382, 1993.

[16] E. William Moore, Janice M. Warmke, and Lonny R. Gorban. The indispensable role of management science in centralizing freight operations at reynolds metals company. *Interfaces*, 21(1):107–129, 1991.

[17] Matthew J. Harding. *Can Shippers and Carriers Benefit from More Robust Transportation Planning Methodologies?* Master thesis, Massachusetts Institute of Technology, 2005.

[18] Warren B. Powell, Yosef Sheffi, Kenneth S. Nickerson, Kevin Butterbaugh, and Susan Atherton. Maximizing profits for north american van lines' truckload division: A new framework for pricing and operations. *Interfaces*, 18(1):21–41, 1988.